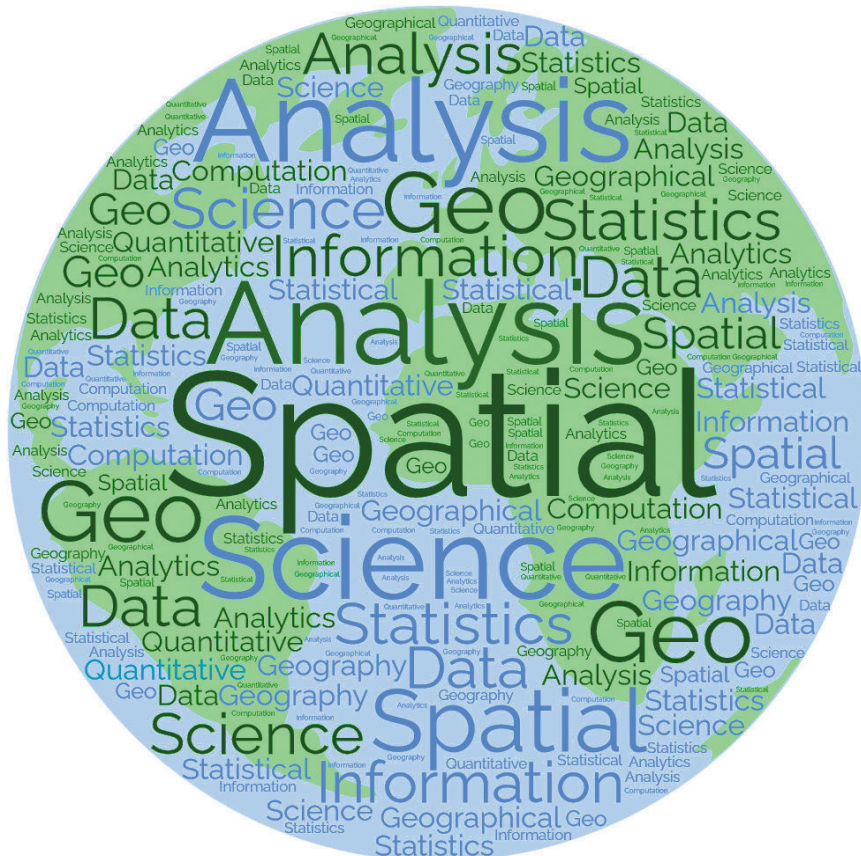


Stefania Bertazzon

L'ANALISI SPAZIALE

La Geografia che... conta



FrancoAngeli 

Nuove Geografie. Strumenti di lavoro

Nuove Geografie. Strumenti di lavoro

Collana diretta da Marina Bertoncin (Università di Padova)

Coordinamento del Comitato scientifico: Andrea Pase (Università di Padova)

Comitato scientifico:

Tiziana Banini (Sapienza Università di Roma), Raffaele Cattedra (Università di Cagliari), Egidio Dansero (Università di Torino), Elena Dell’Agnese (Università di Milano Bicocca), Giulia De Spuches (Università di Palermo), Floriana Galluccio (Università di Napoli L’Orientale), Francesca Governa (Politecnico di Torino), Mirella Loda (Università di Firenze), Claudio Minca (Università di Bologna), Paola Minoia (Università di Torino), Davide Papotti (Università di Parma)

Questa collana intende proporre esplorazioni sul terreno dei nuovi modi di rappresentare, studiare e discutere il territorio. Nuovi modi perché gli oggetti della ricerca geografica cambiano: mutano gli assetti territoriali, si affacciano altri attori, si identificano tematiche inedite o interpretate con inedite sensibilità.

Il nuovo richiede superamento.

Da un lato come capacità di oltrepassare i limiti disciplinari per collegarsi a quanto si sta elaborando nelle scienze vicine e che utilmente possiamo incrociare. Dall’altro come disponibilità ad andare oltre le più consolidate costruzioni teoriche che la disciplina ha sinora prodotto per saggiare ipotesi diverse.

Il nuovo richiede aderenza.

Aderenza al lavoro di terreno, all’indagine di campo, all’ascolto del territorio e delle soggettività che in esso si esprimono. Aderenza al rigore metodologico, da unire al gusto per la sperimentazione.

La collana proporrà strumenti di lavoro, perché nuove geografie chiedono sguardi diversi rispetto a quelli praticati sinora. Senza alcuna pretesa di esaustività e senza alcun accantonamento del lavoro compiuto sinora dalla geografia. Piuttosto, appunto, con la disponibilità a praticare, a maneggiare nuovi attrezzi, sapendo che il lavoro è in corso e che a loro volta queste nuove geografie domani verranno superate.

La collana si rivolge in primo luogo alla comunità dei geografi e ai colleghi di altre discipline interessati al territorio, ma ha l’obiettivo di allargare la platea degli interessati a questi nuovi “discorsi sul mondo”. Un’attenzione particolare verrà data al linguaggio, per contaminarlo con apporti differenti e per renderlo fruibile ad occhi diversi e non solo agli “esperti”.

I testi da pubblicare sono sottoposti a un doppio referaggio, al fine di certificare la qualità del prodotto e la sua congruenza agli obiettivi della collana. Il referaggio è inteso come un momento di crescita e di ulteriore sviluppo del lavoro scientifico e non come una mera attività di valutazione.



Il presente volume è pubblicato in open access, ossia il file dell'intero lavoro è liberamente scaricabile dalla piattaforma **FrancoAngeli Open Access** (<http://bit.ly/francoangeli-oa>).

FrancoAngeli Open Access è la piattaforma per pubblicare articoli e monografie, rispettando gli standard etici e qualitativi e la messa a disposizione dei contenuti ad accesso aperto. Oltre a garantire il deposito nei maggiori archivi e repository internazionali OA, la sua integrazione con tutto il ricco catalogo di riviste e collane FrancoAngeli massimizza la visibilità, favorisce facilità di ricerca per l'utente e possibilità di impatto per l'autore.

Per saperne di più:

http://www.francoangeli.it/come_publicare/publicare_19.asp

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati possono consultare il nostro sito Internet: www.francoangeli.it e iscriversi nella home page al servizio "Informatemi" per ricevere via e-mail le segnalazioni delle novità.

Stefania Bertazzon

L'ANALISI SPAZIALE

La Geografia che... conta

FrancoAngeli

OPEN  ACCESS

Nuove Geografie. Strumenti di lavoro

Isbn 9788835139461

Copyright © 2022 by FrancoAngeli s.r.l., Milano, Italy.

Pubblicato con licenza *Creative Commons Attribuzione-Non Commerciale-Non opere derivate*
4.0 Internazionale (CC-BY-NC-ND 4.0)

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore. L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito

<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.it>

Indice

Introduzione: perché questo libro?	pag. 13
---	---------

Parte I Origini e metodi

1. Le origini	» 21
1.1. Premessa	» 21
1.2. L'analisi spaziale all'ombra del GIS	» 23
1.3. Le radici quantitative	» 30
1.4. Radici interdisciplinari: statistica ed economia	» 35
1.5. Le tecnologie e gli strumenti dell'informazione geografica	» 41
1.6. L'analisi spaziale nella geografia	» 45
2. Oggetto e metodi dell'analisi spaziale: cenni	» 49
2.1. Premessa	» 49
2.2. La "prima legge della geografia"	» 50
2.2.1. I processi spaziali	» 51
2.2.2. L'autocorrelazione spaziale	» 52
2.2.3. La matrice dei pesi spaziali	» 53
2.2.4. L'eterogeneità spaziale	» 54
2.2.5. Relazione tra autocorrelazione ed eterogeneità spaziale	» 55

2.2.6. Il problema dell'unità areale modificabile (<i>MAUP</i>)	pag.	56
2.3. L'analisi di regressione e gli effetti spaziali	»	57
2.3.1. Il modello di regressione	»	58
2.3.2. Le ipotesi del modello di regressione e i processi spaziali	»	59
2.3.3. Effetti spaziali e inaffidabilità delle stime	»	60
2.3.4. Effetti spaziali e il problema dell'inferenza	»	61
2.4. Effetti spaziali e metodi analitici spaziali	»	62
2.4.1. Il modello <i>SAR</i> in risposta all'autocorrelazione spaziale	»	62
2.4.2. Il modello <i>GWR</i> in risposta all'eterogeneità spaziale	»	64
2.5. Definizioni matematiche e statistiche	»	65
2.5.1. I processi spaziali	»	65
2.5.2. L'autocorrelazione spaziale	»	67
2.5.3. La matrice dei pesi spaziali	»	68
2.5.4. L'eterogeneità (non-stazionarietà) spaziale	»	68
2.5.5. Relazione tra autocorrelazione ed eterogeneità spaziale	»	70
2.5.6. Il problema dell'unità areale modificabile (<i>MAUP</i>)	»	71
2.5.7. Il modello di regressione	»	71
2.5.8. Le ipotesi del modello di regressione e i processi spaziali	»	73
2.5.9. Effetti spaziali e inaffidabilità delle stime	»	74
2.5.10. Effetti spaziali e il problema dell'inferenza	»	75
2.5.11. Il modello <i>SAR</i> in risposta all'autocorrelazione spaziale	»	75
2.5.12. Il modello <i>GWR</i> in risposta all'eterogeneità spaziale	»	78
2.6. Oltre i metodi quantitativi: i limiti delle soluzioni parziali	»	79

3. Oltre i metodi quantitativi: relatività dell'analisi spaziale	pag.	83
3.1. Premessa	»	83
3.2. Dalla regionalizzazione all'aggregazione spaziale	»	84
3.3. La dipendenza spaziale e le leggi della geografia	»	91
3.4. Le due facce della medaglia: autocorrelazione ed eterogeneità spaziale	»	96
3.5. Lo spazio dell'analisi spaziale	»	102
3.6. Dal relativismo la nuova analisi spaziale	»	105

Parte II

Definizione, identità, autocritica

4. Verso una definizione di analisi spaziale	»	111
4.1. Premessa	»	111
4.2. Che cos'è la [moderna] geografia quantitativa?	»	112
4.3. Che cos'è [e che cosa non è] il GIS?	»	117
4.4. E la scienza dell'informazione geografica?	»	119
4.5. Che cos'è dunque l'analisi spaziale?	»	128
4.6. Il nome dell'analisi spaziale	»	134
4.7. Il cammino dell'analisi spaziale	»	144
5. La critica e le sfide del presente e del futuro	»	149
5.1. Premessa	»	149
5.2. Dalla critica all'autocritica	»	151
5.2.1. <i>Critical GIScience</i>	»	153
5.2.2. <i>Critical spatial data science, critical spatial data analytics</i>	»	157
5.2.3. <i>Critical spatial thinking e critical mapping</i>	»	160
5.3. L'analisi spaziale e la sfida dei dati	»	162
5.3.1. <i>La geoetica</i>	»	164
5.3.2. <i>Big data, proprietà stocastiche e statistica spaziale</i>	»	168

5.3.3. <i>Data driven, analytics, data science:</i> quale analisi?	pag. 171
5.3.4. Molteplicità di paradigmi e metodi	» 175
5.4. Le nuove frontiere dell'analisi spaziale	» 179

Parte III

La pratica

6. La pratica: dati, attrezzi, personale	» 183
6.1. Premessa	» 183
6.2. Dati, grandi dati, dati spaziali, analisi dei dati	» 185
6.2.1. I dati spaziali e le loro caratteristiche	» 188
6.2.2. Dati spaziali, processi stocastici, errore e incertezza	» 193
6.3. La cassetta degli attrezzi dell'analisi spaziale	» 197
6.3.1. Fabbricare gli attrezzi dell'analisi spaziale	» 200
6.3.2. Dall'elitismo al populismo: gli at- trezzi come strumento per diffondere l'analisi spaziale	» 204
6.4. <i>Chi è speciale</i> nell'analisi spaziale	» 208
6.4.1. Insegnare l'analisi spaziale	» 210
6.4.2. Potenzialità e limiti di chi fa analisi spaziale	» 213
6.5. Verso la pratica	» 215
7. L'analisi spaziale in azione: qualche esempio	» 217
7.1. Premessa	» 217
7.2. Malattie cardiovascolari in contesto urbano	» 219
7.3. Analisi spaziale dell'inquinamento atmosferico	» 223
7.4. Inquinamento da idrocarburi e attività ma- rittime lungo la costa Pacifica canadese	» 226
7.5. Analisi spaziale del rischio ecologico nella laguna di Venezia	» 229

7.6. Analisi spaziotemporale della mortalità nella prima ondata di COVID-19	pag. 232
7.7. Morale delle storie	» 236
8. L'analisi spaziale in Italia	» 239
8.1. Premessa	» 239
8.2. L'analisi spaziale tra il GIS e la geografia italiana	» 240
8.3. La letteratura in lingua italiana	» 242
8.3. L'analisi spaziale nella geografia italiana	» 246
Verso una conclusione: la geografia che conta	» 247
Bibliografia	» 253
Ringraziamenti	» 273

La *filosofia* è scritta in questo grandissimo libro che continuamente ci sta aperto dinnanzi agli occhi (io dico l'universo), ma che non si può intendere se prima non s'impara a intender la lingua, e a conoscer i caratteri, ne' quali è scritto. Egli è scritto in lingua matematica, e i caratteri son triangoli, cerchi, ed altre figure geometriche, senza i quali è impossibile a intenderne umanamente parola, senza i quali è un aggirarsi vanamente per un oscuro laberinto.

Galilei, *Il Saggiatore*, 1977, p. 33.

Dedico questo libro ai miei genitori e a quanti mi hanno insegnato, anche solo un poco, a conoscere il mondo.

Introduzione: perché questo libro?

People die each year because no one bothers to properly analyse disease or traffic accident data or crime data or illness data for unusual localised concentrations that may be preventable once known to exist, predictable and thus avoidable.

Openshaw, *Some Trends and Future Perspectives for Spatial Analysis in GIS*, 1998, p. 7.

Il giovane secolo che stiamo vivendo ha portato sviluppi senza precedenti non solo nella tecnologia, ma ancor più nella sua diffusione, che oramai pervade ogni livello della società, della politica, dell'economia, della ricerca. Questa diffusione della tecnologia, con tutto quanto vi ruota attorno, somiglia ad una valanga, che nel proprio cammino cresce a vista d'occhio, travolgendo e inglobando in sé tutto ciò che incontra. Rimanere fuori (o uscire) dalla valanga tecnologica è sempre più difficile, e se da un lato ci offre possibilità impensabili anche in tempi recenti, dall'altro si fa strada in ciascuno di noi la consapevolezza che la valanga minaccia di travolgere la nostra autodeterminazione, le nostre democrazie, le nostre relazioni e il nostro tempo libero. In parte, la tecnologia erode tanta parte della nostra sfera personale perché molto di ciò che prima si svolgeva senza mediazioni ora ha luogo per mezzo della tecnologia, cosicché ciò che avviene lascia una traccia, formata da ciò che chiamiamo dati, gran parte dei quali sono corredati da informazioni localizzative o spaziali o, con termine più tecnico, georeferenziate.

Frattanto, contemporaneamente a quanto si configura come una rivoluzione tecnologica, questo inizio di secolo è pure costellato di fenomeni problematici, taluni senza precedenti, altri non nuovi, ma tuttavia preoccupanti a causa delle loro dimensioni crescenti. Pensiamo al cambiamento climatico, alle migrazioni di massa, all'inquinamento, alla produzione alimentare, e ultima, ma solo in senso cronologico, alla pandemia del nuovo coronavirus (COVID-19). Anche tutti questi fenomeni, questi problemi alla ricerca di soluzioni, sono localizzati nello spazio geografico e si manifestano con intensità variabile e

caratteristiche locali, interagendo con altri fenomeni su scale diverse. Spesso tali problemi vengono affrontati da chi vi è preposto con gli strumenti a sua disposizione, o che trova più congeniali, benché non sempre quegli strumenti si dimostrino i più efficaci. Per questo si può, e forse addirittura si deve pensare fuori dagli schemi consueti, e proporre un'altra via, che è nel contempo una tentazione e una sfida: quella di pensare i problemi in maniera spaziale, di credere che i problemi localizzati si possano analizzare secondo la loro posizione nello spazio, utilizzando i dati georeferenziati forniti dalla tecnologia, per trovare delle soluzioni locali e a scale diverse. Si tratta di una sfida importante, perché formulare i problemi in maniera spaziale può fornire soluzioni che da altre prospettive non sarebbero visibili, proponendo risposte innovative e aprendo nuove direzioni di ricerca.

L'analisi spaziale è forse la disciplina più preparata a cogliere questa sfida con competenza. Per formulare analisi rilevanti e trovare soluzioni efficaci serve, ma non basta, la disposizione a pensare in maniera spaziale, né bastano i dati georeferenziati: occorre un impianto teorico e concettuale, occorrono strumenti quantitativi supportati da sistemi tecnologici, occorre la consapevolezza delle questioni etiche che sono implicite nell'analisi dei dati, e occorre la volontà di affrontarle compiutamente, assieme ad una riflessione critica, in un dialogo costruttivo con altre discipline, dentro e fuori la geografia. Rispetto a tutto questo l'analisi spaziale possiede poderosi punti di forza ma anche debolezze di cui è cosciente e che ha iniziato ad affrontare non da sola, ma con il contributo dell'intera geografia. Certo non è l'unica disciplina capace di analisi quantitative, ma, diversamente da altre, alle competenze analitiche affianca la capacità di dialogare efficacemente con altri campi disciplinari e applicativi dentro e fuori la geografia, proprio perché sa integrarne le questioni in chiave analitica spaziale. Questa forza si fonda su solide basi teoriche e statistico-matematiche, che sono proprio ciò che le consente un dialogo ampio, ma che pure le conferiscono un'immagine di disciplina ostica, arida, per soli iniziati, che può intimidire chi le si avvicina. Quella che è forse la sua maggiore debolezza è racchiusa proprio in quest'immagine, lontanissima dal fascino di strumenti alla portata di tutti, eppure capaci di produrre grafiche futuristiche e mappe sfavillanti, che sono associati, nell'immaginario di molti, alle tecnologie dell'informazione geografica. Se si sfatasse quel mito, liberando l'analisi spaziale da quella patina di ostilità,

molti più le si avvicinerebbero con fiducia e senza pregiudizi. Si potrebbe allora comprendere e forse superare la dicotomia apparente che relega l'analisi rigorosa al di là di ciò che è comprensibile, raggiungibile e desiderabile, apprezzandone le potenzialità, in simbiosi con l'intera geografia. Allora, quella cui si è accennato non è solo una tentazione e nemmeno solo una sfida, ma anche una formidabile opportunità per la geografia, che può ambire ad una centralità nel panorama scientifico, se saprà comprendere che l'analisi rigorosa e competente dei dati geografici è una delle chiavi che potranno garantirle un futuro duraturo. Per questo riteniamo necessario, in questo momento storico, aggiungere agli strumenti della geografia un testo che possa aiutare i geografi a comprendere il senso, le potenzialità e i limiti dell'analisi spaziale, ponendosi nei suoi confronti senza pregiudizi, consentendole così di svolgere il proprio ruolo a vantaggio della geografia intera.

Alla metà del secolo scorso una rivoluzione tecnologica, quella del calcolatore elettronico, il *computer*, contribuì all'inesco di un cambiamento così radicale nel modo di fare geografia, che quel periodo fu chiamato rivoluzione quantitativa. Sarebbe riduttivo e persino ingenuo ascrivere alla sola tecnologia un cambiamento tanto profondo, ma senza la tecnologia quel cambiamento non sarebbe potuto avvenire. Pochi decenni dopo, gli anni '80 furono caratterizzati dalla diffusione di massa del computer, favorita, se non addirittura resa possibile, dalla drastica diminuzione dei prezzi di *hardware* e *software*, dall'agilità del personal computer, il PC, e dalla popolarizzazione dell'interfaccia, che divenne intuitiva, *user-friendly*, e quindi accessibile ai più. Quelli furono, per la geografia, gli anni della diffusione dei sistemi d'informazione geografica, i GIS, che, ancor giovani, seguirono in tempi brevi l'onda rivoluzionaria del PC, su cui giravano veloci, venduti a prezzi ampiamente accessibili, anch'essi popolarizzati da interfacce intuitive e semplificazioni che risparmiavano all'utente il fardello di teorie, calcoli e altre complicazioni, in nome del tutto-e-subito, e pure del successo e del profitto, dentro e fuori l'università.

In questa storia, lunga ormai 70 anni, poco o punto si parla di analisi spaziale. Pochi sanno persino cosa sia l'analisi spaziale, ancor meno la sanno fare, certi se ne credono capaci, confondendola con tabelle e statistiche descrittive, e altri ancora, con tronfia ingenuità, la credono il mestiere di piazzare pallini sulle carte geografiche. Eppure, la

letteratura di decenni è costellata di contributi importanti e voci autorevoli che ci ricordano che l'analisi spaziale fu la motivazione e rimane lo scopo ultimo dei GIS; che nonostante le critiche, talvolta anche aspre, ha continuato a resistere con tenacia; e oggi, proprio in questa fase storica, l'analisi spaziale potrebbe salvare ancora la geografia dalla valanga tecnologica che minaccia di travolgerla.

Siamo sommersi dai dati. I dati sono ovunque e di una mole impensabile, per questo ricorriamo ad un anglicismo dal suono banale, *big data*: masse enormi di dati creati dai cosiddetti *social media*, dati trasmessi di continuo da una miriade di satelliti, fino alla cessione gratuita di dati sul *web* (*open data*), e alla cessione volontaria di dati correlati della propria posizione geografica (*VGI - volunteered geographical information*). In tutti i casi non si tratta semplicemente di dati, ma di dati riferiti (o riferibili) ad una posizione geografica (georeferenziati). E per i dati vi è la tecnologia: *web2* permette ad un pubblico assai vasto di accedere alla tecnologia dell'informazione geografica *online*, consentendogli addirittura di analizzare i dati e costruire le proprie mappe. Si tratta di tecnologie pervasive, invasive e gravide di conseguenze solo in parte note o prevedibili. Anche volendo tralasciarne le implicazioni sociali, politiche, economiche, non si può prescindere dalle questioni etiche, dai paradigmi, dalle teorie che dovrebbero puntellare l'analisi dei dati e dalle ipotesi che dovrebbero guidare l'interpretazione delle analisi e definire la loro relazione con il mondo che quei dati hanno la pretesa di rappresentare. Per tutto questo, per trarre sapere dai dati, facendo un uso consapevole e responsabile delle tecnologie, c'è l'analisi spaziale, una disciplina ancor giovane e per certi aspetti sfuggente, di nicchia, conosciuta da pochi, praticata da ancor meno. Elitaria forse, per non aver ceduto a tentazioni populiste e commerciali. Adombrata dai giganti dei poteri commerciali del *software* e del GIS, e offuscata, nella percezione di tanti, dai fantasmi del vecchio quantitativismo e della fisica sociale. Tuttavia, nella geografia, è la disciplina più preparata ad analizzare con competenza e rigore i dati spaziali, capace di teoria e di critica oltre che di statistica e matematica; capace di analisi oltre che di descrizione. E proprio questa disciplina potrebbe, oggi, non solo fornire un contributo fondamentale alla risoluzione dei grandi problemi del nostro tempo, ma anche riaffermare, con la propria competenza e rilevanza, l'utilità e l'importanza, se non la necessità, della geografia.

In questo sta dunque il perché di questo libro, la ragione di questo contributo: porre in luce, esplicitare, rendere più accessibile questa disciplina ancora troppo poco nota, che fu e rimane la motivazione e il fine ultimo del GIS, ma che dopo averlo dato alla luce finì nell'ombra, prigioniera forse di se stessa, del dibattito disciplinare, del populismo e dello strapotere economico del suo stesso rampollo. Farla conoscere per ciò che è ai tanti che in essa non si riconoscono, ma che forse lo farebbero se la comprendessero meglio. L'analisi spaziale non è il semplice utilizzo di un sistema d'informazione geografica, non è l'implementazione meccanica di una tecnica analitica, non è la descrizione superficiale di una serie di dati. È una disciplina complessa, che utilizza sì i dati, la statistica, i sistemi d'informazione geografica, ma che pure s'interroga sulla rappresentazione dei fenomeni, sul rapporto tra realtà, dati, modello, analisi e risultati, su se stessa, i propri strumenti e i propri limiti, che sa e vuole dialogare con la geografia intera. È una disciplina giovane che discende, in parte, da geografia quantitativa e sistemi d'informazione geografica, e che, come questi, ha peccato di meccanicismo, superficialità e fede nel potere dei dati. Ma è ormai anche una disciplina matura, che ha saputo cogliere costruttivamente le critiche e riflettere con coraggio su se stessa, intraprendendo il cammino che può condurla a superare le proprie debolezze.

Nello spirito espresso in queste righe questo libro esce *open access*, per promuovere anche in questo modo la conoscenza e l'applicazione consapevole dell'analisi spaziale. Ho scelto un titolo che, echeggiando toni della letteratura recente in lingua inglese¹, è volutamente ambiguo e intende essere a un tempo provocatorio, giocoso e persino dissacrante. Il testo è scritto quasi interamente al plurale narrativo. Riferendomi a persone generiche come "il geografo" o "i professionisti del GIS" uso il maschile, secondo la regola della lingua italiana, che nel maschile include tutte le persone, indipendentemente dal sesso e genere in cui si riconoscono.

Il testo si snoda in una serie di capitoli raggruppati in tre macro-contenitori. Nella prima parte "Origini e metodi", si traccia la vicenda storica dell'analisi spaziale, partendo dalle radici che affondano nel

¹ Ci si riferisce in particolare a Harris *et Alii*, 2014 e Franklin *et Alii*, 2021.

GIS, nella geografia quantitativa e nella statistica (capitolo 1); si discutono i processi spaziali come oggetto della disciplina, accennando ai metodi quantitativi e statistici per la loro analisi (capitolo 2); si considera poi il senso dell'analisi spaziale oltre il rigore quantitativo, evidenziando la consapevole relatività dei suoi risultati (capitolo 3). Nella seconda parte "Definizione, identità, autocritica" si procede ad una definizione che distingue l'analisi spaziale dal GIS e la colloca in relazione alla geografia quantitativa, esaminando criticamente la molteplicità terminologica con cui, più o meno appropriatamente, la si designa (capitolo 4); infine ci si sofferma sul pensiero critico, sviluppato perlopiù nel contesto della *GIScience*, e a ciò si aggiungono considerazioni sull'etica, il ruolo dei dati nello sviluppo della scienza, l'inclusione di altri paradigmi e metodi (capitolo 5). Nella terza parte "La pratica dell'analisi spaziale" si propone una panoramica di tre aspetti fondamentali della pratica: i dati e i problemi ad essi connessi; la casetta degli attrezzi e la costruzione di tali attrezzi; il personale e la sua formazione (capitolo 6); si presenta quindi una breve rassegna di applicazioni concrete dell'analisi spaziale, illustrandone i contributi alla comprensione e gestione del territorio (capitolo 7); infine si discute la letteratura in lingua italiana e la vicenda dell'analisi spaziale nel nostro Paese (capitolo 8).

Parte I

Origini e metodi

1. *Le origini*

The issues of change and progress in human geography remain of interest not only to those concerned to chart the discipline's history but also those wishing to alter its trajectory.

Johnston, *On Disciplinary History and Textbooks: or Where Has Spatial Analysis Gone*, 2000, p. 125.

1.1. Premessa

Questo volume inizia affrontando un compito non facile: delineare brevemente la storia dell'analisi spaziale. Si tratta di una disciplina ancora giovane, la cui ricerca di una propria identità ed autonomia non è del tutto compiuta, come non lo è del tutto la riflessione profonda su se stessa, le proprie origini e il proprio futuro. Per queste ragioni la letteratura sulla storia dell'analisi spaziale rimane sporadica, non sempre sistematica e soprattutto innestata in filoni diversi, primo e predominante quello del GIS¹. In parte ciò è dovuto al fatto che l'analisi spaziale non ha ancora del tutto acquisito un'autonoma coscienza di sé, e quindi la riflessione sulla sua identità e il discorso su se stessa si trovano solo all'interno di altri discorsi, primariamente in quello sul GIS. Tuttavia, dato che non vi è coincidenza perfetta tra analisi spaziale e GIS, queste storie e queste riflessioni rimangono ancor più parziali e frammentarie². Per quanto abbozzato e offuscato, da questa letteratura emerge il ritratto di un'analisi spaziale che discende anche da altre discipline, mutuandone, in parte, paradigmi e metodi.

¹ Si veda al riguardo: Johnston (2000); Goodchild (2000); Barnes (2010). Il termine GIS sarà discusso ampiamente in seguito. Esso si riferisce generalmente a "sistemi d'informazione geografica" [*geographic information systems*] oppure a "scienza dell'informazione geografica" [*geographic information science*].

² Goodchild (2000, p. 5): «The early development of GIS was driven by spatial analysis, but more recent visions of the role of information technology generally, and GIS in particular, have tended to widen rather than narrow the gap between the two fields. GIS is seen as serving a much more populist vision, in which rigorous and mathematically sound analysis is challenged from a number of directions».

Per poter leggere questa storia, bisogna iniziare con una definizione di analisi spaziale, che sarà ampiamente discussa e affinata in seguito. La voce “analisi spaziale” viene definita, dall’*Enciclopedia Internazionale della Geografia*, come «il processo che genera conoscenza utile per rispondere a domande sull’ordine, la disposizione e la struttura inerenti ai problemi spaziali»³. Essa costituirebbe «un dominio unico, che esiste all’incrocio di molte discipline e prospettive metodologiche diverse», spesso accomunata al GIS, tuttavia, «benché l’analisi spaziale possa comprendere alcune funzionalità GIS, essa va ben oltre le possibilità offerte dall’impiego del GIS»⁴.

In questo capitolo non si ambisce a tracciarne una vera e propria storia dell’analisi spaziale, ma piuttosto ad individuarne le radici che, crediamo, in varia misura, ne abbiano condizionato l’evoluzione. Ci soffermeremo su alcuni nuclei originari: anche se cronologicamente non fu il primo, inizieremo dal GIS, la disciplina più strettamente connessa con l’analisi spaziale; proseguiamo poi con le radici che affondano nella geografia teoretico-quantitativa, e da qui all’apporto dell’econometria spaziale, al legame profondo con la statistica, e alla relazione con le tecnologie dell’informazione geografica. Questa breve panoramica non si spingerà oltre la letteratura anglosassone, in cui l’analisi spaziale emerse e che pure accorpa importanti contributi anche da realtà geografiche diverse da quella anglofona⁵. Un discorso a sé meriterebbe la letteratura in lingua italiana, su cui si ritornerà nel capitolo 8. Concluderemo infine con una breve rilettura di questa vicenda che ne individui i problemi aperti che seguiremo nel resto del volume.

³ Wolf e Murray (2016, p. 1): «Primarily, spatial analysis is the process of generating useful knowledge to answer questions about order, pattern, and structure inherent in spatial problems».

⁴ Wolf e Murray (2016, p. 2): «Because spatial analysis is often strongly coupled with explicit GIS functionality, discussion of spatial analysis is sometimes equated to GIS operations. This is an incomplete view. While spatial analysis may include GIS functionality, it goes beyond capabilities offered through the use of GIS».

⁵ Anselin (2010, p. 12): «In continental Europe, interest comes primarily from researchers in the Netherlands, who are almost all trained in economics and econometrics. Examples include Paelinck, Klaassen, Hordijk, Brandsma, Bartels, Blommestein, Folmer, Nijkamp and Ketellapper, among others».

1.2. L'analisi spaziale all'ombra del GIS

Le origini del GIS rimangono tuttora oggetto di dibattito. Se si considerano come essenza del GIS l'organizzazione, visualizzazione ed interpretazione dei dati spaziali, i primi esempi storici si fanno risalire al cartografo francese Berthier, che avrebbe utilizzato la sovrapposizione [*overlay*] per descrivere i movimenti delle truppe nel corso della battaglia di Yorktown (1781), durante la guerra d'indipendenza americana⁶. Un altro esempio classico è il lavoro del medico inglese John Snow, che durante l'epidemia di colera di Londra del 1854 visualizzò i decessi individuando l'origine dell'epidemia nell'acqua contaminata⁷. Nella sua “Storia del mondo in 12 mappe”, Jerry Brotton azzarda addirittura che il primo GIS sia nientemeno che la “Geografia” di Tolomeo: un testo che contiene –tra l'altro– un elenco di località, corredate delle rispettive coordinate geografiche, accompagnate da una specie di manuale di regole per porle su una mappa.

Realisticamente il GIS è un fenomeno assai più recente i cui inizi si possono far risalire attorno al 1960⁸. Tracciarne una storia sarebbe problematico, innanzitutto perché tale storia dipende dalla definizione che si dà del GIS. Tra le definizioni più classiche e generalmente accettate vi è quella di: “una raccolta integrata di *hardware*, *software*, dati e personale che operano in un contesto istituzionale”⁹. Tuttavia, poiché essa contiene gli elementi *hardware* e *software*, che non comparvero se non dopo la metà degli anni '60, essa ne escluderebbe le prime realizzazioni. Inoltre, la fase iniziale del GIS fu caratterizzata

⁶ Waters (2017, p. 1): «Historical accounts of the origins of GIS invariably pay tribute to the intellectual debt owed to those who sought to organize, visualize, and interpret spatial data. Commonly these discussions reach back into the distant past, frequently referencing the map overlays of troop movements created by the French cartographer, Berthier, during the Battle of Yorktown in the American War of Independence».

⁷ Si veda Johnson (2006). Un'interessante descrizione del lavoro di Snow si può trovare in Garfield (2016, pp. 234-245). In ogni caso il lavoro di Snow costituisce un contributo fondamentale per l'analisi spaziale e il cosiddetto “*spatial thinking*”.

⁸ Sugli inizi e l'evoluzione del GIS si veda anche Bertazzon (2001, p. 423): «Così mentre altri settori e altre discipline si ritagliano un posto nel big business del GIS, la geografia accademica si trova a dover correre ai ripari. Si accorge che durante la sua “formidabile rinascita” (Bertazzon e Waters 1996) ha seguito una politica di breve periodo, e non si è garantita un futuro, un futuro che forse non aveva saputo prevedere».

⁹ Maguire (1991, p. 17): «A Geographical Information System is best described as an integrated collection of hardware, software, data and liveware which operates in an institutional context».

da tentativi e progressi non coordinati, in ambiti diversi, spesso ignari gli uni degli altri. Molto del lavoro che portò allo sviluppo del GIS non si contraddistinse all'inizio come tale e spesso fu caratterizzato da fallimenti o tentativi ripetuti, di cui spesso si persero le tracce¹⁰. Scegliendo una definizione che non contenga gli elementi di *hardware* e *software*, si potrebbero rintracciare i primordi del GIS nel lavoro svolto all'Università del Washington a Seattle, dove alcuni geografi ed ingegneri svilupparono metodi quantitativi negli studi sui trasporti¹¹. A fianco a questi metodi quantitativi va considerato il contributo concettuale apportato dall'architetto del paesaggio Ian McHarg, esposto nel suo famosissimo testo *Design with Nature*. L'intuizione di McHarg, tanto semplice quanto potente, fu quella di scomporre l'informazione geografica in mappe tematiche, che poi andranno sovrapposte per supportare decisioni gestionali. Benché McHarg svolgesse tale compito manualmente, utilizzando fogli trasparenti, l'idea è tanto basilare per il GIS da esserne diventata forse la più nota immagine iconica¹².

Diverse definizioni furono proposte su basi funzionali, intese ad identificare ciò che un GIS può svolgere, che non possa essere svolto da generici sistemi d'informazione (IS). In questo senso il tratto distintivo del GIS è che il suo punto focale sono le entità e le relazioni spaziali, ciò che apre la questione su cosa distingue un GIS da un *software*

¹⁰ Tracce di tale storia sono probabilmente sepolte in documenti riservati di agenzie governative, ditte di *software* e istituti militari. Ben pochi erano gli accademici che vi lavoravano, e anche quei pochi erano troppo impegnati nella realizzazione di progetti per scrivere relazioni. Oltretutto erano spesso sottoposti a vincoli di segretezza e, particolarmente agli inizi, poi, non vi erano riviste, congressi, convegni o conferenze adatte ad ospitare eventuale letteratura sull'argomento. Si veda quanto affermato da Coppock e Rhind (1991, pp. 22-23).

¹¹ Coppock e Rhind (1991, p. 26): «In the 1950s, both geographers (notably Garrison) and transportation engineers (notably Horwood) developed quantitative methods in transportation studies. Garrison's colleagues and students included Berry, Tobler and Marble; Horwood's included Barb and Dueker (see Dueker's important 1974 paper)».

¹² Si vedano ad esempio il tributo di URISA (Urban and Regional Information Systems Association), www.urisa.org/awards/ian-mcharg/, visitato il 18/10/2021, e Goodchild (2010, p.7). «Although there are many historic roots of geographic information systems (GIS; Forsman, 1998), one of the strongest lies in the notion of making design decisions by overlaying maps, each map representing one of the actors important in the decision. The net effect of each of the factors would be represented by the amount of light penetrating the layers at each point, allowing the decision maker to make an intuitive judgment as to the best solution. This is one of the central ideas of McHarg's (1969) *Design with Nature*, and the stack of layers has become an icon of the entire field of GIS, appearing on the front covers of many of its textbooks».

per la produzione di cartografia. In tal senso, la caratteristica distintiva del GIS è la capacità di svolgere operazioni di *analisi e modellizzazione* spaziale, che un pacchetto di cartografia non può svolgere. Queste funzioni, tuttavia, lo avvicinano ai pacchetti di statistica; qui l'elemento distintivo è che il GIS, *diversamente* dai pacchetti statistici, richiede l'accesso non solo al dato attributo, ma anche al dato geografico, o locazionale. Analogamente, il GIS presenta elementi in comune con i CAD (Computer Aided Design), ma, diversamente dai CAD, il GIS può immagazzinare dati *locazionali* e dati attributo in un database ed utilizzarli per realizzare funzioni *analitiche*. La ricerca di funzioni *uniche* del GIS, avviata da punti diversi, conduce costantemente alla *gestione ed analisi di dati geografici*. I GIS sarebbero dunque dei Sistemi di Supporto alle Decisioni (DSS), e più precisamente degli SDSS (*Spatial Decision Support Systems*). Se il GIS è un SDSS, il suo sviluppo comporta, oltre gli elementi tecnici, un apparato concettuale che comprende la conoscenza del contesto e la capacità di valutare rischi, conseguenze ed impatti di ogni decisione alternativa (Beaumont, 1991; Fotheringham, 1991). Centrale allo sviluppo dei SI Geografici, e al loro uso, è proprio tale apparato concettuale, così come la conoscenza e l'*expertise* spaziale, cui gli aspetti tecnici fanno solo da supporto. È interessante notare come, dalle varie definizioni, emerga proprio l'analisi dei dati spaziali come elemento centrale e distintivo del GIS. Proprio questa definizione farà affermare a Michael Goodchild che l'analisi spaziale è ciò che il GIS è stato creato per fare, e sarà uno degli elementi fondamentali della transizione del GIS da sistema a scienza dell'informazione geografica.

Al di là di questi primordi, è generalmente accettato che la prima realizzazione documentata e rispondente alla definizione di GIS si deve a Roger Tomlinson, che alla metà degli anni '60 riuscì a persuadere il governo del Canada ad investire nella realizzazione del primo vero sistema informatizzato per la gestione di dati geografici, il CGIS (Canada Geographic Information System)¹³. Fu proprio Roger

¹³ Roger Tomlinson aveva lavorato in precedenza per una ditta privata interessata ad identificare siti adatti per nuove piantagioni nell'Africa orientale. I costi proibitivi dell'operazione (circa 8 milioni di dollari canadesi ai prezzi del 1965 e la necessità di 556 tecnici per tre anni per realizzare le mappe in scala 1:50.000 del Canada Land Inventory) costrinsero la ditta ad accettare la proposta di Roger Tomlinson di sviluppare una metodologia digitale, per cui egli aveva trovato l'appoggio dell'IBM, che si stava allora interessando alla digitalizzazione di

Tomlinson a coniare il termine GIS per la metodologia da lui realizzata, e a ben ragione: la proposta che egli presentò al ministero dell'agricoltura canadese conteneva già tutti gli elementi concettuali e applicativi che avrebbero determinato il futuro del GIS. Essi spaziavano dai perfezionamenti dell'*hardware* fino a concetti di database e al primo uso documentato del concetto *link/node* per la codifica di linee e segmenti¹⁴. È proprio in virtù di questi elementi concettuali che il CGIS viene considerato il primo, innovativo, sistema d'informazione geografica, e non a caso fu l'elaborazione scientifica di tali concetti che permise a Roger Tomlinson di realizzare il CGIS. L'importanza dell'elemento concettuale è tale che secondo Michael Goodchild (1992), nella realizzazione del CGIS l'elemento fondamentale fu l'essere partiti da un approccio "scientifico" e non "applicativo": l'aver cioè dovuto sviluppare concetti e postulati allo scopo di finalizzare gli strumenti operativi esistenti e organizzarne dei nuovi. Tuttavia, se la realizzazione del primo GIS alla metà degli anni '60 fu la conseguenza del grande fermento verificatosi negli anni precedenti, lo sviluppo successivo fu veramente lento e solo dopo la prima metà degli anni '80, quando i prezzi dell'*hardware* iniziarono a scendere drasticamente e

foto aeree. Roger Tomlinson fu così in grado di proporre la propria metodologia al Dipartimento Canadese per l'Agricoltura, che aveva allora in programma la realizzazione del Canada Land Inventory (CLI), che comportava la realizzazione di tutte le mappe della parte abitata del Canada. Nel 1971 il CGIS divenne operativo, fu successivamente riorganizzato e semplificato e agli inizi degli anni '90 conteneva archivi digitali di circa 10.000 mappe per oltre un centinaio di temi; si vedano al riguardo: Coppock e Rhind (1991) e Bertazzon (2001).

¹⁴ Si veda quanto affermato da Coppock e Rhind (1991, p. 23). Gustosissima è la narrazione in prima persona dello stesso Tomlinson (1984) sulla creazione del CGIS. La rivista "*The Operational Geographer*" ha però cessato di esistere nel 1993, e non ci è stato possibile reperirne una copia digitale. Il contenuto dell'articolo fu inizialmente presentato da Tomlinson come *Keynote Address*, intitolato appunto "Geographic Information System – A New Frontier" all'*International Symposium on Spatial Data Handling* del 1984, ed è cortesemente condiviso da Duane Marble e scaricabile dalla sua pagina ResearchGate (www.researchgate.net/profile/Duane_Marble), di cui riportiamo qualche passo significativo: «The first geographic information system was developed in Canada in the 1960s. At that time, the emphasis was on overcoming the technical problems that arose in handling traditionally graphic data by numerical methods, of putting maps into computers that by today's standards were slow, small and cumbersome» (p. 1); «These comments are not intended to be a detailed review of system capabilities in the sixties. Rather they make the point that given the constraints, the pace of development was surprisingly swift, and many of the basic techniques of handling spatial data with computers were invented and put in place during the decade. In fact, many of the GIS systems that exist today, albeit running on much cheaper computers, to a great degree still exhibit sixties-level functional capabilities for geographical analysis» (p. 5).

una gran quantità di programmi divenne accessibile al vasto pubblico, si assistette alla diffusione quasi di massa del GIS¹⁵. La fine degli anni '60 e gli anni '70 furono descritti da Tomlinson come un periodo di “diffusione laterale”, ossia un'epoca di miglioramenti apportati ai sistemi esistenti e di diffusione degli stessi, ma non di vera e propria innovazione¹⁶.

Di storia del GIS si parla già all'inizio degli anni '90, nei due articoli di apertura del primo fondamentale volume collettivo *Geographical Information Systems*¹⁷. Nel capitolo in cui si definiva il GIS, David Maguire (1991), ne delineava lo sviluppo storico secondo tre fasi: la fase dell'inventario, quella dell'analisi, e quella della gestione [*management*], consapevole tuttavia che tali fasi dovevano ancora compiutamente realizzarsi. Tracciando invece la storia del GIS, Terry Coppock e David Rhind (1991) descrivevano lo sviluppo del GIS secondo quattro fasi parzialmente sovrapposte: la prima, un periodo pionieristico, la cosiddetta *ricerca di frontiera*, copriva all'incirca il periodo tra gli anni '50 e il 1975, si sviluppò esclusivamente negli Stati Uniti e nel Regno Unito; la seconda fase, denominata *sperimentazione formale*, si svolgeva circa dal 1973 ai primi anni '80; la terza fu la *fase commerciale*, iniziata approssimativamente nel 1982; questa fase, a causa della forte concorrenza tra i venditori, avrebbe lasciato il posto alla fase della *dominazione dell'utente*, che avrebbe rappresentato «la fine dell'inizio del GIS».

Da allora vi furono altri contributi alla storiografia e all'interpretazione della vicenda del GIS¹⁸, e tra questi è certo degno di nota quello di Nigel Waters (2017). Secondo questo autore la fase della *dominazione dell'utente* si sarebbe protratta dalla fine degli anni '80 alla metà degli anni '90. A questa sarebbe seguita l'era della *scienza*

¹⁵ Bertazzon e Waters (1996, p. 27): «Bisognava però attendere gli anni '80, quando i prezzi dell'hardware iniziarono a scendere drasticamente e quando una grande quantità di programmi divenne accessibile al vasto pubblico, perché la comunità dei geografi si accorgesse della crescente diffusione dei GIS come strumenti per la risoluzione dei problemi geografici della realtà di ogni giorno nelle aree della pianificazione e dell'amministrazione delle risorse».

¹⁶ Coppock e Rhind (1991, p. 30): «Tomlinson (1988) described the 1970s as a period of lateral diffusion rather than of innovation and there is considerable piecemeal evidence to show an increasing interest among a variety of agencies at all levels of government – federal, state, county and city».

¹⁷ Maguire, Goodchild, Rhind (1991).

¹⁸ Waters (2017); Goodchild (2018); Wilson (2015); Foresman (1998).

dell'informazione geografica (*GIScience*), iniziata verso la metà degli anni '90 e riconducibile al lavoro teorico pubblicato da Michael Goodchild (1992) in cui si auspicava che il GIS si elevasse oltre la mera prospettiva tecnologica, per dedicarsi a questioni scientifiche fondamentali realizzando così la propria vocazione potenziale. Descrivendo gli anni '90 come il periodo della diffusione della scienza dell'informazione geografica Nigel Waters (2017) tuttavia non fornisce una data a cui questa fase si possa dire conclusa; ad essa egli fa seguire la fase delle *reazioni al GIS*. Si tratta prevalentemente di reazioni critiche, anche aspre, che Waters fa risalire agli anni '70 innestandole nel filone meglio rappresentato dal celebre lavoro di John Pickles¹⁹. La reazione della comunità del GIS alla critica non si fece attendere e vide sin dalla metà degli anni '90 iniziative volte ad integrare l'approccio tecnocratico GIS con approcci di tipo qualitativo, in particolare con forme partecipative di GIS (*PPGIS - public participation GIS*). Vi sono pochi dubbi che il risultato più significativo, apparentemente quasi paradossale, di questa controreazione fu il sorgere di un pensiero critico in seno al GIS [*critical GIS*], che consentì di spostarne la critica dall'esterno all'interno, inaugurando così una fase radicalmente nuova, e tuttavia parte integrante del discorso del GIS come scienza²⁰.

L'ultima fase della storia del GIS sarebbe iniziata, sempre secondo Nigel Waters, nel 2005, e si tratterebbe, ancora una volta, di sviluppi trainati dalla tecnologia. Tra di essi i più degni di nota sono la

¹⁹ Pickles (1995a): si tratta di un importante testo dal significativo titolo *Ground Truth* in cui Pickles mette in discussione il GIS e il suo ruolo nell'accademia e nella società. Si veda ad esempio il suo saggio introduttivo Pickles (1995b) in cui afferma (p. 20): «GIS is just one of these new complexes of discourse, practice, and institutional ensemble, among many others, effecting changes in the modalities of power. As a cultural practice, instituted historically, its forms and effects are consequently contingent, ungrounded except in terms of other, prior, contingent historically instituted practices. In this sense, power is as much about the possibilities of modernization – the ways in which identity and differences are constituted – as about the exercise of influence and the formation of new iron cages. As social relations and new subjectivities are embodied, we need to ask how such identities are sustained, how power flows through the capillaries of society in particular settings, and what role new technologies of the self and of society play in this circulation of power. Foucault would have understood well our contemporary fascination with GIS, its “technologies of surveillance”, forms of knowledge engineering, and commitment to the categorizing and normalizing of nature and social life».

²⁰ Particolarmente influenti per la definizione del *critical GIS*, detto anche *GIS2*, furono i lavori di Sheppard *et Alii* (1999) e Schuurman (2000).

volunteered geographic information, o VGI, ossia la cessione volontaria di informazioni geolocalizzate, resa possibile dalla tecnologia e diffusione dello *smartphone*²¹; il cosiddetto *web-based mapping*²², consentito dalla disponibilità di *software GIS online*, che permette all'utente di costruire le proprie mappe usando un applicativo che non è installato nel proprio computer (o tablet, smartphone, o piattaforma equivalente). Al discorso di Nigel Waters sul *web-based mapping* vogliamo aggiungere l'importanza degli *open data*²³, accessibili all'utente senza restrizioni ed utilizzabili tramite gli applicativi *online*. Tornando all'elenco di sviluppi recenti proposto sempre da Nigel Waters, citiamo il *mobile GIS*, per cui basti l'esempio delle applicazioni che si possono realizzare nella propria automobile utilizzando tecnologia GIS e dati registrati dal navigatore satellitare²⁴, il *cloud computing* e i *big data*²⁵. Si può concludere citando il recente riconoscimento dell'importanza del cosiddetto *geodesign*, che potremmo definire sinteticamente come l'integrazione dell'informazione geografica nella pianificazione territoriale e i relativi processi decisionali²⁶.

Concludiamo questo breve excursus sulla storia del GIS notando che sul finire degli anni '90 la comunità accademica sembrò prendere coscienza dell'esistenza dell'analisi spaziale come entità distinta dal GIS. Per questa ragione, l'NCGIA, presieduta da Michael Goodchild, finanziò una conferenza, tenutasi a Santa Barbara nel 1998, intitolata "*Workshop on Status and Trends in Spatial Analysis*"²⁷. Gli interventi presentati alla conferenza furono ripresi dall'allora neonata rivista *Journal of Geographical Systems*, che dedicò il primo numero del nuovo millennio al rapporto tra analisi spaziale e GIS²⁸.

²¹ Goodchild (2007).

²² Waters (2017).

²³ Janssen, Charalabidis, Zuiderwijk (2012).

²⁴ Miller (2007).

²⁵ Miller e Goodchild (2015).

²⁶ Waters (2017); Dangermond e Goodchild (2020).

²⁷ Una selezione di quei contributi è contenuta nel volume riassuntivo, strutturato in quattro parti, ciascuna organizzata intorno ad un contributo significativo: Frank (2000); Goodchild (2000); Anselin (2000); Box (2000); Getis (2000).

²⁸ Getis (2000 p. 1): «from the early days of GIS until now, modellers who work closely with geo-referenced data have had the gnawing feeling that those responsible for the great technological progress of GIS research have paid relatively little attention to spatial data analysis».

1.3. Le radici quantitative

Se l'analisi spaziale è «un processo per generare conoscenza che risponde a domande sull'ordine, la disposizione e la struttura inerenti ai problemi spaziali»²⁹, è importante pure comprendere in che modo tale processo avvenga e quali metodi siano impiegati per raggiungere tale conoscenza. In tal senso Levi Wolf e Alan Murray proseguono: «Le tecniche dell'analisi spaziale cercano di facilitare la risposta a queste domande tramite l'applicazione di metodi formali matematici, statistici e/o computazionali», e ancora: «L'analisi spaziale fornisce modi unici, potenti ed incisivi per generare conoscenza utile perché comprende l'uso di metodi quantitativi come parte di un processo»³⁰. L'impiego dei *quantitative methods* è dunque elemento distintivo e fondamentale dell'analisi spaziale, la cui forza non è il mero impiego di tali metodi, ma il loro utilizzo come parte di un processo, all'interno del quale essi esplicano le capacità, che solo essi possiedono, di generare conoscenza relativa ai fenomeni spaziali. La centralità dei *metodi quantitativi* suggerisce dunque che il secondo filone su cui si sono innestate le radici multidisciplinari dell'analisi spaziale sia la geografia teoretico-quantitativa: sarà dunque nella storia di questa che si possono cercare altre tracce della storia dell'analisi spaziale.

Le origini storiche della geografia teoretico-quantitativa si fanno comunemente risalire agli anni '50 e '60 del secolo scorso: fu un fenomeno complesso³¹, su cui esiste una corposa ed autorevole letteratura³² non solo di natura teorica³³ ma anche critica³⁴. Fu,

²⁹ Wolf e Murray (2016, p. 1): «Primarily, spatial analysis is the process of generating useful knowledge to answer questions about order, pattern, and structure inherent in spatial problems».

³⁰ Le due citazioni sono tratte da Wolf e Murray (2016, p. 1): «Spatial analysis techniques seek to facilitate answering these questions through the application of formal mathematical, statistical, and/or computational methods. [...] Spatial analysis provides unique, powerful, and incisive ways to generate useful knowledge because it involves the use of quantitative methods as part of a process».

³¹ Si vedano: Livingstone (1992, pp. 304-346); Capel (1987, pp. 205-225); Kwan e Schwanen (2009) e Lando (2020, pp. 180-194).

³² Come accennato in premessa, lo scopo di queste pagine non è quello di aggiungere a quanto già scritto ma, se possibile, di individuare spunti che aiutino a comprendere la vicenda dell'analisi spaziale.

³³ Bunge (1962; 1966) e Harvey (1969).

³⁴ Si veda, ad esempio, la piccola ma gustosa collezione “*The Legends Papers*” pubblicata dalla rivista *Geographical Analysis* in occasione del proprio quarantesimo anniversario

sostanzialmente, una “rivoluzione” che cercava di trasformare la tradizionale geografia descrittiva in una scienza, introducendovi esplicitamente forme di teorizzazione scientifica (da cui il termine teoretica) e tecniche di verifica empirica (da cui il termine quantitativa), utilizzando quindi un approccio analitico, che in gran parte poggiava su un paradigma logico-positivista³⁵.

Una circostanza storica di rilievo è la realizzazione del CGIS di Roger Tomlinson nel 1967, ossia proprio nel *clou* della rivoluzione quantitativa, che potrebbe far supporre un’interazione tra GIS e quantitativismo in geografia, entrambi tradizionalmente collocati all’interno del paradigma neopositivista, molto legati agli sviluppi dell’informatica ed accomunati dal fulcro posto sull’analisi dei dati geografici. Invece Terry Coppock e David Rhind (1991) notavano che la vicenda della rivoluzione quantitativa ebbe scarsissime interazioni con la cartografia computerizzata, soprattutto perché allora le metodologie statistiche erano anche più di oggi numeriche e largamente a-spaziali, e solo dalla metà degli anni ’80 se ne ottenne una rappresentazione grafica. Anche Luke Bergmann e David O’Sullivan (2018) imputano all’enfasi geometrica del GIS il paradosso per cui le due principali anime dell’analisi spaziale sarebbero cresciute, storicamente, in reciproco isolamento. Bisogna anche ricordare che fino agli anni ’80, in effetti, il GIS fu poco più che cartografia computerizzata, spesso percepito dai geografi quantitativi come uno strumento tecnico, avulso dal proprio campo d’interesse³⁶.

Come visto poco sopra, l’essenza dell’analisi spaziale è l’impiego di tecniche quantitative in un processo di generazione di conoscenza: ciò che qui interessa sono dunque le tecniche emerse dalla rivoluzione teoretico-quantitativa e tuttora utilizzate (*mutatis mutandis*)

(2008), che narrano in forma personale e aneddotica il lavoro quotidiano che contribuì a costruire quella che fu poi detta rivoluzione quantitativa in geografia: Haggett (2008a; 2008b) e Stimson (2008).

³⁵ Barnes (2004) e Lando (2020, pp. 180-194).

³⁶ Nel 1992 l’NCGIA lanciò l’Iniziativa di ricerca 14, “Spatial Analysis and GIS I” (Rogerson, Fotheringham, Pellegrini, 1995). A partire da quel periodo uscirono numerosi testi, tra cui *Spatial Analysis and GIS* (Fotheringham e Rogerson 1994), *Spatial Analytical Perspectives on Gis* (Gisdata Series, Vol 4) di Fisher, Scholten, Unwin (1996), *Spatial Analysis: Modelling in a GIS Environment* di Longley e Batty (1997) e *Spatial Models and GIS: New Potential and New Models* (Gisdata) di Fotheringham e Wegener (1999, e infine il già citato workshop del 1998 (Getis, 2000).

dall'analisi spaziale. Nel corso di numerosi decenni il filone quantitativo continuò a fare geografia utilizzando metodi rigorosi, di tipo matematico, statistico e computazionale, producendo ed affinando quelle tecniche che Trevor Barnes (2004) definisce di verifica empirica e che secondo Levi Wolf e Alan Murray (2016) cercano di rispondere a domande su ordine, disposizione e struttura dei problemi spaziali, agendo come parte di un processo capace di generare conoscenza utile. Un elenco delle varie tecniche usate sarebbe lungo e tedioso, oltre che, con ogni probabilità, incompleto, data anche la natura dinamica della disciplina, per cui le tecniche sono mutate, superandosi a vicenda nel corso del tempo.

Tuttavia, vale la pena cercare di comprendere in cosa consistano i metodi tradizionali della geografia teoretico-quantitativa, da cui l'analisi spaziale avrebbe mutuato le proprie tecniche³⁷. Si tratta di una serie molto articolata di tecniche analitiche, che si pongono una vasta gamma di obiettivi, spaziando dall'interpolazione di osservazioni spaziali alla riduzione multivariata di dati, alla correlazione e all'associazione tra variabili. Nonostante l'ampiezza della gamma degli obiettivi e degli strumenti, non è difficile osservare³⁸ che la gran parte delle tecniche usate in geografia teoretico-quantitativa possono essere formulate in termini di analisi di regressione, o comunque sono ad essa riconducibili³⁹.

L'affermazione secondo cui tutte le tecniche quantitative possono essere formulate in termini di regressione consente il passaggio concettuale da un insieme eterogeneo di tecniche disparate ad un'unica

³⁷ Un interessante spaccato su queste tecniche si ottiene scorrendo l'elenco dei titoli della collana CATMOG (Concepts and Techniques in Modern Geography), che fu curata dal 1975 al 1996 dall'Institute of British Geographers (Quantitative Methods Research Group (QMRG) – Royal Geographical Society with IBG -- <https://quantile.info/catmog/>). Si tratta di ben 59 volumi, la gran parte dei quali erano dei manualetti tascabili che illustravano operativamente una di queste tecniche.

³⁸ Griffith e Amrhein, 1997, p.79: «There is little question that “[r]egression analysis is the most widely used and versatile dependence technique.” (Hair *et Alii*, 1995, p.79), and as such is the “...workhorse of much non-spatial data analysis” (Bailey e Gatrell, 1995, p.274). In fact, most every technique examined in this book can be restated as a variation of the regression model». Le tecniche trattate nel testo sono: *multiple regression and correlation; multivariate analysis of variance; principal components and factor analysis; discriminant function analysis; cluster analysis; canonical correlation analysis*; oltre a varie forme di regressione (n.d.a.).

³⁹ Data la fondamentale importanza di tale tecnica, essa sarà definita sia in termini concettuali che computazionali nel capitolo 2.

metodologia portante che le accomuna e definisce. Questo passaggio ci consente di soffermarci su due questioni importanti, a cavallo tra il piano metodologico e quello teorico. La prima riguarda la validità, ovvero la “relatività delle relazioni” calcolate col metodo dell’analisi di regressione e la loro generalizzabilità. Invero, i parametri del modello di regressione⁴⁰ sono calcolati dai dati contenuti nel campione e riferiti alle variabili dipendente ed esplicative. Questo implica che nessun modello di regressione, e quindi nessuna tecnica analitica quantitativa, può calcolare, non già una legge universale, ma neppure una relazione di carattere generale che leghi le variabili esplicative alla dipendente. Viceversa, le tecniche analitiche quantitative possono solo stimare relazioni numeriche valide per un determinato campione, e non generalizzabili oltre l’ambito rappresentato da quel campione. Questa osservazione è importante per comprendere la profonda trasformazione di pensiero che accompagna la lunga transizione dalla rivoluzione quantitativa all’analisi spaziale.

La seconda questione riguarda l’applicazione dell’analisi di regressione ai dati spaziali. L’analisi di regressione, infatti, poggia su una serie di ipotesi che, se soddisfatte, garantiscono le proprietà ottimali delle stime ottenute. Tuttavia, è raro che i dati spaziali soddisfino quelle ipotesi⁴¹, e ciò implica che, generalmente, il calcolo dei parametri del modello da campioni di dati spaziali sia effettuato in violazione di tali ipotesi, cosicché le stime di regressione su dati spaziali non godono di proprietà ottimali, ma tendono ad essere affette da una maggiore variabilità, e quindi da una minore affidabilità, rispetto a quelle ottenute da dati a-spaziali⁴². Si tratta evidentemente di una questione di amplissima portata, che evidenzia una contraddizione tra le proprietà dei dati spaziali e le ipotesi dello strumento utilizzato per analizzarli. La consapevolezza di tale contraddizione fu raggiunta gradualmente, e documentata sporadicamente già dagli ultimi anni ’50⁴³.

⁴⁰ Essi comprendono in particolare il coefficiente di regressione, e suo tramite la stima della variabile dipendente, l’errore, la varianza e tutte le grandezze diagnostiche ed inferenziali del modello (cfr. capitolo 2).

⁴¹ Come si dirà diffusamente nel capitolo 2.

⁴² Anselin (1988, p. XV): «Spatial econometrics deals with spatial dependence and spatial heterogeneity, critical aspects of the data used by regional scientists. These characteristics may cause standard econometric techniques to become inappropriate».

⁴³ Johnston e Jones (2020, pp. 359-360): «In 1956, however, Reynolds raised a number of specific issues regarding the geographical analysis of statistical data, in a brief paper that

Un limite della vecchia geografia teoretico-quantitativa era dunque proprio la debolezza delle stime ottenute con metodi statistici: benché le analisi fossero applicate a dati spaziali, non vi è evidenza che si fosse raggiunta la piena consapevolezza delle proprietà specifiche di quei dati, per cui spesso si applicavano le tecniche come se essi fossero semplicemente privi di riferimento spaziale⁴⁴. Sarà proprio il raggiungimento della consapevolezza di quei limiti, intrinseci dell'analisi quantitativa, a segnare la maturazione e, di conseguenza, lo spostamento della vecchia geografia teoretico-quantitativa verso la moderna analisi spaziale. L'elemento principale di questa transizione è l'affermazione di un limite connesso alla ricerca di un metodo capace di superarlo, pur rimanendo entro il paradigma quantitativo.

received little attention, with only 20 citations according to Google Scholar©: Barnes (1998:216) termed it an 'anodyne review of the potential of statistics in geography [...] which included a few mild cautionary remarks'. Reynolds claimed that geographers needed to develop 'distinctive tools' specific to their discipline, but Garrison (1956: 428) responded that the 'present methods of statistical inference are applicable to spatial-type problems' such as correlation and regression, that 'In short, there is ample evidence that present tools are adequate to our present state of development. No type of problem has been proposed that could not be treated with available tools'. He concluded that 'The logical methods of science are universal. If assertions that geography is somehow different are accepted without proof, we may lose the benefits derived from the findings of others' (p. 429). Indeed, that geography was not considered different is indicated by Hepple (2001: 385) noting that Yule 'in the 1890s [...] constructed both the theory and application of multiple regression analysis, using geographical' rather than individual biometric data (see also Denis and Docherty, 2007). By the 1970s, however, it was clear to some that spatial data of various types needed bespoke methods and developments were put in train».

⁴⁴ Riportiamo come esempio la discussione di Unwin (1975), nel numero 5 della serie CATMOG dedicato alla Trend Surface Analysis (p. 20, corsivi dell'autore): «As Poole and O'Farrell (1971) point out, and contrary to what is often held in the geographical literature, these assumptions are made about the residuals and not directly about the original data. They can be stated as follows: The residuals have an expected mean value of zero. B) They are uncorrelated with a constant variance σ^2 . The assumption of constant variance is that of *homoscedasticity*. It implies, for example, that the variation of the original spatial series about the fitted trend surface is roughly the same no matter what part of the map we examine. The assumption of uncorrelated residuals is also a very strong one, especially in our spatial series analysis where it is likely that the residuals will show *spatial autocorrelation* (Cliff and Ord, 1970) evidenced by the grouping together of similar residual values above or below the trend. Recently, tests for spatial autocorrelation in least squares residuals have been developed and could be used to test whether or not this assumption is valid».

1.4. Radici interdisciplinari: statistica ed economia

Per comprendere la profondità del legame tra analisi spaziale e statistica, basti osservare l'uso corrente, nella letteratura anglosassone, del termine *spatial statistics* come sinonimo e sostituto di *spatial analysis*⁴⁵. Già nei paragrafi precedenti si sono fatte alcune osservazioni: l'indipendenza dello sviluppo storico della geografia teoretico-quantitativa rispetto al GIS e l'uso di metodologie statistico-matematiche proprie di quella geografia, convogliate poi nell'analisi spaziale. La prima suggerisce che il nucleo dell'analisi spaziale giaccia altrove, rispetto al GIS, e la seconda chiarisce dove: è ben vero che i metodi della geografia teoretico-quantitativa sono di tipo statistico-matematico, ma va chiarito che la matematica viene tradizionalmente usata solamente come strumento, linguaggio⁴⁶, mentre ben diversa è la questione per quanto riguarda la statistica⁴⁷.

Come detto in apertura, ormai da alcuni decenni è in uso il termine *statistica spaziale*. Benché in larga misura sinonimo di *analisi spaziale*, il termine da sé suggerisce che l'analisi dei dati spaziali abbia trainato sviluppi della statistica che possono essere ritrovati anche in quella disciplina, o all'intersezione tra le due⁴⁸.

⁴⁵ Per una definizione più dettagliata del termine si rimanda al capitolo 4.

⁴⁶ Per questo le connessioni tra la matematica e la geografia teoretico-quantitativa saranno trascurate in questa sede.

⁴⁷ Come nel caso della geografia quantitativa, non s'intende, in questo volume, tracciare una storia della statistica, ma semplicemente individuarne alcuni elementi rilevanti per la vicenda dell'analisi spaziale.

⁴⁸ Unwin (1996, p. 541): «Secondly, work by statisticians has developed a substantial body of statistical theory about spatial data to which GIS users can turn but which did not exist in geography's so-called "quantitative revolution" (see, for example, Ripley (1981; 1988), Upton (1985; 1989), Haining (1990), Cressie (1993), Walden e Guttorp (1992), Diggle (1983)). The statistical view is characterized by the notion that spatially distributed information can be regarded as the outcome of some stochastic process operating in the plane. If we can postulate the nature of the process in mathematical terms, we can deduce its spatial outcomes and examine whether or not an observed pattern is a plausible realization of it. As Harvey forcefully pointed out many years ago (Harvey 1966), a much more difficult alternative is to identify the process and model it appropriately from the evidence of a single mapped realization, yet this is frequently what is required». Cressie (1993, pp. XV-XVI): «Design, inference, and diagnostics are natural consequences of the data-model symbiosis, and all play an important role in Statistics for spatial data. [...] This may be the last time spatial Statistics will be squeezed between two covers. A healthy exponential growth of the literature is apparent from the bibliography». Gelfand *et Alii* (2010, p. 12): «The challenges posed by the need to

La statistica impiegata in analisi spaziale, e con essa intrecciata, non è semplicemente statistica elementare o descrittiva, ma comprende “analisi del residuo”, “statistica inferenziale” e in molti casi l’impiego di strumenti matematico-statistici relativamente sofisticati⁴⁹. L’impiego di questo apparato per l’analisi dei dati spaziali richiede un loro perfezionamento che, in questo modo, stimola sviluppi originali della statistica. Questo rapporto profondo con la statistica differenzia quindi l’analisi spaziale da altri settori della geografia che impiegano comunemente analisi descrittive e dallo stesso GIS, che tende a farne un uso più meccanico. In altri termini: l’analisi spaziale non è un semplice utente passivo della statistica, ma una disciplina che ne stimola lo sviluppo, per rispondere alle questioni poste dall’impiego dei dati spaziali. Questa differenza, accomunando l’analisi spaziale a discipline più rigorosamente quantitative, contribuisce a conferirle quel sapore di disciplina elitaria, per soli iniziati, che tendenzialmente non giova alla sua popolarità.

Come vedremo, esistono delle differenze fondamentali tra dati spaziali e dati a-spaziali, in quanto gli ultimi, diversamente dai primi, di norma soddisfano alcune ipotesi importanti per la statistica. Come si è già accennato, si tratta in particolare di due ipotesi: la prima, anche in senso cronologico, a porsi all’attenzione dei geografi quantitativi, fu quella di indipendenza all’interno del campione. L’ipotesi, o assunzione, di indipendenza è alla base non solo di varie tecniche, tra cui la regressione, ma, ciò che è ancor più importante, dell’inferenza e della statistica inferenziale; eppure, i dati spaziali tendono a violarla perché caratterizzati da dipendenza spaziale. La seconda proprietà è l’eterogeneità spaziale, che presenta analogie con la prima, sia perché entrambe le proprietà sono legate alla covarianza dei processi spaziali, sia perché entrambe producono effetti analoghi sulle analisi statistiche ed inferenziali.

La consapevolezza di questa peculiarità dei dati spaziali fu raggiunta gradualmente, dopo il culmine della rivoluzione quantitativa, quando le tecniche venivano applicate in maniera che si potrebbe dire più naïve. La presa di coscienza delle proprietà dei dati spaziali e delle

analyze spatially referenced data have led to their playing a prominent role in the development of modern statistical methodology».

⁴⁹ Si veda in proposito il paragrafo 5.

loro conseguenze sulle analisi statistiche stimolò la ricerca finalizzata allo sviluppo di metodi che possiamo dire correttivi atti a mitigare gli effetti dei dati spaziali sulle analisi statistiche e le procedure inferenziali. Per entrambe le proprietà, gli sviluppi che condussero alla definizione di metodi statistici specifici si rintracciano nella ricerca economica ed econometrica⁵⁰.

Negli stessi anni in cui la geografia, assieme ad altre scienze sociali, viveva la rivoluzione quantitativa, in economia si sviluppavano i metodi econometrici che, come suggerisce il nome, cercavano di misurare, ovvero stimare le relazioni tra grandezze economiche⁵¹. Di norma, l'obiettivo della ricerca economica ed econometrica è quello di stimare tali relazioni allo scopo di poter effettuare previsioni.

Benché i modelli econometrici raggiungano una notevole complessità, la loro struttura tradizionale di base è, ancora una volta, il modello di regressione. Lo scopo per cui i modelli econometrici venivano stimati, ossia quello di consentire previsioni, fu una forte motivazione a definire modelli il più possibile accurati ed affidabili, e questo diede un forte impulso alla ricerca metodologica nell'ambito dell'analisi di regressione. In particolare, la ricerca econometrica si dedicò all'analisi del residuo⁵², nella consapevolezza che il residuo può rappresentare ciò che sfugge al modello, e soprattutto le caratteristiche dei dati e delle loro relazioni che violano le assunzioni del modello stesso. Analizzando il residuo è quindi possibile individuare quelle caratteristiche dei dati, e di conseguenza affinare il modello con l'uso di metodi correttivi che rettificano la violazione delle ipotesi ripristinando le

⁵⁰ Anselin (2010, pp. 9-10): «I see the origins of the field [of spatial econometrics] as coming from two important sources. One dates back to the quantitative revolution in geography, with the stage set by the important Berry and Marble book on spatial analysis (Berry and Marble 1968). [...] The second origin stems from work in regional science and regional and urban economics that reflected a need to incorporate spatial effects into operational models. Interestingly, the main statement on regional science methods of the time, i.e., Walter Isard's *Methods of Regional Analysis* (Isard 1960, pp. 19-27), only briefly mentions regression and analysis of covariance, but otherwise does not touch upon statistical concerns related to model parameters. Early approaches towards explicit spatial methods were reflected in Granger (1969; 1974) and Fisher (1971). The latter dealt with "Econometric Estimation with Spatial Dependence"».

⁵¹ Si veda ad esempio Theil (1971).

⁵² Johnston e DiNardo (2000).

proprietà delle stime che, è importante ribadirlo, sono l'obiettivo principale della disciplina⁵³.

Anche i processi temporali di cui tradizionalmente si occupano l'economia e con essa l'econometria⁵⁴, tendono a manifestare le proprietà di autocorrelazione e eterogeneità, pur con importanti differenze, rispetto ai processi e ai dati spaziali⁵⁵. L'autocorrelazione tra eventi temporali è abbastanza intuitiva: eventi che accadono in momenti ravvicinati nel tempo tendono ad essere più simili di eventi che accadono ad una distanza temporale maggiore (autocorrelazione temporale). Basta pensare alla temperatura media giornaliera o al prezzo di un bene entro un dato arco temporale. Altrettanto intuitiva è l'esistenza di eterogeneità temporali (ossia variabilità nel tempo). Si pensi, di nuovo, alle variazioni della temperatura media giornaliera o al prezzo di un bene entro un dato arco temporale. Quest'ultima proprietà, la non stazionarietà, è ancora più centrale nella ricerca economica ed econometrica, poiché, oltre a violare alcune ipotesi del modello di regressione, essa viola le condizioni di stazionarietà, che sono un presupposto imprescindibile per effettuare previsioni tramite questo apparato modellistico⁵⁶.

Autocorrelazione ed eterogeneità erano dunque note all'econometria, che aveva ben chiare anche le loro conseguenze sulle stime dei modelli: aumento della varianza, e quindi dell'incertezza associata a tali stime, che ne indebolisce l'affidabilità nei modelli previsionali. Per questa importantissima ragione, l'econometria aveva approntato metodi sia per individuare e misurare l'autocorrelazione e l'eterogeneità nei residui dei modelli di regressione, sia per rettificare il modello, così da ripristinare l'affidabilità delle sue stime, e con essa la sua capacità di fornire una base solida per effettuare previsioni. Tali previsioni, come si è visto, cercavano, in quanto poggiavano su serie temporali, di prevedere e quindi poter anticipare, le dinamiche future di quelle

⁵³ Gujarati (2003).

⁵⁴ Granger (1981).

⁵⁵ Peuquet (1994, p. 443): «Any effective spatiotemporal representation must take the special properties of space and time into account, yet space and time can be viewed in many ways».

⁵⁶ Breusch e Pagan (1979).

serie e per questo la ricerca econometrica si occupò principalmente e inizialmente di autocorrelazione ed eterogeneità temporali⁵⁷.

I metodi econometrici furono applicati a varie branche dell'economia, e tra queste anche all'economia regionale che, occupandosi di regioni, come porzioni di sistemi territoriali più ampi, si avvicina ed incrocia con la geografia per l'interesse specifico ai dati spaziali, alla loro distribuzione ed ai processi che essi rappresentano⁵⁸. Va da sé che l'economia regionale s'interessa, ove possibile, ai processi spazio-temporali, la cui complessità è di molto superiore a quella dei processi spaziali e temporali (presi singolarmente), poiché autocorrelazione e eterogeneità si manifestano anche nei processi spazio-temporali, ove non si limitano a sommarsi linearmente, ma si accumulano e moltiplicano con modalità complesse e difficili da modellizzare. È dunque in questo ambito che, sulla scia dei metodi per il trattamento di autocorrelazione ed eterogeneità temporali, matura la riflessione sull'autocorrelazione spaziale, e appaiono dei metodi per la sua misurazione ed il suo trattamento statistico⁵⁹.

L'inizio della storia dell'econometria spaziale generalmente si fa coincidere con la pubblicazione del lavoro di Jean Paelnick e Leo Klaassen (1979)⁶⁰, intitolato appunto *Spatial Econometrics* e dedicato all'applicazione di metodi econometrici ai dati spaziali⁶¹. La ricerca metodologica e applicativa si protrarrà senza sosta nei decenni successivi, stimolata anche dalla crescente disponibilità di dati, ultimi in ordine cronologico i ricchissimi *database* spazio-temporali relativi alla pandemia da COVID-19⁶². Molta della ricerca è intesa ad adattare le metodologie econometriche, costruite per i dati temporali, ai dati ed ai processi spaziali. Oltre la componente applicativa, la ricerca affronta questioni metodologiche e quantitative, con importanti riflessioni teoriche, spesso profondamente innestate nella statistica.

Progressivamente le due letterature, *spatial econometrics* e *spatial statistics*, tendono a diversificarsi, assumendo proprie identità

⁵⁷ Theil (1971).

⁵⁸ McCann e Mudambi (2005).

⁵⁹ Anselin (2010).

⁶⁰ Anche se Anselin (2010, p. 10) identifica lavori precedenti: «Early approaches towards explicit spatial methods were reflected in Granger (1969, 1974) and Fisher (1971)».

⁶¹ Il filone si consoliderà poi lungo tutti gli anni '80 e '90, con gli interventi più notevoli e influenti di Anselin (1988), poi di Haining (1990) e Cressie (1993).

⁶² Krisztin, Piribauer, Wögerer (2020).

distinte⁶³. Daniel Griffith e Jean Paelinck (2007) affermano che esse riflettono le tradizioni delle rispettive discipline d'origine, e quindi hanno molto in comune, ma con importanti differenze di enfasi⁶⁴.

Lo scostamento tra *spatial statistics* e *spatial econometrics* farà slittare l'analisi spaziale verso un approccio più *data driven*, ossia centrato sull'utilizzo di un insieme di tecniche e di metodi intesi a saggiare possibili relazioni tra i dati a disposizione, senza che vi siano necessariamente delle ipotesi vincolanti, né che le indagini facciano riferimento a teorie accettate⁶⁵.

L'approccio *data driven* aprì all'analisi spaziale la possibilità di applicazioni in contesti dove l'inquadramento teorico era più debole⁶⁶, consentendo sia fertili interazioni con altre discipline e applicazioni, sia l'affinamento delle tecniche analitiche, rendendole più flessibili e capaci di affrontare contesti eterogenei, come per esempio l'analisi dei dati *raster*⁶⁷. Per contro, essa porterà con sé anche la tendenza all'utilizzo delle tecniche analitiche spaziali in maniera esplorativa (*data mining*)⁶⁸ e sempre più distante da ipotesi e teorie, forse anche inconsapevolmente⁶⁹, pur tuttavia esponendo l'analisi spaziale a critiche dall'interno della geografia.

⁶³ Cressie (1993), Anselin (2000) e Haining (2003).

⁶⁴ Griffith e Paelinck (2007, p. 210 corsivi miei): «In practice, spatial econometrics and spatial statistics reflect traditions of their respective parent disciplines, namely econometrics and statistics. In other words, they share much in common, with some notably differing emphases». Essi poi, nella Table 1, elencano i problemi centrali delle due discipline, individuandone 12 in comune, 8 specifici della statistica spaziale, e 4 specifici dell'econometria spaziale: «*Central problems in spatial statistics and spatial econometrics*: Super population perspective (i.e., realizations from a theoretical population); model-based inference; Properties of estimators; Specification of geographic neighborhood structure; Modifiable areal unit problem (MAUP); Quantifying spatial autocorrelation; Variable transformations: Box–Cox, Box–Tidwell; Spatially adjusted statistical techniques; Cluster detection: hot and cold spots; LISA statistics; Distance as a covariate; Bayesian hierarchical models; Exploratory spatial data analysis; Space-time modeling. *Central problems in spatial statistics*: Sampling network structure; design-based inference; Ecological fallacy; Map generalization: spatial interpolation; Missing spatial data imputation; Auto-model specification: normal, Poisson, binomial; Spatial structure as a covariate (spatial filtering); Bayesian smoothing of map values; Error propagation. *Central problems in spatial econometrics*: Constrained parameter estimation; Optimization models; Endogenous versus exogenous variables; Spatial complexity and spatial regimes».

⁶⁵ Questa lenta svolta avrà molteplici conseguenze per l'analisi spaziale, la cui portata è ben lungi dall'essere esaurita, che verranno sviluppate successivamente in questo volume.

⁶⁶ Anselin (2012).

⁶⁷ Boots e Tiefelsdorf (2000).

⁶⁸ Li *et Alii* (2015).

⁶⁹ Miller e Goodchild (2015).

Nel già citato “*Workshop on Status and Trends in Spatial Analysis*” del 1998, Luc Anselin (2000) delinea il rapporto tra GIS, econometria spaziale e scienze sociali, tracciando linee di sviluppo per una comunità scientifica la cui consistenza numerica andava già allora aumentando, che aveva accettato, se non fatto propri, gli strumenti del GIS e dell’analisi spaziale. Così, come dimostra il poderoso lavoro di Noel Cressie (1993), allontanandosi dall’econometria spaziale ed avvicinandosi alla statistica spaziale, l’analisi spaziale manterrà un fortissimo legame con la statistica, e, seppur *data driven*, rimarrà radicata nella teoria statistica nell’utilizzo e l’elaborazione metodologica dell’inferenza, nella definizione e nel controllo dell’errore persino in GIS⁷⁰. Bisognerà attendere la fine degli anni ’90 perché si trovino concetti e metodi per misurare la non-stazionarietà o eterogeneità spaziale e si approntino metodi per affrontarne le conseguenze⁷¹. Come si vedrà più avanti, anche in questo caso le soluzioni saranno adattate da quelle sviluppate dalla teoria economica.

1.5. Le tecnologie e gli strumenti dell’informazione geografica

Nei paragrafi precedenti, in particolare nel contesto dei cenni storici sul GIS e sulla geografia teoretico-quantitativa, si è parlato a lungo dell’importanza della tecnologia nello sviluppo, anche recente, di quelle che sono dette, appunto, tecnologie dell’informazione geografica⁷². Benché sia evidente che l’analisi spaziale è qualcosa di molto diverso dalla mera tecnologia, è tuttavia importante riconoscere che non può prescindere, e che il passo del suo stesso sviluppo è stato a lungo condizionato dalla disponibilità (o meno) di dati ed adeguati strumenti computazionali. Inoltre, come è già avvenuto per altre discipline, tra cui il GIS, la tecnologia potrebbe in futuro presentare delle opportunità anche per l’analisi spaziale e, suo tramite, per la geografia. Ci soffermiamo qui su alcuni sviluppi rilevanti per l’analisi spaziale, senza pretesa di esaurire la gamma di sviluppi tecnici e tecnologici degli ultimi decenni.

⁷⁰ Burrough (1986).

⁷¹ Fotheringham, Charlton, Brunson (1998).

⁷² Waters (2017).

Per le tecnologie dell'informazione geografica, il nuovo millennio si aprì all'insegna dell'ottimismo, particolarmente negli Stati Uniti, con l'annuncio del progetto Digital Earth⁷³: una proposta innovativa per la costruzione di un mondo digitale, che poteva ospitare, al suo interno, rappresentazioni, condivisione, idee scientifiche e pedagogia, il cui paladino fu l'allora senatore Al Gore⁷⁴. Nel frattempo, si andavano consolidando i recenti progressi delle scienze affini; basti pensare alla geodesia e alla realizzazione del datum geodetico WGS84 (World Geodetic System 1984)⁷⁵, su cui ancora oggi poggia il sistema GPS (Global Positioning System). È ben vero che quest'ultimo era già stato sviluppato negli Stati Uniti e in Europa⁷⁶, ma la sua diffusione in ambito civile è della fine degli anni '90, non a caso quando, nel 2000, la difesa americana rimosse l'errore introdotto artificialmente negli anni '90, proprio per limitarne l'efficacia nelle applicazioni civili⁷⁷. A tutto questo contribuì anche la creazione di nuove reti satellitari, che consentono di migliorare sostanzialmente la capacità di posizionamento. In particolare, va ricordato il sistema europeo Galileo, lanciato nel 2005⁷⁸, seguito dal cinese BeiDou⁷⁹.

La nuova precisione ed accuratezza dei sistemi di posizionamento globale ne consentirono un uso sempre più sofisticato, che si innestava, come già osservato, con gli impieghi innovativi degli strumenti GIS. Se l'installazione di navigatori nelle automobili civili è ormai la norma, applicazioni sempre più sofisticate vengono realizzate in ambiti già consolidati, come la meteorologia e lo studio del clima, spingendosi fino allo studio dell'inquinamento atmosferico⁸⁰, e favorendo applicazioni sempre più innovative in agricoltura, ove consentono tra l'altro di ottimizzare l'uso di fertilizzanti, riducendo consumi e inquinamento e nel contempo aumentando la produzione agricola⁸¹.

⁷³ <https://digitalearth2021.org/>

⁷⁴ Sconfitto alle presidenziali del 2000, il senatore non abbandonò la passione per il pianeta, abbracciando la causa della lotta al cambiamento climatico (Gore, 2006).

⁷⁵ Kumar (1988).

⁷⁶ In particolare, il francese SPOT, *Système Probatoire pour l'Observation de la Terre*, del 1986, si veda Parkinson *et Alii* (1996).

⁷⁷ La cosiddetta “*selective availability*”, si veda Adrados *et Alii* (2002).

⁷⁸ Benedicto *et Alii* (2000).

⁷⁹ Cai *et Alii* (2015).

⁸⁰ Evans *et Alii* (2013).

⁸¹ Gebbers e Adamchuk (2010).

Permangono applicazioni di tipo militare, che ora comprendono anche le operazioni di *peace keeping*, *search and rescue* e *disaster management*⁸². Parallelamente a tutte queste applicazioni è importante lo sviluppo di sempre migliori strumenti per la definizione e la rappresentazione cartografica, che consentono la realizzazione di proiezioni sempre più in grado di controllare e minimizzare la distorsione⁸³.

Infine, d'importanza fondamentale per l'analisi spaziale è lo sviluppo del *software* per l'analisi dei dati spaziali, più spesso però indirizzato al più ampio mercato del GIS. Se negli ultimi decenni del secolo scorso era presente sul mercato una pluralità di prodotti commerciali, da Idrisi⁸⁴ a Geomedica⁸⁵, negli ultimi decenni la Esri⁸⁶ si impone sempre più come il leader pressoché monopolistico del mercato del GIS. L'ultimo prodotto di Esri è ArcGIS Pro che si rivolge ad un mercato professionale d'alto livello e ambisce a fornire capacità analitiche statistiche avanzate avvalendosi del linguaggio Python⁸⁷, nel tentativo di attrarre anche il potenziale bacino di utenza dell'analisi spaziale.

Notiamo qui infine il successo e la diffusione del prodotto commerciale di Esri Storymap⁸⁸, che, come suggerisce il nome, consente di raccontare storie attraverso mappe. Cariche di potenzialità, pur non prive di limitazioni, le *storymaps* si vanno affermando, trainando con sé diffusione, popolarità e consapevolezza della mappa, e con essa del GIS e della geografia, nella comprensione e comunicazione di svariati fenomeni⁸⁹. Tuttavia, va anche osservato che, mentre sul fronte del *software* commerciale per il GIS si afferma il monopolio di Esri, al suo esterno cresce la diffusione dei *software freeware* e *open source*,

⁸² De Albuquerque *et Alii* (2015).

⁸³ Šavrič, Patterson, Jenny (2019).

⁸⁴ <https://clarklabs.org/terrset/idrisi-gis/>

⁸⁵ www.hexagongeospatial.com/products/power-portfolio/geomedica

⁸⁶ Environmental Systems Research Institute di Redlands, California, a tutt'oggi guidato da Jack Dangermond (www.esri.com/en-us/home).

⁸⁷ www.python.org/

⁸⁸ <https://storymaps.arcgis.com/>

⁸⁹ È possibile osservare una notevole varietà di *storymaps* sul sito di Esri (<https://storymaps-classic.arcgis.com/en/gallery/#s=0>). Benché la valutazione sia necessariamente soggettiva, sottolineiamo l'efficacia dell'uso dello strumento nella *storymap* che illustra la crisi da uso di oppioidi, su <https://urbanobservatory.maps.arcgis.com/apps/Cascade/index.html?appid=f86499d99e4340b68229eacfb02b29f> (visitato il 18/10/2021), quella sulla diffusione del virus Zika su <https://bureau.maps.arcgis.com/apps/MapJournal/index.html?appid=9e9ca7c6957f4616a05a4331f99a0c22-> (visitato il 18/10/2021)

il più noto dei quali è Quantum GIS, o QGIS⁹⁰, che offre potenzialità sempre più vicine a quelle del *software* commerciale. A questi va aggiunta la crescente disponibilità di dati, i cosiddetti *open data*, su cui si tornerà nei prossimi capitoli.

Il fenomeno del *software freeware* e *open source* investe direttamente l'analisi spaziale, poiché proprio *freeware* ed *open source* è il pacchetto che fornisce le maggiori potenzialità e le routine più avanzate per l'analisi statistica dei dati spaziali, interfacciandosi anche con i più diffusi pacchetti GIS. Si tratta di R⁹¹ o RStudio⁹², di cui esiste anche una versione commerciale. Diversamente dal mercato del GIS, l'*open source* R è di fatto il pacchetto leader dell'analisi statistica spaziale, senza validi rivali tra i *software* commerciali, nonostante la vasta gamma di prodotti statistici presenti in commercio⁹³. Anche nella sua versione più *user-friendly*, RStudio, R rimane un *software* “*command-driven*”, in cui l'utente deve scrivere, compilare ed eseguire le varie analisi utilizzando stringhe di comando o aggregandole in *scripts*. Si tratta quindi di *software* accessibili da un punto di vista economico, ma che richiedono un minimo di preparazione di base, o un investimento nell'apprendimento, generalmente superiore a quanto richiesto dai più comuni pacchetti “*menu-driven*” generalmente in commercio. Notiamo che questa necessità spesso costituisce, nella percezione di molti, un ostacolo, che ne inibisce l'avvicinamento all'analisi spaziale.

A questa breve panoramica aggiungiamo alcuni pacchetti *software* altamente specializzati, per l'esecuzione di forme avanzate di analisi spaziale, distribuiti a partire dai primi anni di questo secolo. Tra essi spiccano *GWR*, realizzato da Stuart Fotheringham e colleghi, per la stima di modelli di regressione *GWR*, un acronimo che sta per *geographically weighted regression*. *GWR* fu inizialmente commercializzato, e poi distribuito come *freeware*. L'ultima versione, *GWR4*, è stata recentemente aggiornata con *MGWR*, ossia *multiscale geographically weighted regression*, ora distribuito dalla *School of Geographical Sciences and Urban Planning della Arizona State*

⁹⁰ www.qgis.org/en/site/

⁹¹ www.r-project.org/

⁹² <https://rstudio.com/>

⁹³ Questi vanno dal classico SPSS a SAS, STATA, etc., nessuno dei quali offre capacità analitiche *spaziali* comparabili a quelle di R.

University⁹⁴. Analogamente GeoDa⁹⁵, sviluppato da Luc Anselin e distribuito fin dall'inizio come *freeware*, svolge una serie di funzioni di analisi spaziale e locale, da quelle descrittive fino a regressioni di tipo spaziale, come la *GWR* e i modelli autoregressivi spaziali (*SAR*). Questa situazione ha delle implicazioni importanti per lo sviluppo dell'analisi spaziale, su cui si tornerà nei capitoli successivi.

1.6. L'analisi spaziale nella geografia

Fin qui si sono individuate le origini dell'analisi spaziale in alcuni nuclei distinti, dei quali si sono ripercorse le tappe salienti. Ci chiediamo a questo punto se si possano intrecciare i fili di queste storie, individuandovi alcuni punti di riflessione che aiutino a situare l'analisi spaziale nel presente e nel futuro della geografia.

Dov'è finita l'analisi spaziale? si chiedeva Ron Johnston (2000), notandone la relativa marginalità secondo una sua lettura della storiografia della geografia⁹⁶. L'aspetto che qui ci preme sottolineare non è la lotta per il predominio disciplinare o l'opinione secondo cui la storia degli *spatial analysts* sarebbe stata relegata in riviste di minore risonanza rispetto alle posizioni occupate dai *social theorists*⁹⁷; ciò che più importa è che da queste riflessioni emerge, innegabile, l'affermazione dell'ininterrotta vitalità dell'analisi spaziale nella geografia, persino nei momenti in cui le circostanze le erano meno favorevoli⁹⁸. La scarsa

⁹⁴ <https://sgsup.asu.edu/sparc/mgwr>

⁹⁵ <https://spatial.uchicago.edu/software>

⁹⁶ Johnston (2000, p.125): «Analysis of a small number of recent texts shows that they reflect a current contest within human geography between two groups, stereotyped as “spatial analysts” and “social theorists”. The former are being “written out” of disciplinary history, despite their continued vitality».

⁹⁷ Secondo Johnston (2000), alla rivoluzione teoretica e quantitativa degli anni '50 e '60 avrebbe fatto seguito una disillusione per il quantitativismo, che non era riuscito a produrre un mondo migliore, e che sfociò, pochi decenni dopo la controrivoluzione della teoria sociale. Si vedano anche Livingstone (1992, pp. 304-346) e Pickles (1995a, pp. 1-30).

⁹⁸ Ne sarebbe seguita una lotta per il predominio disciplinare tra *spatial analysts* e *social theorists* in cui i secondi avrebbero avuto la meglio, relegando i primi a scrivere la propria storia in riviste di minore risonanza. Tant'è che ripetute affermazioni secondo cui era stata la geografia quantitativa a salvare la geografia dall'estinzione come disciplina scientifica furono fondamentalmente ignorate. Johnston (2000, p. 128): «In the former [*Urban Geography*], for example, Getis (1993, 517) claimed that quantitative geography “[...] is the key to the survival of geography as a discipline”, Morrill (1993, 442) having previously argued that it “[...] saved

presenza dell'analisi spaziale nella storiografia della geografia spine anche Trevor Barnes (2010), ad interrogarsi provocatoriamente sul suo stato di salute, ma anche dalla sua riflessione non emergono che rassicurazioni sulla continua vitalità dell'analisi spaziale e del quantitativismo in geografia, sostenuto anche dalla vitalità della scienza dell'informazione geografica (*GISci*)⁹⁹.

Nella riflessione presentata nel 1998 al “*Workshop on Status and Trends in Spatial Analysis*”, Michael Goodchild (2000) affermava che, benché proprio l'analisi spaziale sia stata la vera motivazione che portò alla nascita del GIS, in seguito il GIS si allontanò progressivamente dalla linea del rigore dell'analisi, prendendo una direzione che egli stesso definisce populista. Egli articola quindi la complessità del rapporto tra GIS e analisi spaziale in quattro tensioni: tra populismo ed elitismo, tra visuale e numerico, tra aperto e chiuso, e tra locale e globale. La prima tensione si gioca in termini di accessibilità, e contrappone il GIS, aperto ed accessibile a chiunque, all'analisi spaziale, riservata ai soli iniziati e perciò ritenuta elitaria¹⁰⁰. Alcuni sviluppi del *software*, in epoche relativamente recenti, hanno tentato di ridurre questo divario utilizzando strategie che potremmo dire bilaterali: da un lato il *software* analitico propone analisi rigorose attraverso interfacce relativamente *user-friendly*¹⁰¹, mentre dall'altro il GIS¹⁰² spinge i propri menù tradizionalmente intuitivi e accattivanti ad offrire analisi spaziali sempre più sofisticate. Si tratta, a nostro avviso, di risposte parziali e ancora superficiali, che non risolvono le altre tensioni, ossia quelle tra la centralità della mappa (visuale), propria del GIS, contrapposta a quella analitica (numerica) propria dell'analisi spaziale, e quella relativa all'opportunità o meno di dichiarare (apertamente) i passaggi matematici e le scelte computazionali dell'analisi, in contrapposizione al confezionare l'analisi in un pacchetto (chiuso), la

geography from extinction as a serious university discipline [...] by gaining a foothold in science at large [...]”. But these drew virtually no response (to a considerable extent because they were directed at relatively restricted audiences)».

⁹⁹ L'acronimo *GSci* viene spesso usato in riferimento a *geographic information science*, in contrapposizione a *geographic information systems*. Su questa definizione torneremo nel capitolo 4.

¹⁰⁰ Per una riflessione su questo punto si veda anche Bertazzon (2001; 2005).

¹⁰¹ L'esempio migliore è GeoDa di Luc Anselin (<https://spatial.uchicago.edu/software>).

¹⁰² Qui l'esempio è ancora ArcGIS Pro (www.esri.com/en-us/arcgis/products/arcgis-pro/overview).

cosiddetta *black box*, di cui si conoscono solo i dati immessi e i risultati ottenuti. Anche qui la tensione emerge dalla possibilità di rendere trasparenti i meccanismi e i parametri dell'analisi, consentendo a chiunque di replicarne i risultati, ed è riconducibile, come la precedente, a quella tra elitismo e populismo.

L'ultima delle tensioni individuate (il rapporto tra locale e globale) riporta invece al relativismo dell'analisi spaziale, alla presa di coscienza della relatività della validità dell'analisi spaziale, che conduce all'importante dibattito tra nomotetico e idiografico, tra la ricerca di leggi universali e l'analisi dettagliata a grande scala, ed è qui che Michael Goodchild intravede una terza via, quella che delinea relazioni che possono sì essere universali, ma con caratteristiche specifiche che variano localmente¹⁰³. Sull'importanza dell'analisi locale si era già soffermato Paul Longley (2000), per il quale tuttavia non si tratterebbe di una tensione, ma di una potenzialità, che potrebbe consentire un superamento della tradizionale dicotomia tra nomotetico e idiografico.

Molto più di recente Harvey Miller (2018) definisce il GIS (o meglio la *GIScience*) e l'analisi spaziale come i figliastri (*stepchildren*) della vecchia e criticata fisica sociale. Inoltre, rievocando anch'egli la classica dicotomia tra nomotetico e idiografico, definisce una microgeografia, dettagliata e descrittiva, in contrasto con una macrogeografia, alla ricerca di leggi universali e incurante del contesto. Secondo questo autore, il percorso storico dei due figliastri, GIS e analisi spaziale, potrebbe far emergere quella che egli definisce una mesogeografia, in posizione intermedia tra microgeografia e macrogeografia. Il fulcro della mesogeografia sarebbe il ruolo del contesto spaziale nel determinare il comportamento di sistemi collettivi come città e regioni. La mesogeografia infatti ricerca generalizzazioni, non leggi universali, ossia relazioni causali valide entro una gamma definita di ambienti. Harvey Miller si spinge poi ad affermare che la *GIScience* è ben posizionata per produrre sapere mesogeografico, grazie alla sua capacità di perseguire sia leggi generali che descrizioni, le prime tramite algoritmi e metodi, le seconde tramite l'enfasi sui dati; il tutto grazie anche alla flessibilità consentita dalla dinamica tra analisi globale e locale.

¹⁰³ La questione, influenzata anche dall'allora nascente analisi locale (Anselin, 1995; Ord e Getis, 1995), sarà dibattuta ampiamente nel capitolo 3.

Per quanto riguarda infine il rapporto tra analisi spaziale e geografia, Paul Longley (2000) vi si sofferma analizzando il triangolo definito da analisi spaziale, GIS e geografia. Egli nota che mentre la rivoluzione quantitativa pervase la geografia e si protrasse per un quarantennio, l'infusione del GIS nella geografia fu rapida, e la geografia ne fu più utente che autrice. Per questo non è chiaro se la parabola del GIS sia (stata) positiva o negativa per la geografia: da un lato potrebbe aver contribuito in parte al suo "*de-skilling*" ossia all'impoverimento delle sue competenze, tuttavia la capacità del GIS di gestire quantità enormi e crescenti di dati potrebbe essere importante per lo sviluppo futuro della geografia. Grazie alle sue caratteristiche di disciplina populista e accattivante, il GIS potrebbe anche contribuire a divulgare l'analisi spaziale, forse più ostica ma certo capace di analisi rigorose che possono fornire soluzioni innovative a problemi importanti. Secondo Barry Boots (2000), proprio per questa ragione il GIS potrebbe svolgere un ruolo strumentale nel promuovere l'analisi spaziale tra gli scienziati fisici e sociali.

La storia dell'analisi spaziale che emerge da questo capitolo è dunque fatta di radici profonde in discipline diverse, di una presenza forse troppo discreta nella storia della geografia, e di un rapporto complesso con la sua controparte più prorompente, il GIS, con i suoi utenti, sia attuali che potenziali, e con la geografia. Ma da tutti questi aspetti, in qualche modo problematici, emergono l'importanza e le potenzialità di questa disciplina, capace di giocare un ruolo irrinunciabile in un futuro in cui la geografia dovrà misurarsi con grandi moli di dati, e in cui sarà chiamata ad utilizzare strumenti per la loro gestione [GIS] e a svolgere un ruolo di protagonista nella loro analisi, grazie al proprio rigore, alla profondità analitica, alla capacità di gestione e controllo dell'errore che le vengono dalle proprie intricate radici che affondano, tra l'altro, nella statistica. Tuttavia, i problemi individuati mostrano anche che il successo dell'analisi spaziale e della geografia stanno anche nella capacità di dialogo interdisciplinare entro e oltre la geografia.

2. Oggetto e metodi dell'analisi spaziale: cenni

Formula itaque art. praec. sponte perducit ad egregium THEOREMA.
Gauss, *Disquisitiones generales circa superficies
curvas*, 1827¹.

2.1. Premessa

Il discorso tracciato in questo volume, benché in larga parte teorico, poggia anche su alcuni concetti matematici e statistici che costituiscono il nocciolo dell'analisi spaziale. Comprendere questi concetti, anche assaporandone le problematiche statistiche, è importante per afferrare il senso profondo dell'analisi spaziale, e quindi anche i suoi aspetti teorici e la sua vicenda storica. Per questi motivi un discorso sull'analisi spaziale non può prescindere da alcune definizioni formali e riferimenti a concetti statistici chiave. Tuttavia, poiché questo volume ambisce rivolgersi potenzialmente all'intera comunità di geografi, si è optato per una sorta di duplice esposizione dei contenuti di questo capitolo: nei paragrafi (2.1 – 2.4) si sono dipanati i concetti in un linguaggio narrativo piano, rimuovendo definizioni formali e formule matematiche, che sono state raccolte nel paragrafo 2.5. I paragrafi (2.1 – 2.4) forniscono comunque richiami puntuali alle formule presentate nel paragrafo 2.5.

In questo capitolo, sia i paragrafi (2.1 – 2.4) che il paragrafo 2.5 sono suddivisi in numerosi sotto-paragrafi, per consentirne una consultazione rapida anche di richiamo ai concetti discussi negli altri capitoli. Analogamente, il paragrafo 2.5 riproduce struttura e titoli dei paragrafi (2.1 – 2.4), per facilitarne la consultazione incrociata. Le parti narrative del paragrafo 2.5 presentano necessariamente qualche ridondanza con quelle dei paragrafi (2.1 – 2.4): tale ridondanza riflette

¹ Citato in M. Andreatta, *Archimede, l'arte della Misura*. Bologna, Il Mulino, 2021. p. 13.

anche l'intento di rendere le due parti il più possibile autonome l'una dall'altra. Per la lettura di questo capitolo dunque ogni lettore può scegliere, in base alla propria preparazione e preferenza, tra il linguaggio narrativo (2.1 – 2.4), quello formale matematico-statistico (2.5), o un'integrazione dei due.

Si osservi infine che il materiale presentato in questo capitolo è limitato allo scopo di questo volume, quindi non fornisce concetti base (ad esempio media, varianza, etc.), né si spinge oltre le principali definizioni delle problematiche e delle tecniche qui esposte.

2.2. La “prima legge della geografia”

All'inizio degli anni '70 Waldo Tobler² coniò, in un saggio sulla simulazione computerizzata della crescita dell'area metropolitana di Detroit, l'espressione “prima legge della geografia” che recita:

I invoke the first law of geography: everything is related to everything else, but near things are more related than distant things³.

Da allora quella relazione porta il nome di “legge”, spesso anche legato al suo nome, “legge di Tobler”, o anche “prima legge di Tobler” (Tobler's First Law, o TFL). Nei decenni successivi alcuni hanno persino enunciato una seconda⁴, e pure una terza legge, sulla validità delle quali si può vedere ad esempio l'interessante analisi di Michael Goodchild (2004).

² Il testo di Tobler (1970) si può trovare in:

<https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118786352.wbieg1011>

³ Tobler (1970, p. 236).

⁴ Foresman e Luscombe (2017, p.982) «Second law of geography: Things that know where they are can act on their locational knowledge. Spatially enabled things have increased financial and functional utility». Zhu, Lu, Liu, Qin, Zhou (2018, p. 225): «Current methods of spatial prediction are based on either the First Law of Geography or the statistical principle or the combination of these two. The Second Law of Geography contributes to the revision of these methods so they are adaptive to local conditions but at the cost of increasing demand for samples. This paper presents a new thinking about spatial prediction based on the Third Law of Geography which focuses on the similarity of geographic configuration of locations. Under the Third Law of Geography, spatial prediction can be made on the basis of the similarity of geographic configurations between a sample and a prediction point».

Al di là del contenuto e valore di queste “leggi”, è interessante notare come se ne continui a parlare ben 50 anni dopo l’enunciato di Waldo Tobler⁵ e, in ogni caso, rimane innegabile che, dietro una facciata un po’ pomposa, essa cela un contenuto importante: evidenzia come i rapporti e le reazioni di vicinanza territoriale/spaziale rappresentino le proprietà basilari che contraddistinguono i processi spaziali e costituiscono la ragion d’essere dell’analisi spaziale⁶.

2.2.1. I processi spaziali

Intuitivamente, col termine processo spaziale si designa qualsiasi fenomeno che si dispiega sullo spazio geografico, come ad esempio la temperatura o la vegetazione, la popolazione residente o l’incidenza di una malattia. In termini più rigorosi, un processo spaziale è un modello di variazione spaziale (Equazione 1)⁷ che rappresenta un fenomeno, continuo o discontinuo, osservato su un insieme di punti dello spazio, e descritto da un insieme di triplette che per ciascun punto campionato ne indicano le coordinate geografiche⁸ e l’attributo, ossia il valore che il fenomeno considerato (ad esempio l’incidenza di una malattia) assume in quel punto. I processi spaziali sono caratterizzati dalla proprietà che in essi nessun valore nello spazio è indipendente dagli altri valori. Formalmente, questa proprietà è espressa come covarianza non-nulla del processo (Equazione 2). Questa covarianza non-nulla diventa interessante se interpretata in termini di quelli che sono generalmente detti effetti spaziali: struttura spaziale (eterogeneità spaziale), interazione spaziale (autocorrelazione spaziale) e ordine spaziale delle

⁵ Si veda anche il numero speciale di *Geographical Analysis* per commemorare Waldo Tobler in occasione della sua recente scomparsa ed in particolare il testo introduttivo di Clarke e Murray (2020).

⁶ Tra i primi a formalizzare il concetto di processi spaziali furono Cliff e Ord (1981), a cui seguirono i lavori dei primi anni ’90, radicati anche nella statistica e nell’econometria, di Anselin (1988), Haining (1990) e Cressie (1993).

⁷ Come anticipato nella premessa a questo capitolo, le relazioni matematico-statistiche sono qui espresse dalle equazioni contenute nel paragrafo 2.5 e numerate progressivamente. Sia per questa equazione che per le successive si rimanda quindi al paragrafo 2.5.

⁸ Generalmente le coordinate geografiche sono almeno due (spesso latitudine e longitudine), ma potrebbero aumentare se si considera l’altitudine, di norma misurata rispetto al livello del mare. In taluni processi spazio-temporali, il tempo può essere rappresentato da una quarta coordinata.

osservazioni⁹. Si può facilmente osservare che la covarianza non-nulla tra valori del processo spaziale è semplicemente un modo diverso per esprimere la dipendenza, ossia una definizione più rigorosa di quanto espresso informalmente dalla prima parte della legge di Tobler: ogni cosa è collegata ad ogni altra cosa. La seconda parte della legge di Tobler si può invece esprimere più rigorosamente con una delle numerose funzioni di attrito della distanza (*distance decay function*) (Equazione 3), definite già durante la rivoluzione quantitativa¹⁰.

2.2.2. *L'autocorrelazione spaziale*

Le definizioni di autocorrelazione spaziale¹¹ e di dipendenza spaziale sono da lungo tempo oggetto di dibattito¹². Secondo Arthur Getis¹³ il punto cruciale è relativo al fatto che l'autocorrelazione spaziale è un termine, per sua natura geometrico, che misura, in riferimento ad una singola variabile, come certe unità spaziali sono correlate (positivamente) o inversamente-correlate (negativamente) ad altre unità spaziali e come vi è dipendenza tra unità spaziali. Esistono diversi indici di autocorrelazione spaziale, di cui i più noti sono quelli di Moran e di Geary. In questo testo facciamo riferimento all'indice di Moran, la cui formulazione (Equazione 4) ne consente un'interpretazione intuitiva: secondo tale indice l'autocorrelazione è positiva quando gli attributi vicini sono più simili (attrazione), negativa quando questi sono più dissimili (repulsione), o prossima allo zero, quando non vi è relazione tra similitudine e distanza, e in quel caso il processo spaziale si configura un processo casuale (*complete spatial randomness*)¹⁴.

⁹ Come nota Anselin (2001, p. 4): «Spatial autocorrelation, or the coincidence of value similarity with locational similarity, can be formally expressed by the moment condition».

¹⁰ Si vedano al riguardo: Costa e Canestrelli (1983, pp. 67-77) ed in particolare Getis e Ord (1992).

¹¹ Si veda l'analisi che ne fa Griffith (1992).

¹² Anselin (2001, p. 3) osserva: «Obviously, the two [spatial dependence and spatial autocorrelation] are not identical, but typically, the weaker form is used [spatial autocorrelation], in the sense of a moment of a joint distribution».

¹³ Getis (2007, p. 493): «The point here is that spatial auto-correlation is a term, geometrical in character, that says that with regard to a single variable certain spatial units are related (positively) or unrelated (negatively) to other spatial units and that there is dependence between spatial units».

¹⁴ Si veda Diggle (1979).

Come discusso ampiamente nel capitolo 3, gli indici di autocorrelazione spaziale forniscono solo un'indicazione della similitudine in rapporto alla distanza, senza tuttavia spiegarne le cause.

2.2.3. *La matrice dei pesi spaziali*

Un elemento fondamentale per la stima dell'autocorrelazione spaziale è la definizione di prossimità, o contiguità utilizzata nel calcolo degli indici di autocorrelazione spaziale (ad esempio, Equazione 4). Anzi, tale definizione è uno degli elementi più critici della stima dell'autocorrelazione spaziale e anche delle tecniche analitiche spaziali definite per alleviarne le conseguenze. Operativamente la contiguità, presente anche nell'indice di Moran (Equazione 4), è definita dalla matrice dei pesi spaziali (*spatial weights matrix*) (Equazione 5).

La matrice dei pesi spaziali è formata di tanti elementi quante sono le unità spaziali campionate e ciascuno dei suoi elementi definisce la contiguità tra due unità spaziali. In genere la matrice è composta di una struttura binaria moltiplicata per una funzione di attrito della distanza. La struttura binaria ha il compito stabilire quali unità siano considerate contigue. La matrice che si ottiene moltiplicando la struttura binaria per la funzione di attrito della distanza pesa l'interazione osservata per la distanza tra due unità spaziali in tutti i casi in cui esse sono definite come contigue. Per questa ragione, la matrice di contiguità è spesso chiamata anche matrice dei pesi spaziali (*spatial weights matrix*)¹⁵.

Esistono molti modi per definire la contiguità, sia in termini binari che pesati¹⁶. Se le unità spaziali in questione sono poligoni, è frequente considerare contigui i poligoni adiacenti, ossia quelli che presentano un lato in comune. Se invece le unità spaziali sono punti¹⁷, questo metodo non è applicabile ed esistono due criteri alternativi: il primo è l'utilizzo di una distanza soglia, entro la quale le unità sono considerate contigue; il secondo è l'utilizzo di un numero (generalmente indicato dalla lettera *k*) di vicini più prossimi (*k nearest neighbours*). I

¹⁵ Si veda Anselin, Murray, Rey (2000).

¹⁶ Come notano Getis e Aldstadt (2004).

¹⁷ Analogamente, anche se meno frequente, è il caso in cui le unità spaziali siano segmenti, come tratti di strade o formazioni geomorfologiche.

metodi della distanza soglia e dei vicini più prossimi sono applicabili anche ad unità spaziali poligonali, di cui si fa riferimento al centroide. Nel caso dei poligoni adiacenti, una formalizzazione frequente dell'attrito della distanza è riportare l'interazione alla lunghezza del lato in comune; negli altri casi si utilizzano funzioni specifiche (Equazione 3)¹⁸. In tutti i casi si tratta di rappresentazioni la cui definizione concettuale ha un impatto rilevante sul calcolo dell'indice.

2.2.4. *L'eterogeneità spaziale*

La seconda proprietà intrinseca ai processi spaziali è nota come eterogeneità (o non-stazionarietà) spaziale. Di nuovo, la definizione intuitiva è molto semplice; parafrasando la Legge di Tobler si potrebbe dire semplicemente che tutte le cose variano nello spazio (e che anche le relazioni tra le cose variano nello spazio). In termini più formali, un processo spaziale si dice spazialmente omogeneo (o stazionario) se soddisfa quattro proprietà (Equazioni 6 – 8)¹⁹:

1. ha media costante nello spazio;
2. ha varianza costante nello spazio;
3. ha covarianza costante nello spazio;
4. è isotropico, ovvero, le proprietà 1 - 3 sono soddisfatte in tutte le direzioni.

Un processo che violi una o più di queste quattro proprietà si dice spazialmente eterogeneo. Benché in questo caso non si parli di leggi della geografia, anche l'eterogeneità spaziale è una caratteristica intrinseca dei processi spaziali, al pari della dipendenza spaziale²⁰. Anche per questa proprietà l'osservazione empirica mostra che la maggior parte dei processi spaziali, siano essi di geografia umana o fisica, sono affetti da eterogeneità spaziale. È infatti alquanto improbabile

¹⁸ Getis e Aldstadt (2004).

¹⁹ Si veda Fotheringham, Charlton, Brunson (1998).

²⁰ Anselin (2001, pp. 3-4): «The latter [spatial heterogeneity] is simply structural instability in the form of non-constant error variances (heteroskedasticity) or model coefficients (variable coefficients, spatial regimes) and can be tackled by means of the standard econometric toolbox. However, there are three reasons why it is important to consider spatial heterogeneity explicitly. First, the “structure” behind the instability is *spatial* (or geographic) in the sense that the location of the observations is crucial in determining the form of the instability».

che un fenomeno geografico soddisfi tutte e quattro le proprietà (si veda anche l'Esempio 2.1).

Una misura quantitativa dell'eterogeneità spaziale è fornita dagli indici di clustering, il più comune dei quali è l'indice G di Getis-Ord, descritto dall'Equazione 9, di cui è meglio nota la versione G_i^* , utilizzata nella cosiddetta *hotspot analysis*.

2.2.5. Relazione tra autocorrelazione ed eterogeneità spaziale

Una delle proprietà richieste per l'omogeneità spaziale di un processo (proprietà 3) è che la covarianza spaziale sia costante. Come detto all'inizio del capitolo, anche l'autocorrelazione spaziale è una forma di covarianza spaziale. Quindi la covarianza spaziale è forse, da sola, la proprietà che meglio caratterizza i processi spaziali: se non nulla, la covarianza costituisce dipendenza, o autocorrelazione, spaziale; se incostante essa costituisce eterogeneità spaziale.

ESEMPIO 2.1. Malattie cardiovascolari in ambiente urbano²¹.

Il tasso d'incidenza della malattia varia nello spazio urbano:

➤ **processo spaziale.**

La variazione spaziale dell'incidenza della malattia è caratterizzata da dipendenza ed eterogeneità spaziali:

➤ **covarianza non-nulla ed incostante del processo spaziale.**

L'incidenza della malattia manifesta valori simili all'interno di uno stesso quartiere o entro quartieri limitrofi:

➤ **autocorrelazione spaziale.**

➤ [la proprietà è riconducibile alla similitudine di età, reddito e accesso alle strutture sanitarie dei residenti di quei quartieri].

L'incidenza della malattia è incostante nello spazio urbano:

➤ media incostante: **eterogeneità spaziale** (Proprietà 1):

➤ [l'incidenza è inferiore nei quartieri a prevalenza di persone giovani o facoltose rispetto a quartieri con prevalenza di residenti anziani o con reddito medio-basso].

²¹ L'esempio è trattato in dettaglio nel capitolo 7. Si vedano anche: Bertazzon (2009) e Bertazzon, Olson, Knudtson (2010).

La variabilità dell'incidenza della malattia è incostante nello spazio urbano:

- varianza incostante: **eterogeneità spaziale** (Proprietà 2):
- [la proprietà è riconducibile alla maggiore variabilità nei quartieri a prevalenza di residenti giovani rispetto ai quartieri a prevalenza di anziani].

L'estensione spaziale della similitudine (*autocorrelazione spaziale*) è incostante nello spazio urbano:

- covarianza incostante: **eterogeneità spaziale** (Proprietà 3):
- [la proprietà è riconducibile alla differenza nella densità di popolazione e alla destinazione ed intensità dell'uso del suolo, per cui ciò che è "lontano" nei quartieri centrali è "vicino" in quelli periferici].

La variazione della media, quella della varianza, e quella dell'estensione spaziale della similitudine (*autocorrelazione spaziale*) sono incostanti nello spazio urbano:

- anisotropia: **eterogeneità spaziale** (Proprietà 4):
- [la proprietà è riconducibile alle differenze nella densità abitativa, nel tipo di lottizzazione, nella rete stradale, nelle eventuali barriere fisiche quali fiumi, ferrovie o parchi che interrompano la continuità residenziale].

Si osservi, inoltre, come nell'Esempio 2.1 l'incidenza di malattie cardiovascolari manifesti autocorrelazione spaziale positiva, anche se in realtà si tratta di un processo *interattivo*, poiché la relazione tra prossimità residenziale e incidenza della malattia è riconducibile alla prossimità residenziale di persone di età simile e quindi alla relazione tra prossimità residenziale ed età dei residenti, soprattutto perché l'età media dei residenti rispecchia il periodo in cui i quartieri furono lottizzati e vi si stanziarono persone di una certa fascia d'età. Ciascuno di questi fattori è in realtà un processo spaziale a sé stante, e per comprenderne l'influenza sul processo d'interesse è utile fare riferimento ad uno schema analitico multivariato, quale l'analisi di regressione.

2.2.6. Il problema dell'unità areale modificabile (MAUP)

La presenza di autocorrelazione ed eterogeneità nei processi spaziali è alla base del cosiddetto problema dell'unità areale modificabile (*modifiable areal unit problem*, o *MAUP*), secondo cui i risultati

analitici variano, in una medesima area di studio, al variare del numero e dell'aggregazione delle unità spaziali utilizzate per l'analisi. Il problema dell'unità areale modificabile (*modifiable areal unit problem, o MAUP*) è infatti composto di due elementi, o problemi (*problems*), separati ma strettamente collegati: la scala e l'aggregazione (*scale and aggregation*). Il problema della scala si riferisce alla variazione nei risultati che si tende ad ottenere quando i dati riferiti ad un insieme di unità areali vengono progressivamente aggregati in aree più grandi e meno numerose per poi essere ri-analizzati. Il problema dell'aggregazione si riferisce a qualsiasi variazione nei risultati dovuta all'uso di combinazioni alternative di unità areali, pur se il numero di unità è mantenuto costante²². È chiaro nessuno dei due problemi sussisterebbe in assenza di autocorrelazione ed eterogeneità spaziali.

2.3. L'analisi di regressione e gli effetti spaziali

L'analisi di regressione è una metodologia che la geografia quantitativa prima e l'analisi spaziale poi hanno mutuato dalla statistica²³. La centralità dell'analisi di regressione nella nostra disciplina è legata alla sua importanza metodologica poiché, come si è visto, la maggior parte delle tecniche utilizzate dalla geografia quantitativa sono forme di regressione.

Alla base del modello di regressione vi è l'ipotesi che un processo spaziale, rappresentato da una serie di osservazioni su una variabile, possa essere concettualizzato come funzione, intesa come relazione matematica, di uno o più altri processi spaziali. Il processo d'interesse viene detto variabile dipendente o risposta (*response variable*), mentre i processi ad esso associati sono detti variabili indipendenti, esplicative o regressori (*predictors*)²⁴.

Un esempio caro ai geografi quantitativi²⁵ è il cosiddetto modello edonico, in cui i prezzi delle abitazioni vengono stimati in funzione di fattori quali la superficie, il numero di vani o la localizzazione: la stima del prezzo di una casa poggia sul calcolo della relazione tra ciascuna

²² Si veda Openshaw (1977).

²³ Barnes (1998).

²⁴ Si veda Burt, Barber, Rigby (2009).

²⁵ Si veda ad esempio Fotheringham, Brunson, Charlton (2000).

variabile esplicativa (i fattori appena elencati) e la dipendente (il prezzo) osservati per un campione di abitazioni. Analogamente, consideriamo il modello di regressione LUR (*land use regression*) per la stima dell'inquinamento atmosferico a grande scala spaziale²⁶. Come ogni modello di regressione, esso poggia su un campione di osservazioni sulla variabile dipendente, in questo caso la concentrazione di un certo inquinante nell'atmosfera, e sull'insieme delle variabili esplicative, in questo caso le cosiddette variabili d'uso del suolo (*land use variables*), ossia prossimità a strade e ferrovie, presenza di stabilimenti industriali, densità di popolazione, ecc. Tutte le variabili, dipendente ed esplicative, sono osservazioni relative a processi spaziali misurati in siti campione. Sulla base delle osservazioni campionarie, il modello stima la relazione tra ciascuna esplicativa e la dipendente. I valori numerici di tali relazioni (definiti di seguito) serviranno poi a stimare l'inquinamento anche nei siti non campionati all'interno dell'area di studio della regressione²⁷.

2.3.1. Il modello di regressione

Il modello di regressione è dunque una relazione matematica che esprime la variabile dipendente in funzione delle variabili esplicative (Equazione 10). Nell'esempio del modello LUR la variabile dipendente è la concentrazione di inquinante nell'atmosfera nei siti campionati; come variabili esplicative consideriamo la lunghezza complessiva delle strade entro un certo raggio da ogni sito campionato, il numero di stabilimenti industriali entro il medesimo raggio, la densità di popolazione nella porzione residenziale entro quel raggio. La relazione matematica è definita dai coefficienti di regressione, che legano (funzionalmente, o matematicamente) ciascuna variabile esplicativa alla dipendente. L'errore, o residuo, è definito come la differenza tra i valori osservati della dipendente e quelli calcolati come prodotto delle variabili esplicative per i rispettivi coefficienti.

²⁶ Hoek *et Alii* (2008).

²⁷ Un modello di regressione LUR è illustrato nel capitolo 7. Si vedano anche Bertazzon e Couloigner (2018) e Bertazzon, Couloigner, Underwood (2019).

Obiettivo dell'analisi è la stima, il più possibile corretta ed affidabile, dei coefficienti di regressione. Questi sono ottenuti come combinazione delle osservazioni sulla dipendente e sulle esplicative, tramite una procedura stocastica volta a minimizzare l'errore, o residuo²⁸. Ciascun coefficiente di regressione, β , è infatti una combinazione delle osservazioni sulla variabile dipendente e sulle esplicative (Equazione 11). Si tratta di una relazione estremamente importante, che mostra come il legame tra ciascuna esplicativa e la dipendente (ad esempio tra la concentrazione di inquinante in atmosfera e il numero di stabilimenti industriali nel raggio definito) è determinato dalle osservazioni sulle variabili ai siti campionati. Questo implica che con la specificazione e la stima del modello di regressione si assume sì che esista una relazione tra dipendente ed esplicative, ma anche che il valore numerico di tale relazione (coefficiente di regressione) vari al variare del campione empirico utilizzato. Di fondamentale importanza per i risultati del modello è quindi il campione statistico sulla cui base viene calcolato il coefficiente di regressione²⁹. Questo significa che il modello non fornisce una legge universale che leghi alla dipendente ciascuna esplicativa, ma soltanto una relazione numerica valida per quel campione, e difficilmente generalizzabile (estrapolabile) oltre l'ambito del campione da cui è stata ottenuta.

2.3.2. Le ipotesi del modello di regressione e i processi spaziali

Se i coefficienti di regressione sono combinazioni delle variabili dipendente ed esplicative, sono anch'essi quantità stocastiche. Per questo la loro stima, o calcolo, è subordinata alla verifica di una serie di ipotesi, la cui violazione influisce negativamente sulle proprietà delle stime del modello. Le ipotesi concernono la relazione funzionale tra dipendente ed esplicative, le variabili esplicative e l'errore³⁰. Tra esse, le più rilevanti per l'analisi spaziale sono quelle relative

²⁸ Burt, Barber, Rigby (2009, p. 475).

²⁹ Più precisamente, il valore numerico del coefficiente è dato dal rapporto tra la somma dei prodotti delle variabili dipendenti ed esplicative e la somma dei quadrati delle esplicative, come mostrato dalle Equazioni 11 e 12.

³⁰ Burt, Barber, Rigby (2009, p. 495): «Testing the statistical significance of the equation or using the derived model for prediction purposes requires four specific assumptions».

all'errore, in particolare che l'errore segua una distribuzione identica ed indipendente. Come i coefficienti di regressione, anche l'errore (Equazione 13) è funzione delle variabili dipendente ed esplicative, e quindi le proprietà statistiche dell'errore sono legate a quelle delle variabili del modello, tanto che spesso un'attenta analisi statistica descrittiva (o esplorativa) spaziale³¹ delle variabili dipendente ed esplicative consente di identificare potenziali violazioni delle ipotesi del modello prima ancora che esso sia stimato. L'analisi esplorativa spaziale può infatti ispezionare le variabili del modello per comprendere se esse soddisfino le ipotesi di distribuzione identica e indipendente. Una distribuzione identica presenta le medesime caratteristiche, o proprietà stocastiche, nell'intera area di studio, ossia rimane invariata in tutti i punti del campione, e quindi, nel caso di variabili spaziali, in tutti i siti campionati e sull'intera area di studio. Una distribuzione indipendente è tale che le proprietà stocastiche di ogni osservazione siano indipendenti da quelle di ogni altra osservazione nel campione³². Quindi, in tutti i casi in cui il modello di regressione viene applicato a dati spaziali, si richiede che tali ipotesi siano soddisfatte da ciascuno dei processi spaziali coinvolti (variabili dipendente ed esplicative).

2.3.3. Effetti spaziali e inaffidabilità delle stime

Se si considerano a questo punto le proprietà intrinseche dei processi spaziali, autocorrelazione (o dipendenza) spaziale ed eterogeneità spaziale, è evidente come l'ipotesi di *indipendenza* sia violata dall'autocorrelazione (o *dipendenza*) spaziale, e l'ipotesi di *identità* sia violata dall'*eterogeneità* spaziale. Ne consegue che l'analisi di regressione, calcolata su variabili spaziali, incorre nella violazione delle ipotesi sulla distribuzione dell'errore ed ancora, poiché l'analisi di regressione sorregge la gran parte delle tecniche di geografia quantitativa, questa violazione tende a verificarsi in tutte le analisi quantitative spaziali³³.

³¹ Exploratory spatial data analysis, o ESDA; si veda Burrough (1986).

³² Rogerson (2014, p. 127): «One of the assumptions of the statistical hypothesis tests described in this chapter is that observations are independent. This means that the observed value of one observation is not affected by the value of another observation».

³³ Griffith e Amrhein (1977).

Analizziamo ora brevemente quelli che all'inizio di questo capitolo sono stati chiamati effetti spaziali, ossia le conseguenze della violazione delle ipotesi dell'analisi di regressione da parte dei processi spaziali. La varianza del coefficiente di regressione ottenuto in violazione delle ipotesi di indipendenza o identità dell'errore (Equazione 15) è maggiore della varianza del coefficiente di regressione ottenuto da un modello in cui tali ipotesi siano soddisfatte (Equazione 14), come un semplice raffronto visivo delle due formule consente di intuire. Tale incremento di varianza del coefficiente comporta un incremento dell'incertezza ad esso associata e quindi una diminuzione della sua affidabilità (*reliability*).

2.3.4. Effetti spaziali e il problema dell'inferenza

Legato all'incremento di varianza delle stime c' è un secondo problema: il suo impatto sull'inferenza statistica. La valutazione del modello di regressione e la verifica delle sue ipotesi poggiano su una gamma di test statistici inferenziali, dai classici test F e t sulla significatività della regressione e di ciascuna esplicativa³⁴, fino alla mole di test sui residui sviluppati soprattutto in ambito econometrico³⁵. Per comprendere il senso di questo problema ci soffermiamo proprio sul semplice test t utilizzato nel processo di selezione delle variabili per valutare la significatività di ciascun coefficiente nel modello (Equazione 16), al cui denominatore compare la varianza del coefficiente. Nel caso in cui il modello sia stato ottenuto da processi spaziali, che violano le ipotesi sull'errore del modello di regressione, l'inflazione della varianza del coefficiente comporta una maggiore tendenza a rifiutare l'ipotesi nulla quando essa è vera (errore di tipo I), ossia a considerare significative nel modello variabili che in effetti non lo sono. Il risultato è che si tende a ritenere valido un modello che contiene variabili esplicative che non sono in relazione con la dipendente. I due problemi generati dall'inflazione della varianza (inaffidabilità delle stime ed errore di tipo I nell'inferenza), sommati, conducono a

³⁴ Burt, Barber, Rigby (2009); Rogerson (2014).

³⁵ Anselin (2002); Griffith (2008); Bivand e Piras (2015); Florax, Folmer, Rey (2003).

considerare come “veri” dei coefficienti inaffidabili e riferiti a relazioni inesistenti.

Tutto questo spiega, seppur brevemente, la ragion d’essere dell’analisi spaziale: assumendo consapevolezza che ogni analisi quantitativa su processi spaziali sia affetta da un’eccessiva variabilità, che ne mina l’affidabilità, essa si prefigge di formulare delle soluzioni capaci di ripristinarne l’affidabilità.

2.4. Effetti spaziali e metodi analitici spaziali

Benché i due effetti spaziali (autocorrelazione o dipendenza ed eterogeneità) siano strettamente legati tramite la covarianza spaziale, allo stato attuale, la ricerca ha prodotto solo metodi correttivi che risolvono i due diversi effetti spaziali separatamente: si tratta, rispettivamente, del modello *SAR* (*spatial autoregression*) e del modello *GWR* (*geographically weighted regression*).

2.4.1. Il modello SAR in risposta all’autocorrelazione spaziale

Il problema della dipendenza nei processi spaziali, noto ai geografi quantitativi, fu affrontato inizialmente dagli econometrici e affinato nei decenni seguenti³⁶. Come osservato, l’econometria disponeva già di metodi per affrontare la dipendenza nei processi temporali, e su quelli poggia il metodo per trattare la dipendenza nei processi spaziali. Essi consistono nell’utilizzo di modelli autoregressivi, ossia modelli in cui il valore di un processo in un dato istante temporale è funzione dei valori che l’hanno preceduto. Com’è evidente, il metodo poggia su una concezione del tempo lineare ed unidirezionale, che procede dal passato al presente ed al futuro, per cui ogni realizzazione del processo è funzione del proprio passato. L’idea dell’econometria spaziale fu, dunque, di estendere questo tipo di modelli ai processi spaziali³⁷. Tuttavia, lo spazio presenta una complessità maggiore rispetto al tempo,

³⁶ Per ciascuna delle tre affermazioni si veda rispettivamente: Cliff e Ord (1970), poi Paelnik e Klaassen (1979) e per l’ultima Anselin (2010).

³⁷ Paelnik e Klaassen (1979).

essendo lo spazio per lo meno bidimensionale e privo di una direzione preferenziale. Nel processo spaziale, dunque, ogni realizzazione è funzione di altre realizzazioni secondo modalità regolate dalla distanza. La trasposizione allo spazio del modello autoregressivo temporale richiede quindi una struttura concettuale e computazionale che definisca le relazioni tra le diverse realizzazioni del processo, ossia le osservazioni spaziali. Per questo il modello autoregressivo spaziale (Equazione 18), noto anche come *SAR* (*Spatial AutoRegression*) contiene anche la matrice dei pesi spaziali. Integrando la componente *SAR* nel modello di regressione definito in precedenza (Equazione 10) si ottiene l'intero modello, che comprende sia il termine autoregressivo spaziale che le variabili esplicative (Equazione 19).

Il modello così definito, in apparenza è costituito dalla somma del modello standard e della componente autoregressiva. Tuttavia, la struttura autoregressiva comporta la presenza della variabile dipendente anche tra le esplicative (Equazione 19): ciò che modifica la forma funzionale del nuovo modello violandone alcune delle ipotesi³⁸. La maggiore complessità del modello *SAR*, rispetto al modello standard, ricade sull'errore, (ε), che per definizione contiene la variabilità non spiegata dal modello³⁹. La stima del modello richiede una specificazione del tipo di struttura autoregressiva ipotizzata, l'utilizzo di un diverso metodo di stima e di un diverso apparato inferenziale per la verifica delle ipotesi⁴⁰. In questo modo il modello autoregressivo spaziale risolve, almeno da un punto di vista statistico e computazionale, il problema dell'inflazione della varianza indotta dalla dipendenza spaziale (Equazioni 19 – 24): il termine autoregressivo consente infatti di rimuovere l'inflazione della varianza indotta dall'autocorrelazione spaziale. In tal modo, utilizzando un metodo autoregressivo spaziale, la struttura spaziale o autocovarianza del processo spaziale viene scorporata dal modello di regressione attraverso l'inverso di una matrice analoga a quella dei pesi spaziali (Equazione 21).

In conclusione, con il modello *SAR* si ottengono stime affidabili ed inferenza corretta, anche se a prezzo di scorporare la struttura spaziale dal modello, come argomentato da Stefania Bertazzon (2003). Inoltre,

³⁸ Burt, Barber, Rigby (2009); Fotheringham, Brunson, Charlton (2000).

³⁹ Rogerson (2014).

⁴⁰ Anselin (1988).

come osservato in apertura, il modello *SAR* allevia l'inflazione della varianza indotta dalla dipendenza spaziale, ma non quella indotta dall'eterogeneità spaziale.

2.4.2. Il modello *GWR* in risposta all'eterogeneità spaziale

Anche il problema dell'eterogeneità spaziale fu affrontato metodologicamente nell'ambito dell'econometria a partire dai primi anni '70 da Emilio Casetti (1972), poi pienamente sviluppato in ambito geografico sul finire degli anni '90 da Stuart Fotheringham, Martin Charlton e Chris Brunsdon (1998), integrandovi anche il recente sviluppo della cosiddetta analisi locale ad opera di Luc Anselin (1995).

L'effetto dell'eterogeneità spaziale è legato all'instabilità nello spazio della relazione tra variabile dipendente ed esplicative, che induce, ancora una volta, un'inflazione della varianza delle stime del modello, oltre ad incrementare la probabilità di errore di tipo I nelle procedure inferenziali. La metodologia che gestisce questo problema è il cosiddetto modello *GWR* (*geographically weighted regression*). Si tratta di un modello molto semplice dal punto di vista matematico e statistico, che poggia sull'idea di suddividere lo spazio del processo in frazioni sufficientemente piccole da poter ipotizzare che in ciascuna di esse il processo sia spazialmente omogeneo, ossia vi siano soddisfatte le quattro proprietà che definiscono l'omogeneità spaziale.

Il modello *GWR* è identico a quello della regressione standard (Equazione 10), con l'unica differenza che nel modello *GWR* sono precisate le coordinate geografiche di ciascun coefficiente di regressione (Equazione 25). Infatti, nella regressione *GWR* i coefficienti variano nello spazio, perché il modello *GWR* stima una miriade di micro-regressioni locali nell'area di studio.

Per ottenere questa molteplicità di modelli, *GWR* usa una funzione, detta *kernel* (seme) allo scopo di selezionare le unità spaziali circostanti ciascuna unità, che saranno usate per stimare ogni regressione locale⁴¹. Di questa funzione esistono due tipi principali: il *kernel* fisso (*fixed*) e quello adattivo (*adaptive*). Il *kernel* fisso, agendo come un compasso, definisce un raggio attorno a ciascuna unità spaziale,

⁴¹ Fotheringham, Brunsdon, Charlton (2003).

selezionando tutte le unità che si trovano all'interno del cerchio. Il metodo funziona bene se le unità spaziali sono disposte in maniera abbastanza regolare, ma può generare notevoli disparità numeriche tra regressioni locali in contesti in cui le unità spaziali siano disposte in maniera irregolare. Per questo secondo caso si utilizza il *kernel* adattivo, che, partendo da un punto, ne seleziona gli n vicini più prossimi. Nel *kernel* fisso, quindi, è data la distanza ma non il numero di unità, mentre in quello variabile è dato il numero di unità ma non la distanza⁴². Il parametro che definisce il *kernel*, ossia il raggio (per il fisso) o il numero di vicini (per l'adattivo), è chiamato larghezza di banda (*bandwidth*). Si tratta di un parametro fondamentale: se troppo piccolo, le regressioni locali saranno penalizzate in termini statistici dalla scarsa numerosità del campione; se troppo grande, paradossalmente ricondurrà *GWR* ad una regressione standard, o, nel gergo dell'analisi locale, ad una regressione globale⁴³. Il metodo *GWR*, come visto, è più recente del metodo *SAR*, tuttavia ha riscontrato una certa popolarità, in relazione alla quale il gruppo di Stewart Fotheringham continua ad affinarlo⁴⁴.

2.5. Definizioni matematiche e statistiche

2.5.1. I processi spaziali

Un processo spaziale o, più precisamente, processo stocastico spaziale è un modello probabilistico definito da una collezione di variabili aleatorie e dalla loro funzione di distribuzione congiunta⁴⁵. Si tratta di un modello di variazione spaziale, definito dall'Equazione 1:

⁴² Si noti la corrispondenza con le definizioni di contiguità basate sulla distanza soglia e sul numero di vicini più prossimi.

⁴³ Generalmente si utilizzano dei criteri per definire la larghezza di banda in modo tale da ottimizzare il modello *GWR*, che sono descritti nei lavori di Fotheringham, Brunson, Charlton (2003) e che pure accompagnano il software da loro commercializzato.

⁴⁴ Si vedano a questo proposito i loro lavori più recenti come, ad esempio, Fotheringham, Crespo, Yao (2015).

⁴⁵ Cressie (1993).

$$\{y_i, i \in D\}$$

Equazione 1

dove y_i rappresenta il valore del processo spaziale (y) nel punto i dello spazio D , che può essere una superficie continua o un insieme finito di punti. Proprietà qualificante dei processi spaziali è la covarianza, o momento secondo, definita da Luc Anselin⁴⁶ come una quantità diversa da zero, secondo l'Equazione 2:

$$\text{Cov}(y_i, y_j) = E[y_i - \mu_y]E[y_j - \mu_y] \neq 0 \text{ per } i \neq j$$

Equazione 2

dove il simbolo E indica il valore atteso⁴⁷ e μ_y la media del processo spaziale y . L'equazione esprime la covarianza come relazione tra diverse realizzazioni del processo, per $i \neq j$, ovvero escludendo il caso banale in cui si confronti un punto dello spazio con se stesso ($i = j$).

La covarianza non-nulla rappresenta rigorosamente la prima parte dell'enunciato della legge di Tobler; la seconda parte dell'enunciato, invece, si può esprimere rigorosamente come funzione di attrito della distanza (*distance decay function*). La più generica di queste relazioni è espressa dall'Equazione 3:

$$I = d^{-k}$$

Equazione 3

dove l'interazione, I , è espressa come funzione inversa della distanza (d), modulata da un coefficiente (k).

⁴⁶ Anselin (2001, pp. 4-5): "This covariance becomes meaningful from a spatial perspective when the particular configuration of nonzero i, j pairs has an interpretation in terms of spatial structure, spatial interaction or the spatial arrangement of the observations."

⁴⁷ L'operatore "valore atteso" calcola la media.

2.5.2. L'autocorrelazione spaziale

Tra i diversi indici di autocorrelazione spaziale⁴⁸, l'indice I di Moran è descritto dall'Equazione 4⁴⁹:

$$I = \left(\frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \right) \left(\frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \mu_x)(x_j - \mu_x)}{\sum_{i=1}^n (x_i - \mu_x)^2} \right)$$

Equazione 4

dove I è il simbolo comunemente detto “ I di Moran”⁵⁰, n indica il numero di oggetti (punti spaziali) contenuti nel campione (*sample size*); i e j sono due qualsiasi di tali oggetti; x_i è il valore dell'attributo x nel punto i , μ_x è la media dei valori di tale attributo e w_{ij} rappresenta la contiguità (detta anche *prossimità* o *distanza*) tra la localizzazione di i e quella di j ; dove $w_{ij} = 0$ per ogni i non contiguo a j , e per $i=j$. L'indice I di Moran, come qualsiasi coefficiente di correlazione, è definito nell'intervallo $[-1 \leq I \leq 1]$ ⁵¹ ed indica autocorrelazione negativa nell'intervallo $[-1 \leq I < 0]$, positiva nell'intervallo $[0 < I \leq 1]$ e nulla per $I = 0$.

Il secondo indice di autocorrelazione spaziale più comunemente usato è l'indice di Geary, di interpretazione meno immediata⁵².

⁴⁸ Per un'analisi dettagliata si vedano: Chun e Griffith (2016) e Getis (2008).

⁴⁹ Goodchild (1986).

⁵⁰ Questa I ha un significato diverso da quella dell'Equazione 3: mentre in quel caso denota l'interazione, in questo caso denota l'indice di autocorrelazione spaziale.

⁵¹ Per motivi legati alla formulazione matematica dell'indice di Moran, è possibile, nella pratica, che il valore assuma valori di poco superiori a 1.

⁵² L'indice c di Geary è definito dalla seguente espressione:

$$c = \frac{(n-1) \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - x_j)^2}{2 \left(\sum_{i=1}^n \sum_{j=1}^n w_{ij} \right) \sum_{i=1}^n (x_i - \mu_x)^2}$$

dove la notazione corrisponde a quella definita per l'Equazione 4. L'indice di Geary varia tra 0 e 2; la sua interpretazione, tuttavia, è meno intuitiva: se $0 < c < 1$ l'autocorrelazione spaziale è positiva, se $1 < c < 2$ l'autocorrelazione spaziale è negativa e se $c=1$ è nulla. Si veda: Goodchild (1986).

2.5.3. La matrice dei pesi spaziali

Nell'Equazione 4 la contiguità tra ciascuna possibile coppia di unità spaziali è espressa dal termine w_{ij} . Il termine w_{ij} dell'Equazione 4 è un elemento di una matrice dei pesi spaziali (o di contiguità), W , formata di tanti elementi quante sono le unità spaziali presenti nel campione, come descritto dall'Equazione 5:

$$W = [w_{ij}], \quad w_{ij} = \begin{cases} 1 & \text{se le unità spaziali } i \text{ e } j \text{ sono contigue} \\ 0 & \text{se le unità spaziali } i \text{ e } j \text{ non sono contigue} \end{cases}$$

Equazione 5

dove W indica la matrice e w_{ij} ciascuno dei suoi elementi. Nella sua forma più semplice W è una struttura binaria che stabilisce se due unità spaziali siano o meno contigue. Nelle formulazioni più complesse, tipicamente usate nella pratica, la matrice dei pesi spaziali binaria viene moltiplicata per una funzione di attrito della distanza, analoga a quella definita nell'Equazione 3. Moltiplicando la matrice binaria per la funzione di attrito della distanza si ottiene una struttura che pesa l'interazione spaziale per la distanza tra due unità spaziali, ammesso che esse siano definite contigue secondo l'Equazione 5. Per questa ragione, la matrice di contiguità è spesso chiamata anche matrice dei pesi spaziali (*spatial weights matrix*)⁵³.

La matrice dei pesi spaziali definita nell'Equazione 5 serve a definire i diversi tipi di contiguità, ad esempio utilizzando il numero di vicini più prossimi (*nearest neighbours*) o una distanza soglia (*threshold distance*)⁵⁴.

2.5.4. L'eterogeneità (non-stazionarietà) spaziale

Un processo spaziale si dice stazionario (o omogeneo) se soddisfa quattro proprietà:

⁵³ Anselin, Murray, Rey (2000).

⁵⁴ Getis e Aldstadt (2004).

1. Ha media costante nello spazio:

$$E(y_i) = \mu_y \text{ per ogni } i$$

Equazione 6

2. Ha varianza costante nello spazio:

$$\text{Var}(y_i) = \sigma_y^2 \text{ per ogni } i$$

Equazione 7

3. Ha covarianza costante nello spazio:

$$\text{Cov}(y_i, y_{i+k}) = E[y_i - \mu_y]E[y_{i+k} - \mu_y] = \gamma_k \text{ per ogni } i, k$$

Equazione 8

4. È isotropico: le proprietà 1 – 3 sono soddisfatte in ogni direzione.

Nelle Equazioni 6 – 8 y_i è il processo spaziale presentato nell'Equazione 1. Secondo l'Equazione 6 la media del processo è costante e uguale al termine μ_y nell'intera area di studio. Anche la varianza è costante e uguale a σ_y^2 in tutta l'area di studio, come definito dall'Equazione 7. Infine, la covarianza, definita dall'Equazione 8 analogamente all'Equazione 2, è costante e uguale a γ , per ogni i , ossia in qualsiasi punto dell'area di studio, e per ogni k , ossia per qualsiasi intervallo, o distanza tra i e j . La covarianza è comunque definita come funzione della distanza k tra i e j , γ_k , ossia, γ rimane costante per ogni dato valore di k ma varia al variare della distanza k tra i e j . L'isotropia, proprietà 4, richiede inoltre che la covarianza rimanga costante indipendentemente dalla direzione della distanza k tra i e j .

Un processo spaziale che violi almeno una di queste quattro proprietà si dice *non-stazionario*, o *eterogeneo* nello spazio, ovvero caratterizzato da *eterogeneità spaziale*⁵⁵.

⁵⁵ Anselin, Murray, Rey (2000).

Come l'autocorrelazione spaziale viene misurata dagli indici di Moran o di Geary, anche per l'eterogeneità spaziale si utilizzano degli indici di *clustering*, il più comune dei quali è l'indice G di Getis-Ord⁵⁶, descritto dall'Equazione 9⁵⁷:

$$G = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} x_i x_j}{\sum_{i=1}^N \sum_{j=1}^N x_i x_j}$$

Equazione 9

dove la notazione è la stessa dell'Equazione 4. Di tale indice sono notevoli la versione locale, G_i , e l'indice locale G_i^* , divenuto popolare in ambito GIS per il suo utilizzo nella *hotspot analysis*⁵⁸.

2.5.5. Relazione tra autocorrelazione ed eterogeneità spaziale

L'Equazione 2 può essere riscritta sostituendo al termine j il termine $i+k$. La differenza tra le due espressioni, infatti, è legata al fatto che nell'Equazione 8 la differenza tra i e j viene esplicitata nel termine k , al fine di descrivere la covarianza spaziale come funzione della distanza (γ_k). Come già affermato, l'autocorrelazione e l'eterogeneità spaziale sono due aspetti diversi della covarianza spaziale.

Osservando l'Equazione 8 si può infine notare che γ_k varia con la distanza, e può così rappresentare, da sola, la legge di Tobler.

⁵⁶ Getis e Ord (1992).

⁵⁷ Bivand e Wong (2018).

⁵⁸ In proposito si vedano ad esempio la documentazione fornita dal software R sull'uso di G_i (local G) e G_i^* [pacchetto "spdep" versione 1.1-5]: www.rdocumentation.org/packages/spdep/versions/1.1-5/topics/localG (visitato il 13/1/2022) e quella fornita dal software ArcGIS Pro: [https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-high-low-clustering-getis-ord-general-g-spat.htm#:~:text=The%20High%2FLow%20Clustering%20%28Getis-Ord%20General%20G%29%20statistic%20is,there%20is%20no%20spatial%20clustering%20of%20feature%20values](https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-high-low-clustering-getis-ord-general-g-spat.htm#:~:text=The%20High%2FLow%20Clustering%20%28Getis-Ord%20General%20G%29%20statistic%20is,there%20is%20no%20spatial%20clustering%20of%20feature%20values.). Visitato il 13/01/2022.

2.5.6. Il problema dell'unità areale modificabile (MAUP)

Il coefficiente di correlazione⁵⁹ è un indice che misura la covarianza tra due variabili⁶⁰. Tale coefficiente è di cruciale importanza nell'analisi bivariata e multivariata, e di conseguenza nell'analisi di regressione e in numerose analisi quantitative. Come affermato in un noto saggio⁶¹, il coefficiente di correlazione varia al variare della scala e aggregazione delle unità areali. Nel problema dell'unità areale modificabile (*modifiable areal unit problem, o MAUP*), quello della scala (*scale*) è anche un problema matematico, relativo al numero di unità considerate nell'analisi, mentre quello dell'aggregazione è indipendente dal numero di unità considerate, ma riconducibile ad eterogeneità e dipendenze spaziali⁶². Osserviamo inoltre che le proprietà statistiche delle stime, ad esempio nei modelli di regressione, sono legate anche alla dimensione del campione e quindi non sono indipendenti dalla scala. Generalmente le proprietà statistiche tendono a migliorare con l'aumentare della numerosità del campione; tuttavia, nel contesto spaziale, la numerosità del campione implica maggiore frammentazione dell'area di studio e maggiore instabilità (eterogeneità) delle relazioni.

2.5.7. Il modello di regressione

Il modello di regressione è descritto dall'Equazione 10, che costituisce il più semplice caso di una relazione funzionale lineare:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i$$

Equazione 10

dove:

⁵⁹ La correlazione, come è noto, non è che una covarianza standardizzata (Burt, Barber, Rigby, 2009)

⁶⁰ Burt, Barber, Rigby (2009, pp. 168-171).

⁶¹ Il saggio (Openshaw e Taylor, 1979) s'intitola proprio: «A million or so correlation coefficients».

⁶² Openshaw (1977).

y_i è la variabile dipendente;

$x_{i1}, x_{i2}, \dots, x_{ik}$ sono le variabili esplicative;

β_0 è l'intercetta;

$\beta_1, \beta_2, \dots, \beta_k$ sono i coefficienti di regressione associati a ciascuna variabile esplicativa;

ε_i è l'errore, o residuo, definito come la differenza tra i valori osservati della dipendente e quelli calcolati come prodotto delle variabili esplicative per i rispettivi coefficienti.

Obiettivo dell'analisi di regressione è la stima dei coefficienti (β_0, \dots, β_k) che legano ciascuna variabile esplicativa alla dipendente.

Nell'ipotesi di una relazione lineare, e assumendo che una serie di altre ipotesi siano soddisfatte, è possibile utilizzare il metodo OLS (*ordinary least squares*) per calcolare i coefficienti di regressione in maniera tale da minimizzare l'errore della regressione⁶³. Con tale metodo si ottiene il vettore, ossia l'insieme dei coefficienti di regressione (β_0, \dots, β_k) dalla relazione:

$$\beta = (X'X)^{-1}(X'Y)$$

Equazione 11

L'Equazione 11, in cui, per semplicità, il vettore dei coefficienti è espresso in forma matriciale⁶⁴, evidenzia come ciascun coefficiente di regressione β altro non è che una combinazione delle osservazioni sulla variabile dipendente (Y) e sulle esplicative (X). A titolo esplicativo, la formula in livelli è illustrata in Equazione 12; tuttavia, la forma matriciale sarà più utile nel seguito.

$$\beta = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n (x_i)^2}$$

Equazione 12

L'Equazione 12 è equivalente all'Equazione 11.

⁶³ Burt, Barber, Rigby (2009).

⁶⁴ Per questa ragione vi si utilizzano le lettere maiuscole, che nell'algebra lineare denotano vettori e matrici.

2.5.8. Le ipotesi del modello di regressione e i processi spaziali

Poiché ogni coefficiente (definito dall'Equazione 11) è una combinazione delle osservazioni sulla variabile dipendente (Y) e sulle esplicative (X), anch'esso, come le variabili che lo formano, è una quantità stocastica, le cui proprietà stocastiche (distribuzione di probabilità, momenti, o misure di tendenza centrale e dispersione) sono collegate a quelle delle variabili che lo formano. Per questo, la stima dei coefficienti di regressione poggia su una serie di ipotesi (o assunzioni). L'eventuale violazione di una o più di tali ipotesi influisce negativamente sulle proprietà delle stime del modello. Le ipotesi concernono: (1) la relazione funzionale tra dipendente ed esplicative, che si assume lineare; (2) le variabili esplicative, che si assumono incorrelate tra loro⁶⁵ ed incorrelate con l'errore o residuo del modello, (3) l'errore (o residuo), che si assume avere distribuzione normale, indipendente ed identica⁶⁶. Di questo insieme di ipotesi, le più rilevanti per l'analisi spaziale sono quelle relative all'errore, e in particolare le ipotesi di indipendenza ed identità della sua distribuzione.

Riscrivendo l'errore del modello di regressione (Equazione 10) come differenza tra osservazioni e stime si ottiene l'Equazione 13.

$$\varepsilon_i = y_i - (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik})$$

Equazione 13

Dall'equazione è facile osservare che anche l'errore è funzione delle variabili esplicative (X), e della dipendente (Y), tramite i coefficienti β definiti dall'Equazione 11; quindi anche le proprietà stocastiche dell'errore sono collegate a quelle di tali variabili.

⁶⁵ Questa è l'unica delle ipotesi la cui violazione non si limita ad influire sulla qualità delle stime, ma può comportare l'impossibilità di stimare il modello, nel caso di correlazione perfetta tra le variabili esplicative, o multicollinearità perfetta.

⁶⁶ Burt, Barber, Rigby (2009).

2.5.9. Effetti spaziali e inaffidabilità delle stime

Applicando l'analisi di regressione ai processi spaziali, alcune ipotesi sull'errore del modello di regressione tendono ad essere violate: come suggerito dalla terminologia, l'ipotesi di *indipendenza* dell'errore è violata dalla *dipendenza* (o autocorrelazione) spaziale e l'ipotesi di *identicità* è violata dall'*eterogeneità* spaziale. L'Equazione 14 descrive la varianza del coefficiente di regressione β ottenuto da un modello in cui le ipotesi di indipendenza ed identità dell'errore siano soddisfatte, ammettendo che siano soddisfatte anche le altre ipotesi del modello di regressione:

$$\text{Var}(\beta) = \sigma_{\beta}^2 = \sigma^2(X'X)^{-1}$$

Equazione 14

dove la notazione è la medesima che nell'Equazione 11.

L'Equazione 15, invece, descrive la varianza del coefficiente di regressione β ottenuto da un modello in cui tali ipotesi sono violate:

$$\text{Var}(\beta) = \sigma_{\beta}^2 = \sigma^2(X'X)^{-1}X'\Omega X(X'X)^{-1}$$

Equazione 15

dove, oltre la notazione definita in precedenza, il termine Ω rappresenta la matrice di varianza-covarianza dell'errore. Il semplice raffronto visivo tra le equazioni 14 e 15 suggerisce che la varianza espressa dall'Equazione 15 sia maggiore di quella espressa dall'Equazione 14. Infatti, l'Equazione 15 contiene il termine Ω , che rappresenta la matrice di varianza-covarianza dell'errore nel caso di covarianza spaziale non-nulla o non-stazionaria⁶⁷.

⁶⁷ Nel caso di distribuzione identica e indipendente dell'errore la matrice Ω è una matrice identità, ovvero, contiene solo elementi unitari lungo la diagonale principale e nulli fuori la diagonale. Invece, nel caso di eterogeneità spaziale gli elementi lungo la diagonale principale sono diversi da 1, mentre nel caso di dipendenza spaziale gli elementi fuori la diagonale sono diversi da 0. Entrambe le circostanze (che possono anche verificarsi simultaneamente) impediscono di semplificare l'Equazione 15, ciò che la ricondurrebbe all'Equazione 14.

2.5.10. Effetti spaziali e il problema dell'inferenza

La valutazione del modello di regressione e la verifica delle sue ipotesi poggiano su test statistici inferenziali⁶⁸. Consideriamo, come esempio, il semplice test t che valuta la significatività di ciascun coefficiente β , definito dall'Equazione 16.

$$t = \frac{\beta - 0}{\sigma_{\beta}}$$

Equazione 16

Nell'equazione, al denominatore del test compare la varianza del coefficiente β , definita dall'Equazione 14 nel caso in cui le ipotesi del modello siano soddisfatte, e dall'Equazione 15 nel caso non siano soddisfatte (a causa degli effetti spaziali). L'ipotesi nulla del test è che β sia uguale a zero ($H_0: \beta = 0$). Nel caso in cui il denominatore del test sia la varianza descritta dall'Equazione 15 (presenza di eterogeneità e/o dipendenza spaziale), l'inflazione della varianza di β comporta una maggiore tendenza a rifiutare l'ipotesi nulla quando essa è vera (errore di tipo I) rispetto al caso in cui le ipotesi di omogeneità ed indipendenza spaziale siano soddisfatte.

2.5.11. Il modello SAR in risposta all'autocorrelazione spaziale

Consideriamo, dapprima, un modello autoregressivo temporale⁶⁹, più intuitivo di quello spaziale. Qui il valore del processo in un dato momento è funzione dei valori che l'hanno preceduto, come mostra l'Equazione 17:

$$y_t = \rho y_{t-k} + \varepsilon_t$$

Equazione 17

⁶⁸ Burt, Barber, Rigby (2009); Rogerson (2014).

⁶⁹ È il caso tipico dei modelli econometrici; si, veda ad esempio, Johnston e DiNardo (2000).

dove ρ è un coefficiente di autocorrelazione, t denota un istante temporale, e k è un intervallo di tempo. Come già osservato, lo spazio, nella consueta rappresentazione, è definito da almeno due dimensioni oltre l'attributo, e in esso non è data una direzione preferenziale. Per questo, la trasposizione del modello temporale allo spazio richiede l'introduzione di un termine che definisca le relazioni spaziali tra le varie unità spaziali, ossia le diverse realizzazioni del processo. Il modello autoregressivo spaziale, noto anche come *SAR* (*Spatial AutoRegression*), è definito dall'Equazione 18:

$$y_i = \rho W y_j + \varepsilon_i$$

Equazione 18

dove si nota che, oltre al termine autoregressivo ρ definito nell'Equazione 17, compare anche un termine W , che serve a definire quelle relazioni spaziali. Infatti, W altro non è che la matrice dei pesi spaziali (Equazione 5). Analogamente al modello temporale (Equazione 17), in cui ogni realizzazione (Y_t) è funzione delle realizzazioni che l'hanno preceduta, (Y_{t-k}), nel modello spaziale ogni realizzazione (Y_i) è funzione di quelle che la circondano (Y_j). La matrice dei pesi spaziali, W , seleziona, tra le realizzazioni (Y_j), quelle legate alla dipendente (Y_i), e la loro influenza è modellata dalla funzione di attrito della distanza (Equazione 3).

Integrando il modello *SAR* nel modello di regressione multivariato (Equazione 10) e tornando per semplicità alla notazione matriciale si ottiene il modello descritto dall'Equazione 19:

$$Y = X\beta + \rho W Y + \varepsilon$$

Equazione 19

Manipolando in pochi semplici passaggi l'Equazione 19 si può mostrare come il modello autoregressivo spaziale gestisce il problema

dell'inflazione della varianza indotta dalla dipendenza spaziale⁷⁰. Con quei passaggi si ottiene l'Equazione 20.

$$Y = (1 - \rho W)^{-1} X\beta + (1 - \rho W)^{-1} \varepsilon$$

Equazione 20

Dall'Equazione 20 si definisce un termine Ω , o più precisamente la radice quadrata di Ω come da Equazione 21.

$$(1 - \rho W) = \sqrt{\Omega}$$

Equazione 21

Il modello *SAR* può allora essere riscritto sostituendo l'inverso della radice quadrata di Ω al termine autoregressivo ottenendo l'Equazione 22:

$$Y = \sqrt{\Omega}^{-1} X\beta + \sqrt{\Omega}^{-1} \varepsilon$$

Equazione 22

La formula per la stima dei coefficienti β conterrà allora al suo interno l'inverso di Ω , come mostrato dall'Equazione 23.

$$\beta = (X' \Omega^{-1} X)^{-1} (X' \Omega^{-1} Y)$$

Equazione 23

Si può mostrare facilmente che la varianza del coefficiente β del modello *SAR* è uguale a quella descritta dall'Equazione 14, ossia

⁷⁰ L'Equazione 19 può essere riscritta portando il termine autoregressivo a sinistra dell'uguaglianza:

$$Y - \rho WY = X\beta + \varepsilon$$

La formula può essere riscritta come segue:

$$Y(1 - \rho W) = X\beta + \varepsilon$$

Entrambi i termini dell'equazione possono poi essere divisi per $(1 - \rho W)$.

quella ottenuta quando le ipotesi del modello di regressione sono soddisfatte, come mostrato dall'Equazione 24:

$$\text{Var}(\beta) = \text{Var}((X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}Y)) = \sigma^2(X'X)^{-1}$$

Equazione 24

La stima dei coefficienti di regressione del modello *SAR* non può avvenire con il metodo OLS, poiché la presenza della variabile dipendente tra le esplicative (nel termine autoregressivo) comporta la violazione di alcune delle ipotesi. Perciò questo modello si stima generalmente il metodo della massima verosimiglianza (*maximum likelihood*, o *ML*), mentre i tradizionali test inferenziali *t* ed *F*, che poggiano sull'ipotesi di normalità dell'errore, sono sostituiti dalla statistica di Wald e dal test del rapporto della verosimiglianza (*likelihood ratio test* o *LRT*)⁷¹. La stima del modello richiede inoltre alcune ipotesi sulla struttura autoregressiva del modello: le principali ipotesi considerate in letteratura sono (1) che sia la variabile dipendente a manifestare autocorrelazione spaziale (è questo il cosiddetto *spatial lag model*, ossia il caso del modello descritto dall'Equazione 19; (2) che sia l'errore o residuo del modello a manifestare autocorrelazione spaziale (*spatial error model*)⁷².

Con il metodo *SAR*, in entrambe le specificazioni, la struttura spaziale o autocovarianza del processo spaziale viene così scorporata dal modello di regressione tramite l'inverso del termine Ω , che altro non è che la matrice dei pesi spaziali (Equazione 5) moltiplicata per il coefficiente autoregressivo spaziale (Equazione 21).

2.5.12. Il modello *GWR* in risposta all'eterogeneità spaziale

Il modello *GWR* (*geographically weighted regression*), espresso dall'Equazione 25, è matematicamente molto più semplice del modello *SAR*.

⁷¹ Anselin (1988).

⁷² Anselin (2005, pp. 93-99); Bivand, Pebesma, Gómez-Rubio (2008, pp. 290-303).

$$y_i = \beta_0(u_i v_i) + \sum_k \beta_k(u_i v_i) x_{ik} + \varepsilon_i$$

Equazione 25

L'equazione rappresenta un modello identico alla regressione standard (Equazione 10)⁷³, con l'unica differenza che nell'Equazione 25 sono precisate le coordinate geografiche di ciascun coefficiente β , (u_i e v_i). Nella regressione *GWR*, infatti, i coefficienti variano nello spazio e questo perché il modello *GWR* stima una molteplicità di regressioni locali. Per definire i modelli locali, *GWR* utilizza una matrice, detta W , diversa e più semplice della matrice dei pesi spaziali (Equazione 5), nonostante sia indicata dal medesimo simbolo. La matrice W di *GWR* ha il solo scopo di selezionare le unità spaziali circostanti ciascuna unità, che saranno usate per stimare ogni regressione locale⁷⁴. Il modello *GWR* viene stimato semplicemente col metodo OLS (salvo specificazioni di *GWR* diverse dalla lineare), poiché le proprietà matematiche e statistiche delle regressioni locali sono le medesime di quella globale. Il parametro che definisce la dimensione del *kernel* (raggio per il *kernel* fisso e numero di vicini per quello adattivo) è chiamato larghezza di banda (*bandwidth*). Per la sua definizione si utilizzano dei criteri che ottimizzano i risultati di *GWR*, i più comuni dei quali sono AIC (*Akaike Information Criterion*) e CV (*cross-validation*)⁷⁵. Utilizzando i vari tipi di *kernel* è possibile frazionare l'area di studio in svariati modi; tuttavia, è prassi comune definire tante regressioni locali quante sono le unità spaziali del campione, ovvero una regressione per ciascuna unità campionata.

2.6. Oltre i metodi quantitativi: i limiti delle soluzioni parziali

Come si è visto nei paragrafi precedenti, l'analisi spaziale ha gradualmente preso consapevolezza dei cosiddetti effetti spaziali, ossia

⁷³ Nell'Equazione 25 si è usato il simbolo di sommatoria, mentre nell'Equazione 10 si sono elencati i termini della somma.

⁷⁴ Fotheringham, Brunson, Charlton (2003).

⁷⁵ Fotheringham, Brunson, Charlton (2003).

della violazione delle ipotesi del modello di regressione indotta dalle proprietà intrinseche dei processi spaziali: dipendenza (o autocorrelazione) ed eterogeneità (o non stazionarietà) spaziali. Tali violazioni, come si è illustrato, gonfiano la varianza delle stime del modello, compromettendone anche le procedure inferenziali.

Nonostante i progressi compiuti, né il modello *SAR* né il modello *GWR* risolvono completamente i problemi creati dagli effetti spaziali. Il modello *SAR* allevia le conseguenze della dipendenza spaziale solo in termini computazionali. Analogamente, il modello *GWR* lascia irrisolte questioni riguardanti la scarsa numerosità campionaria e la potenziale multicollinearità delle regressioni locali⁷⁶. Collegato a quest'ultimo inoltre è il problema della significatività delle variabili: poiché anche la selezione del modello avviene a livello globale, è possibile che una o più delle variabili esplicative risultino non-significative nelle regressioni locali, fino al paradosso che alcune regressioni possono contenere solo variabili non-significative. Da ultimo, consideriamo il problema, che è forse il più serio ma anche il più interessante dal punto di vista concettuale per l'analisi spaziale: nelle microaree selezionate dal *kernel* (o in alcune di esse) l'autocorrelazione spaziale del processo potrebbe essere esacerbata proprio dalla loro dimensione ridotta, per cui se tutte le unità spaziali si trovano a distanze molto piccole tra loro, sono quindi più suscettibili di autocorrelazione spaziale. Questa circostanza implica che *GWR* non solo ignora la dipendenza spaziale (al pari di *SAR* per l'eterogeneità) ma potrebbe addirittura aggravarla.

Dunque, in assenza di una soluzione complessiva agli effetti spaziali, ciascun metodo allevia l'inflazione di varianza generata da una proprietà spaziale ma, nel migliore dei casi, ignora quella generata dall'altra; nei casi peggiori l'uso di una metodologia per gestire una forma di covarianza non solo trascura, ma addirittura aggrava l'altra forma. Benché le due proprietà siano intimamente legate, poiché entrambe discendono dalla covarianza del processo spaziale, esse possono manifestarsi con modalità distinte, e comunque vengono identificate da procedure analitiche diverse. Per la loro risoluzione,

⁷⁶ Generalmente la correlazione tra variabili viene testata solo a livello globale, ma se presente in una o più delle regressioni locali può generare multicollinearità.

storicamente, si fece ricorso ai metodi sviluppati in altre discipline⁷⁷, che furono adattati al contesto spaziale. In realtà, autocorrelazione ed eterogeneità spaziale sono diverse manifestazioni quantitative di quelle che sono anche caratteristiche qualitative dei processi spaziali. L'insoddisfazione rispetto alle tecniche di cui disponiamo a tutt'oggi (*SAR* e *GWR*) deriva non tanto dal fatto che nessuna ne risolve completamente la manifestazione quantitativa, ma soprattutto dal fatto che nessuna affronta i problemi concettuali e qualitativi che ne stanno alla base. Solo affrontando compiutamente tali aspetti, partendo non da altre discipline, ma da una prospettiva rigorosamente spaziale, sarebbe, forse, possibile dare una soluzione soddisfacente anche ai problemi statistici e quantitativi. All'inizio di questo secolo Stewart Fotheringham e colleghi (2000) parlavano di soluzioni prese a prestito da altri, auspicando un'inversione di rotta per cui sarebbe stata in futuro l'analisi spaziale a costruire soluzioni da prestare ad altre discipline. Analogamente Paul Longley (2000) si soffermava sulla transizione, che stimava imminente, da una geografia quantitativa che tradizionalmente prende a prestito tecniche da altre discipline, alla possibilità di sviluppare un paradigma analitico spaziale proprio, esportandolo poi ad altre discipline.

Anche se, una ventina d'anni dopo, non vi sono segni evidenti che questo programma ambizioso sia stato realizzato, i progressi compiuti dall'analisi spaziale, di cui si discuterà nei prossimi capitoli, spaziano dalla critica della rappresentazione all'impiego di metodologie miste, facendo supporre che quell'inversione di rotta non sia poi così lontana dal compiersi.

⁷⁷ In particolare l'econometria.

3. Oltre i metodi quantitativi: relatività dell'analisi spaziale

Everything is related to everything else, but near things are more related than distant things.

Tobler, *A Computer Movie Simulating Urban Growth in the Detroit Region*, 1970, p. 236.

3.1. Premessa

Nel capitolo precedente abbiamo riassunto le proprietà caratteristiche dei processi spaziali, ossia autocorrelazione (dipendenza) ed eterogeneità (non-stazionarietà) spaziali, concentrandoci sui cosiddetti “effetti spaziali”, ossia i problemi che proprio quelle caratteristiche pongono per l'analisi quantitativa e statistica, la cui soluzione costituisce il nucleo dell'analisi spaziale. Abbiamo concluso quel capitolo soffermandoci su quelli che appaiono a tutt'oggi i due limiti principali di quelle soluzioni. Il primo è relativo al fatto che, nonostante gli effetti spaziali tendano a manifestarsi simultaneamente, esistono, ad oggi, solo soluzioni distinte che ne gestiscono ciascuno separatamente; inoltre, la soluzione all'autocorrelazione spaziale (il modello *SAR*) si limita ad ignorare il problema dell'eterogeneità spaziale, mentre la soluzione all'eterogeneità (il modello *GWR*) può addirittura aggravare il problema dell'autocorrelazione spaziale. Il secondo limite è che entrambe le soluzioni risolvono il problema – al più – dal punto di vista statistico, lasciando tuttavia irrisolte importanti questioni concettuali: per questo aspetto – contrariamente al primo – la critica principale è mossa al modello *SAR*, la cui soluzione consiste nello scorporre dall'analisi statistica la componente spaziale del modello.

Come è stato auspicato già da molto tempo¹, l'analisi spaziale necessiterebbe di una soluzione organica relativa agli “effetti spaziali” o, meglio, di una teoria sviluppata dal suo interno, capace di affrontare

¹ Fotheringham, Brunson, Charlton (2000).

gli effetti spaziali simultaneamente. Il compito, di per sé non facile, consentirebbe di superare almeno il primo dei due limiti.

Per affrontare il secondo limite, è importante spingersi oltre gli aspetti quantitativi dei processi spaziali, e comprenderne le proprietà per ciò che effettivamente sono: si tratta infatti di caratteristiche qualitative, intrinseche e innate, ovvero ben più che mere manifestazioni quantitative stimabili, quantificabili, o persino misurabili. Comprendere la natura qualitativa delle proprietà dei fenomeni spaziali potrebbe consentirne una migliore quantificazione e rappresentazione anche numerica, che potrebbe, potenzialmente, superare quel limite.

Si comprende quindi il legame profondo tra aspetti quantitativi e qualitativi non solo dei processi spaziali, ma dell'intera analisi spaziale. In questo capitolo, dunque, ci soffermiamo proprio su questi aspetti qualitativi e concettuali, che consentiranno anche di mettere meglio a fuoco l'analisi spaziale, nella sua portata teorica e in ciò che la differenzia da approcci quantitativi più tradizionali.

3.2. Dalla regionalizzazione all'aggregazione spaziale

La stagione della geografia teoretico-quantitativa fu un fenomeno di vasta portata che, nato nel Nord America, non lasciò indifferenti l'Europa e l'Italia². Non è certo possibile qui analizzare il fervore di idee, la molteplicità di applicazioni e la vivacità di quel dibattito sviluppatosi in ambito europeo, per questo ci limitiamo a ricordare la creazione, nel 1978, dell'*European Colloquium on Theoretical and Quantitative Geography (ECTQG)*, voluto perché «Gli sviluppi della “geografia teorica e quantitativa” interessavano un numero crescente di accademici in Europa»³. Il relativo ritardo con cui la geografia teoretico-quantitativa si affermò nei Paesi europei non anglo-sassoni in

² In Italia gli studi e le applicazioni dei metodi quantitativi fiorirono soprattutto negli anni '80: si vedano gli interventi di Celant, Ruggiero e Skonieczny, Staluppi, Tinacci Mossello, Turco, Vlora, nella terza sezione del convegno di Varese in Corna Pellegrini e Brusa (1980); i vari interventi in Pagnini (1985) ed i lavori di Zanetto e Lando (1980), Odd-Ambrosetti, Turco, Zanetto (1983) e Zanetto (1979; 1982 e 1986).

³ Pumain (1991, p. 14): «The developments of “theoretical and quantitative geography” interest a growing number of scholars in Europe since the beginning of the seventies. From 1978 on, European geographers decided to meet regularly, about every other year, in order to share their experiences in the field».

qualche modo ne accorciò la vicenda: infatti proprio in quegli anni, in parallelo alla maturazione delle spinte quantitative, cominciarono ad emergere altri fermenti all'interno della geografia, che divisero la disciplina, separando i quantitativi dai teorizzatori sociali, dando luogo ad una divisione che si sarebbe sanata solo in seguito⁴.

Più che i dibattiti sulla *Theoretical and Quantitative Geography*, qui ci interessa analizzare le problematiche legate a quell'importante filone di ricerca connesso alla suddivisione territoriale generalmente definito come *regionalizzazione*. Fu un dibattito che lasciò le tracce più profonde in ambiente anglosassone⁵, dove, con classico pragmatismo, molto lavoro è stato dedicato alla definizione funzionale di regioni, quali i distretti sanitari, scolastici e, tra i più interessanti, i collegi elettorali⁶. A questo lavoro possono essere ricondotte alcune tra le linee di ricerca che più avrebbero influenzato la geografia quantitativa degli ultimi anni. Il lavoro emblematicamente intitolato "*A Million or so Correlation Coefficients*" di Stan Openshaw e Paul Taylor (1979) pose, per certi versi, fine al dibattito sulla regionalizzazione: in quell'articolo gli autori manifestavano una lucida presa di coscienza di quanto i risultati quantitativi possano variare in risposta alla modifica delle unità territoriali di riferimento.

Il problema dell'aggregazione delle unità spaziali in geografia⁷, o, secondo la terminologia della letteratura anglosassone, problema dell'unità areale modificabile (*modifiable areal unit problem*, o *MAUP*) riconosce che per la gran parte delle analisi statistiche, fondate sul coefficiente di correlazione, i risultati analitici cambiano se si cambiano le unità geografiche di riferimento⁸. Se si conduce un'analisi a scala regionale, provinciale o comunale, i risultati cambiano innanzitutto perché cambia il numero di unità considerate, ma la questione va ben oltre gli aspetti quantitativi e numerici, come illustra ancora

⁴ Si veda ad esempio Poorthuis e Zook (2020); Franklin *et Alii* (2021).

⁵ Anche in Italia si è posto il "problema della regionalizzazione" dando origine nel 1980 al gruppo di lavoro A.Ge.I. "Teorie e metodi della regionalizzazione" guidato da Adalberto Vallega, gruppo che, conclusa la sua attività nella seconda metà degli anni ottanta, produsse un numero speciale della *Rivista Geografica Italiana* del 1982 (vol. LXXXIX, fasc. 2) e un Colloquio *internazionale* (Verona 27-28 gennaio 1983) i cui risultati sono confluiti in Turco (1984); significativo è poi il lavoro di Odd-Ambrosetti, Turco, Zanetto (1983).

⁶ Si veda a questo proposito il paragrafo 3.

⁷ Definito nel capitolo 2.

⁸ Openshaw (1977).

meglio la seconda questione del *MAUP*: il coefficiente di correlazione, come gli altri risultati quantitativi, cambia anche quando si modificano i confini delle unità territoriali, pur lasciandone inalterato il numero. Se, ad esempio, si ridisegnassero i confini delle regioni italiane, col vincolo di mantenerne comunque 20, ma aggiungendo alla Lombardia la provincia di Verona e togliendo alla Toscana quella di Massa-Carrara, ciò modificherebbe i risultati analitici basati sul coefficiente di correlazione, pur non avendo alterato il numero delle unità considerate (in questo caso le regioni). Continuando a spostare anche una sola provincia alla volta, si otterrebbe facilmente quel “milione o giù di lì” di risultati analitici. Questa molteplicità di risultati è chiaramente legata all’eterogeneità spaziale, perché se i processi osservati nella provincia di Verona fossero identici a quelli osservati in quella di Brescia, lo spostamento di quel territorio da una regione all’altra non altererebbe i risultati analitici, a parità di unità spaziali.

Il “giochino” di spostare le province ha, in apparenza, poco senso se si vuole condurre una ricerca a scala regionale, ammenoché lo scopo non sia proprio la ridefinizione delle unità amministrative esistenti. È pur vero, tuttavia, che l’accettazione acritica di quelle unità amministrative conduce a privilegiare, tra il “milione o giù di lì” di risultati possibili, proprio quello determinato da quella particolare suddivisione del territorio: è ciò che in passato ha suscitato reazioni critiche nei confronti della geografia quantitativa, che, in tal modo, rafforzerebbe lo *status quo*, assumendo, anche inconsciamente, una posizione politica conservatrice⁹. A questo va peraltro aggiunto che il “giochino”

⁹ A proposito si veda Harris e Weiner (1998) che, tra l’altro, riportano una ricca bibliografia sulla discussione attorno al ruolo e impatto sociale del GIS, p. 68: «We contend that a problematic dualism exists whereby GIS production and use is understood to be either empowering or marginalizing. This has contributed to the polemical nature of recent GIS and Society debate (Abler 1993, Aitken and Michel 1995, Curry 1994, Goodchild 1991, Goss 1995, Lake 1993, McHaffie 1995, Onsrud and Rushton 1995, Obermeyer and Pinto 1994, Openshaw 1991 and 1992, Pickles 1991 and 1995, Sheppard 1993a 1993b and 1995, Smith 1992, Taylor Overton 1991, Taylor and Johnston 1995). We argue instead that GIS is a contradictory technology that simultaneously marginalizes and empowers people and communities. The social and environmental impacts of GIS are therefore contingent upon a particular mix of historical, socioeconomic, political, and technological conditions in particular places. To date, there has been a noticeable dearth of such contextual GIS and Society case study research». Sui condizionamenti e le distorsioni legate ai confini comunali nelle analisi sulla localizzazione industriale si veda anche Lando (1985, p. 246, corsivo dell’autore): «Si considerano *spazialmente associate* (con misura ed entità da trovare) le attività economiche localizzate entro la medesima unità territoriale [stesso comune] mentre non sono considerate tali se sono ubicate in due

non può esaurirsi con tutti gli spostamenti possibili delle province, ma dovrebbe considerare spostamenti di comuni, singolarmente o a gruppi, e addirittura di porzioni di territorio non delimitate da confini amministrativi. Vi sono poi analisi geografiche che non sono condotte su unità amministrative date, ma su unità spaziali definite da altri criteri. L'Esempio 3.1 illustra la definizione di unità spaziali per l'analisi dell'inquinamento marino. Anche in questo caso i risultati dell'analisi sono condizionati non solo dalla porzione di oceano considerata e dalla dimensione delle celle, ma anche dal posizionamento della griglia¹⁰.

La presa di coscienza dell'esistenza di quel "milione o giù di lì" di coefficienti di correlazione porta ad ammettere la relatività di qualsiasi risultato quantitativo su un qualsiasi numero di unità spaziali. Non esiste quindi una regione, o un'unità spaziale 'aurea' alla quale i fenomeni geografici possano essere misurati ed analizzati¹¹. Esistono invece processi spaziali che interagiscono tra loro e concorrono su più scale scavalcando i confini delle unità spaziali, comunque esse siano definite. Per questo, la scelta di quali unità debbano essere campionate ed analizzate dovrebbe essere determinata dal processo spaziale analizzato e dalla questione di ricerca. Tuttavia, tale scelta è spesso condizionata anche da questioni pratiche legate all'esistenza, reperibilità ed affidabilità dei dati relativi a quel fenomeno. Così, ad esempio, le unità censuarie (in Italia come altrove) non sono necessariamente le più adatte a rappresentare processi spaziali pertinenti a dinamiche socioeconomiche, tuttavia il loro uso viene spesso privilegiato poiché

diverse unità territoriali. Questo anche se la distanza che le separa può essere, nel secondo caso, minore (attività contigue in quanto poste una a fianco all'altra su due comuni confinanti) che nel primo (attività poste agli estremi di un medesimo comune) [...] tali distorsioni saranno tanto più rilevanti quanto meno la delimitazione comunale risponde a criteri economico funzionali: e ben si sa che i confini comunali non sono "disegnati" con quei criteri».

¹⁰ Mathur e Bertazzon (2009) considerano la questione del posizionamento ottimale di una griglia di maglie regolari, ossia minimizzando l'errore (o differenza) tra i valori di una variabile osservata su unità censuarie rispetto a quelli osservati su un numero di griglie alternative (p. 267): «The contribution of this work consists of the development of a methodology for the rasterization of census units, which can be optimized according to some set criteria. The construction of an optimal grid may offer a practical answer to many analytical and planning problems. Additionally, in the process of constructing the grid, the population distribution is split into pixels, which provides a homogeneous raster to merge and analyze census data and remotely sensed data».

¹¹ Su questo si veda l'analisi che ne fa Vallega (1984).

contengono dati affidabili, a scala spaziale dettagliata, e ad intervalli temporali regolari.

Ad oggi, il problema dell'unità areale modificabile non ha soluzione. La ricerca applicata generalmente ne è consapevole e tende a gestire il problema dando conto, con trasparenza e in dettaglio, della scala e aggregazione usate, e ribadendo che la validità dei risultati ottenuti può essere affermata solo entro quei parametri, e che non vi è ragione per pretendere che i risultati siano validi ed estensibili ad altre scale o a diverse aggregazioni¹². Se questo tipo di approccio comporta un'accettazione della relatività dei risultati analitici, rimane insoluta la questione dell'uso privilegiato di certe unità, quali le unità censuarie o quelle amministrative, privilegiando quindi, tra tutti quelli possibili, il solo risultato analitico determinato dal 'potere costituito'. Si tratta, per la gran parte delle analisi applicate, di una scelta dettata solo da questioni pratiche, prima fra tutte la disponibilità e qualità dei dati, ma tuttavia non priva di conseguenze 'politiche' quando i risultati analitici vengano utilizzati a scopo di pianificazione o comunque come supporto alle decisioni, oltre il fatto che tende a trascurare istanze territoriali non consolidate¹³. La questione è solo apparentemente meno rilevante quando vengano analizzati fenomeni fisici anziché socioeconomici; tuttavia, come evidenziato anche dall'Esempio 3.1, non è sempre facile separare fenomeni fisici da entità territoriali politiche, e quindi anche in questi casi la scelta delle unità di analisi potrebbe avere conseguenze anche politiche in termini di pianificazione e supporto alle decisioni.

ESEMPIO 3.1. Inquinamento da idrocarburi lungo la costa Pacifica canadese¹⁴

Il programma del governo canadese National Aerial Surveillance Program¹⁵ esegue ricognizioni sistematiche delle acque territoriali, negli oceani di pertinenza, tramite l'impiego di velivoli equipaggiati con

¹² Si veda ad esempio Bertazzon *et Alii* (2015) e la discussione sulle ipotesi in Tallone e Lando (2009) in particolare alle pp. 162-163 e 179-184.

¹³ Ad esempio, da parte di popolazioni native nelle Americhe o in Oceania. Nel caso italiano ci si può riferire alle difficoltà riscontrate nella definizione del "territorio del distretto industriale": un accenno a questo si trova in Tattara e Volpe (2001).

¹⁴ Questo studio è illustrato in maggior dettaglio nel capitolo 7. Si veda anche Bertazzon *et Alii*, 2014.

¹⁵ <https://tc.canada.ca/en/programs/national-aerial-surveillance-program>

fotocamere ad alta risoluzione. Oltre ad altre informazioni¹⁶, il NASP fornisce foto aeree georeferenziate delle chiazze oleose negli oceani. Tali informazioni vengono utilizzate non solo allo scopo di proteggere l'ambiente marino e le specie più a rischio, ma anche per identificare e sanzionare i comportamenti che violano le norme sull'inquinamento.

Col finanziamento dei Dipartimenti Trasporti e Ambiente del governo federale venne costituito il gruppo di lavoro OCW (Oil in Canadian Waters)¹⁷ che elaborò un modello di regressione multivariata per esaminare l'associazione tra chiazze oleose ed attività marittime lungo la costa Pacifica.

L'analisi esplorativa dei dati NASP esemplifica interessanti problematiche relative all'analisi spaziale.

- La variabilità e mobilità dell'ambiente marino vanifica qualsiasi analisi della forma geometrica ed estensione spaziale delle chiazze: per questo il dato analizzabile è il semplice conteggio delle chiazze, enumerate in ciascuna unità spaziale.
- Considerate la dimensione della *economic exclusive zone* (EEZ) (Fig. 3.1), il numero complessivo e la distanza delle chiazze riportate per il periodo 2008-2010, la EEZ fu suddivisa in celle di 5×5 Km. L'area di studio conteneva così 101 chiazze sparse su 37.162 celle (alcune nella medesima cella); la maggior parte delle unità spaziali (celle) conteneva quindi 0 chiazze (osservazione nulla).
- Per isolare i processi d'interesse e ridurre il numero di celle con osservazioni nulle, il gruppo definì il *minimo poligono convesso*¹⁸ delimitato dalla linea di costa e comprendente le osservazioni non-nulle, i cui *buffer* furono stabiliti in 25 Km dalla costa e 50 Km dalle chiazze più al largo. Nonostante la drastica riduzione, l'area di studio conteneva ancora 3,414 celle e 96 osservazioni non-nulle.
- ***Anche nell'area ridotta, secondo l'analisi statistica tradizionale, la distribuzione di probabilità delle chiazze oleose si configurava come una distribuzione Poisson, ossia casuale.***

¹⁶ In particolare, relative alla sicurezza, come l'eventuale presenza di imbarcazioni non autorizzate.

¹⁷ Gruppo di lavoro formato, tra gli altri, da esperti di università, dipartimenti governativi e autorità costiere; parte dei suoi risultati sono apparsi in Bertazzon *et Alii* (2014).

¹⁸ Questo, come il successivo *buffer*, sono ottenuti con operazioni di GIS, discusse nel capitolo 4.

- Solo l'analisi statistica spaziale evidenziava la presenza, statisticamente significativa, di autocorrelazione (Moran's I) e clustering (Getis and Ord G_i^*) spaziali.
- L'evidenza di autocorrelazione ed eterogeneità spaziale richiede l'impegno di metodi analitici spaziali (SAR o GWR) nella regressione Poisson¹⁹, per alleviare l'inflazione della varianza indotta dagli effetti spaziali.

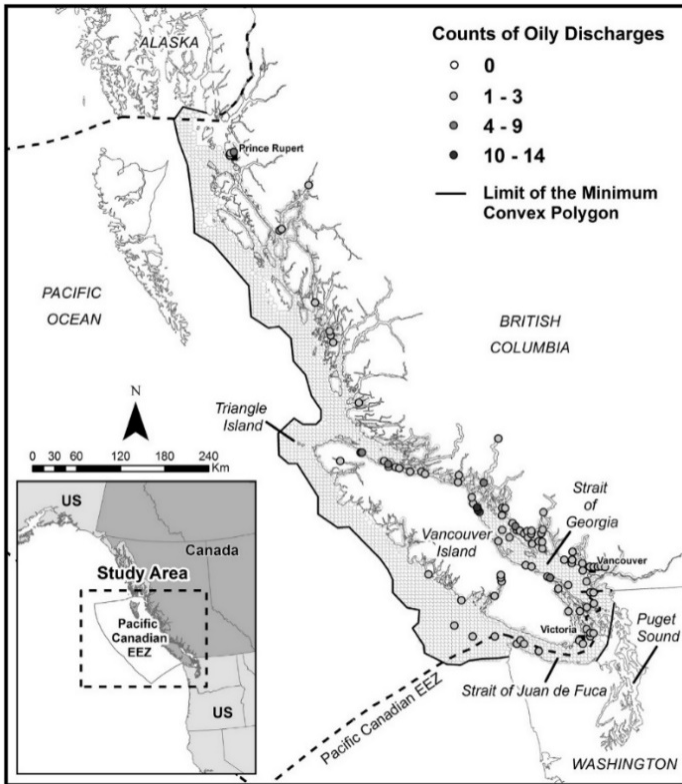


Fig. 3.1 - Analisi spaziale delle chiazze oleose nel Pacifico canadese.
 Riprodotta da <https://doi.org/10.1016/j.apgeog.2014.04.010> su licenza CC-BY-NC-ND

Nell'ambito dello studio sintetizzato nell'Esempio 3.1 furono condotti diversi esperimenti, modificando la dimensione delle celle e la

¹⁹ L'impiego del modello Poisson è imposto dalla distribuzione statistica della dipendente.

posizione della griglia. Tali esperimenti chiaramente evidenziavano non solo il fenomeno *MAUP*, ma anche il cambiamento dei parametri e delle proprietà della distribuzione statistica in risposta a cambiamenti nel numero e nella disposizione delle unità spaziali. Osservando la Fig. 3.1 è evidente che le chiazze oleose non sono distribuite casualmente nella EEZ, e neppure nell'area di studio ridotta, ma manifestano una presenza intensificata (*clustering*), particolarmente nella zona tra gli stretti di Georgia e di Juan de Fuca, in relazione alle attività commerciali e turistiche del porto di Vancouver. Tuttavia, l'analisi statistica tradizionale avrebbe indotto alla conclusione che la distribuzione fosse casuale; su questa base, si sarebbero condotti modelli analitici tradizionali, le cui stime sarebbero state affette da eccesso di incertezza, ossia inflazione della varianza. Questo studio, invece, fornisce un esempio importante di come l'analisi statistica spaziale sia in grado di individuare e valutare statisticamente la presenza e l'entità delle proprietà dei dati spaziali e, di qui, di procedere all'impiego di modelli specifici, capaci di alleviare l'incertezza delle stime indotta dagli effetti spaziali.

3.3. La dipendenza spaziale e le leggi della geografia

L'importanza dei concetti di dipendenza ed autocorrelazione spaziale si può meglio apprezzare ricordando come Waldo Tobler, nel clima della rivoluzione quantitativa, all'inizio degli anni '70 arrivò non solo a formulare, ma persino ad invocare ciò che da lui prese il nome di 'prima legge della geografia': "ogni cosa è collegata ad ogni altra cosa, ma le cose vicine sono più collegate delle cose distanti", e che rimane, se non una legge, un'efficace definizione operativa ed intuitiva della dipendenza spaziale. La ricerca di leggi universali derivava forse dal desiderio di emulare le cosiddette scienze dure, quali la fisica e la chimica, sul cui linguaggio, infatti, appare modellata la "legge di Tobler": legge che rimane forse la più emblematica manifestazione dello spirito di quel tempo²⁰.

A mezzo secolo dalla sua enunciazione è interessante interrogarsi sulle ragioni della persistenza di riferimenti alla prima (e ad ulteriori)

²⁰ Tobler (1970) come discusso nel capitolo 2. Se ne veda anche l'analisi in Miller (2004).

leggi della geografia, che nella letteratura contemporanea appaiono persino in ambiti meno quantitativi della geografia²¹, quando ormai la stessa analisi spaziale ha rinunciato ad ogni ambizione di enunciare o dimostrare leggi universali; ed anche sulle ragioni per cui la più paradossale delle leggi enunciate durante la rivoluzione quantitativa appare più viva che mai. Nel 2003 mi trovavo io stessa a partecipare alla conferenza dell'AAG in cui si tenne una sessione speciale dedicata a quella prima legge, *The First Law of Geography*²², i cui contributi sono raccolti in un volume (Sui, 2004) che uscì l'anno successivo negli *Annals* dell'AAG. Di quel dibattito ricordo in particolare un'affermazione di Michael Goodchild, il quale, si noti, prima di specializzarsi in geografia si era laureato in fisica²³. Secondo Michael Goodchild l'affermazione di Waldo Tobler non può assolutamente essere considerata una legge, poiché di legge non ha le caratteristiche, il rigore, i presupposti, né la generalità; ma con altrettanta certezza bisogna riconoscere che quella relazione è universalmente vera, entro la scala geografica. In sintesi: a meno che non si considerino fenomeni infinitamente piccoli (di tipo microscopico, ad esempio molecolari) oppure infinitamente grandi (di tipo macroscopico, ad esempio cosmici), non si riesce a provare la falsità di quell'affermazione²⁴. Pur consapevoli che di legge non si tratta, generazioni di geografi continuano a far riferimento alla "legge" come se tale fosse, forse per semplice convenienza e convenzione²⁵.

Come affermato da Michael Goodchild, la prima legge cela un contenuto generalmente vero e di fondamentale importanza per l'analisi spaziale e lo definisce in maniera semplice, intuitiva e priva di complessità matematiche²⁶. Essa afferma ciò che in epoca più recente e

²¹ Per esempio, la geografia cognitiva in Klippel, Hardisty, Li (2011).

²² [www.aag.org/galleries/conferencefiles/AAG2003 Printed_Program.pdf](http://www.aag.org/galleries/conferencefiles/AAG2003%20Printed_Program.pdf)

²³ www.geog.ucsb.edu/~good/

²⁴ Goodchild (2004, p. 301): «Perhaps the easiest way to see the validity and value of TFL is through a thought experiment, by trying to imagine a world in which it is not true». Si veda anche l'articolo di Sui (2004) nel cui titolo è sintetizzata l'intera posizione.

²⁵ Tutti gli studenti di geografia incontrati in una ventina d'anni di insegnamento all'università di Calgary conoscevano la prima legge, e questa conoscenza e l'immediatezza del riferimento mi sono state molto utili nell'insegnamento dell'analisi spaziale.

²⁶ Goodchild (2004, pp. 302-303): «TFL is stated in a charmingly informal way, and one wonders if it would be taken more seriously if it were formalized and made a little less accessible. It might state, for example, that for every geographic variable (a function of location $z = f(x)$) there exists some distance d below which covariance is monotonically increasing- or

con linguaggio più rigoroso viene chiamato “dipendenza spaziale”, ossia una di quelle proprietà intrinseche che, rendendo *speciali* i dati *spaziali*²⁷, causano quegli effetti spaziali che minano l’affidabilità delle stime dei modelli, di cui si occupa l’analisi spaziale. Ciò che forse più stride col sentire del geografo contemporaneo è l’idea di “invocare la prima legge” come fece cinquant’anni fa, Waldo Tobler, e come oggi ci aspetteremmo forse da un chimico o da un fisico, ma non certo da un geografo.

Oggi i geografi non invocano leggi, ma misurano relazioni, che definiscono con nomi meno altisonanti, consapevoli che la misurazione, la stima, il calcolo di un indice, possono servire solo ad asserire o suffragare la validità di una relazione per il caso empirico in esame, senza però potere (né volere) dimostrare alcuna legge generale. Smorzati gli echi della rivoluzione quantitativa coi suoi immancabili eccessi, la geografia teoretico-quantitativa seppe crescere ed evolvere, come dimostra, tra l’altro, proprio l’accettazione della relatività dei risultati analitici rispetto all’aggregazione spaziale. E al lavoro di quegli anni rimane il merito di aver richiamato l’attenzione su quelli che poi sarebbero stati chiamati effetti spaziali e che oggi formano il nucleo della ricerca in analisi spaziale.

L’autocorrelazione spaziale misura o, meglio, stima in termini quantitativi la similitudine degli attributi di un processo spaziale in rapporto alla loro distanza: non ne spiega le ragioni. Per apprezzare i limiti di tale valutazione meramente quantitativa si può partire dall’esempio fornito da Michael Goodchild, che ne dà una spiegazione molto lucida in riferimento alla distribuzione spaziale degli immigrati italiani nella città di London (Ontario, Canada)²⁸. La distribuzione si

that there exists at least one scale for which spatial autocorrelation is positive. Formalization might address some of the reluctance to give TFL greater stature; but, at the same time, it would reduce both its accessibility and its charm. [...] But TFL seems rather different and more in the style of the traditional laws of science, which are judged primarily for their empirical validity and obtain their value from their simplicity and their practical usefulness».

²⁷ Il corsivo fa riferimento all’affermazione di Goodchild (2001, p. 1): «While geographic inherits many of its properties from spatial, it also adds new ones, and thus specializes the definition. If “spatial is special” [...] then geographic should be even more special, and a theory of geographic information should be distinct from a theory of spatial information, inheriting all of the generality of the latter, but adding its own specifics».

²⁸ Goodchild (1986, p. 42): «An attribute variable can show spatial autocorrelation in its arrangement across spatial objects either because neighbouring objects influence each other directly, so that the value at one place is caused directly by values at neighbouring places

presenta positivamente autocorrelata, poiché gli immigrati italiani tendono a risiedere in prossimità di altri italiani. Ma perché gli immigrati italiani tendono a risiedere in prossimità di altri connazionali immigrati? Potrebbe trattarsi di un processo *reattivo*, ossia gli immigrati si stanziavano in prossimità di connazionali, oppure di un processo *interattivo*, ossia, gli immigrati si stanziavano in un certo luogo in virtù di altri processi, ad esempio la disponibilità di certi tipi di lavoro, prezzi delle abitazioni accessibili, etc., e poiché in generale gli immigrati italiani hanno simili esigenze lavorative e livelli di reddito, tendono a stanziarsi tutti nella stessa zona. Analogamente, nell'Esempio 2.1²⁹ l'incidenza di malattie cardiovascolari manifesta autocorrelazione spaziale positiva, ma in realtà si tratta di un processo interattivo poiché la relazione tra la prossimità residenziale e l'incidenza di malattie cardiache rivela anche un altro processo: la prossimità residenziale di persone di età simile. È quindi il processo spaziale "età media dei residenti" a manifestare autocorrelazione spaziale positiva, un fenomeno abbastanza comune nelle città del nord America, dove l'età media dei residenti è legata al periodo della lottizzazione: se quegli abitanti non hanno mai cambiato residenza è chiaro che con l'andar del tempo anche la loro età media aumenta e, parallelamente, il rischio di malattie cardiocircolatorie.

L'autocorrelazione spaziale positiva, ossia l'attrazione manifestata dai processi spaziali nei casi in cui gli attributi vicini tendono ad essere più simili, è non solo la situazione più frequente, ma anche la più studiata. Viceversa, la repulsione, ossia il caso in cui gli attributi vicini

(autocorrelation), or because the value at each place is determined by some other variable at the same place which is itself autocorrelated. For example, we do not know whether the Italian ethnic group in the city of London, Ontario (Figure 2c) shows autocorrelation at the Census tract level because people of Italian origin are attracted to areas near other people of the same ethnic group, or because the factor which attracts them is also present in other areas with large Italian populations. Cliff and Ord (1981, pp. 141-ff.) refer to the two interpretations as interactive and reactive respectively. Suppose the process which generated the attributes is solely reactive. Then if all of the causative factors can be found and modelled, the residuals from this model will be completely lacking in spatial autocorrelation. So if a test for autocorrelation of residuals is negative, and if we are willing to ignore the possibility of a Type 2 error, we can conclude that the modelling effort is successful and that the assumption of no interactive effects is valid. On the other hand if a test for autocorrelation is positive, we will be unable to resolve whether the source is interactive effects, or reactive effects not included in the model. So we can in general resolve the ambiguity between interaction and reaction only in those cases where perfect models can be obtained».

²⁹ L'Esempio è discusso nel capitolo 2.

tendono ad essere più dissimili di quelli distanti, rimane, secondo Daniel Griffith (2019), di gran lunga la situazione più trascurata in analisi spaziale. Esistono infine casi empirici in cui i processi spaziali non presentano autocorrelazione spaziale significativa – né positiva né negativa – e pertanto si qualificano come fenomeni casuali (*random*). Va comunque ricordato che le proprietà statistiche spaziali sono condizionate anche dall'aggregazione spaziale e dalla definizione delle unità campionate, per cui un processo spaziale che non presenta autocorrelazione spaziale significativa potrebbe anche non essere un processo casuale³⁰.

Poco sopra si è accennato alla definizione dei collegi elettorali in relazione al problema dell'aggregazione spaziale. Occorre però considerare che, sempre con riferimento all'aggregazione spaziale, lo strumento principale utilizzato dai geografi quantitativi è proprio l'autocorrelazione spaziale. Per comprenderne la portata, si pensi alla vittoria di Donald Trump alle elezioni politiche americane del 2016, che ha sorpreso molti³¹ perché la sconfitta, Hilary Clinton, in realtà aveva ottenuto più consensi del vincitore Trump: nel contesto del sistema elettorale americano, la geografia elettorale è un fattore influente nel determinare la differenza tra maggioranza di consensi e vittoria politica. Ciò che qui interessa evidenziare non sono metodi più o meno espliciti per manipolare i risultati elettorali a proprio favore, come il cosiddetto *gerrymandering*³², ma l'impiego dell'autocorrelazione spaziale come strumento per la definizione o ridefinizione dei distretti elettorali. Utilizzando tale indice, è possibile ri-disegnare distretti o collegi elettorali che siano di volta in volta omogenei oppure eterogenei. Ciò è possibile utilizzando appunto il grado di similitudine degli attributi delle unità situate in prossimità spaziale, in questo caso la similitudine delle opinioni politiche di elettori residenti in prossimità l'uno dell'altro, entro un perimetro ipotizzato.

³⁰ Si veda in proposito l'Esempio 3.1.

³¹ Si vedano, ad esempio, Stein *et Alii* (2020); Monnat e Brown (2017).

³² Il termine si riferisce al senatore del Massachusetts Elbridge Gerry, che nell'Ottocento propose una ri-definizione dei distretti elettorali di Boston al fine di favorire la propria forza politica. Rappresentata su una carta, la proposta di Gerry dava luogo ad una forma che ricordava la salamandra della mitologia, da cui il termine *gerrymandering*.

3.4. Le due facce della medaglia: autocorrelazione ed eterogeneità spaziale

I processi spaziali, come s'è visto nel capitolo 2, sono caratterizzati da proprietà intrinseche: la dipendenza (o autocorrelazione) spaziale e l'eterogeneità (o non-stazionarietà) spaziale, che sono alla base dei cosiddetti "effetti spaziali". Pur essendo l'elemento fondamentale e caratterizzante dei processi spaziali, tali proprietà hanno un impatto negativo sulle analisi statistiche, di cui tendono ad aumentare l'incertezza, tanto da rendere necessario l'impiego di metodi correttivi che ne mitigano le conseguenze. Nella loro apparente semplicità, tali proprietà potrebbero apparire addirittura banali, e per questo è importante considerare che le misure quantitative sono solo un fiavole riflesso delle proprietà qualitative che le sottendono.

L'autocorrelazione spaziale, come si è visto, è solo un'indicazione numerica della dipendenza spaziale, ossia una proprietà molto più complessa e difficile da definire. Tuttavia, una sua definizione è necessaria per la specificazione del modello autoregressivo spaziale (SAR)³³, che allevia l'effetto spaziale, ovvero l'inflazione della varianza indotta dall'autocorrelazione spaziale nelle stime del modello di regressione classico, che, in questo modo, ne diminuisce l'affidabilità. In termini meno formali di quelli usati nel capitolo 2, il problema dell'inaffidabilità delle stime può essere compreso intuitivamente: la dipendenza tra le osservazioni genera una certa ridondanza. Se infatti le osservazioni fossero indipendenti, ciascuna di esse fornirebbe un intero elemento di informazione, e l'informazione complessiva del modello sarebbe proporzionale al numero di osservazioni nel campione. Ma quando le osservazioni sono autocorrelate, l'informazione che ciascuna di esse fornisce coincide in parte con quella fornita dalle osservazioni che le sono vicine, cosicché l'informazione fornita complessivamente da tutte le osservazioni del campione è inferiore a quella che esse fornirebbero se fossero indipendenti: ciò comporta uno spreco, uno spandersi della poca informazione su un numero eccessivo

³³ Si veda il capitolo 2.

di osservazioni, che gonfia le code della distribuzione, aumentandone la varianza³⁴.

Il metodo utilizzato per rimuovere la ridondanza generata dall'autocorrelazione spaziale comporta la sua rimozione dal modello: un'operazione molto delicata che dovrebbe poggiare su uno studio del processo spaziale il più accurato possibile. Questo dovrebbe essere fatto sia in termini computazionali sia, e principalmente, in termini concettuali, in modo da costituire in se stesso un modello analitico del processo spaziale. Infatti, solo un modello accurato e fedele del processo spaziale (che genera l'autocorrelazione) può fornire un'altrettanto accurata quantificazione di tale autocorrelazione, che, introdotta a denominatore nel modello di regressione, possa efficacemente rimuoverne l'inefficienza da essa indotta. La specificazione del modello di autocorrelazione spaziale diviene quindi il passo cruciale nella realizzazione dell'analisi di regressione spaziale. Essa richiede la formalizzazione di un modello di tale autocorrelazione, che contenga una matrice dei pesi spaziali³⁵. Come si è visto poco sopra, tuttavia, l'indice di autocorrelazione spaziale non spiega il processo spaziale che sta alla base dell'autocorrelazione osservata, che, occorre ricordarlo, viene specificata solo con criteri quantitativi. Ed infatti, solo se si comprendessero la natura e le cause dell'autocorrelazione spaziale, si potrebbero definire un modello e una matrice dei pesi spaziali che rappresentino correttamente il processo spaziale. Se invece la matrice è definita solo da criteri numerici – volti a minimizzare la varianza delle stime –, il modello può anche produrre stime affidabili, che però potrebbero non rappresentare correttamente il processo spaziale.

Se si considera l'Esempio 2.1, il modello di autocorrelazione spaziale, definito sulla base dell'autocorrelazione osservata, utilizza una matrice dei pesi spaziali la cui soglia è definita dalla “vicinanza residenziale” dei cardiopatici. Se invece si analizzasse il processo in termini concettuali, identificando la radice dell'autocorrelazione spaziale

³⁴ Consideriamo l'esempio, banale, di un'aula scolastica tradizionale coi banchi disposti in file parallele a partire dalla cattedra. Uno studente seduto in prima fila fornisce l'informazione “apprezza la prima fila”. Ma anche lo studente che gli è seduto accanto fornisce la stessa informazione. Per questo, se lo studio considera un campione di 3 studenti seduti in prima fila, ben 3 unità forniscono la medesima informazione. Vi è quindi ridondanza delle unità spaziali campionate, ma anche ridondanza dell'informazione da esse fornita: si tratta di tre elementi che danno la stessa informazione, in altre parole un'unica informazione e 3 elementi.

³⁵ Si veda il capitolo 2.

nella “vicinanza residenziale” dei coetanei, la soglia della matrice dei pesi spaziali sarebbe diversa: non solo essa costituirebbe una rappresentazione migliore del processo spaziale, ma potrebbe anche portare a risultati quantitativi migliori, ossia ad un’ulteriore riduzione della varianza delle stime.

Analogamente alla dipendenza, o autocorrelazione spaziale, occorre anche considerare tutte le problematiche, anche di tipo concettuale, relative all’individuazione e al trattamento dell’eterogeneità spaziale che, intuitivamente, può essere definita come l’incostanza della variazione dei fenomeni nello spazio. Parafrasando Waldo Tobler, si può affermare che tutte le cose variano in maniera incostante nello spazio, e anche le relazioni tra le cose variano in maniera incostante nello spazio³⁶.

Intuitivo è anche il legame tra le due proprietà: se la variazione spaziale fosse costante non si avrebbe dipendenza spaziale, e se le cose vicine non fossero più simili di quelle distanti non ci sarebbe eterogeneità spaziale. Come osservato nel capitolo 2, un processo spaziale per potersi dire omogeneo (o stazionario) deve soddisfare alcune proprietà, tra cui la costanza della covarianza nello spazio³⁷. Si vede quindi come un elemento dell’eterogeneità sia l’incostanza della covarianza, e quindi della correlazione, ossia dell’autocorrelazione spaziale. Sono queste le ragioni per cui autocorrelazione ed eterogeneità spaziale spesso si presentano simultaneamente nei processi spaziali e per molti aspetti costituiscono le due facce di una stessa medaglia.

Nell’analisi multivariata l’eterogeneità spaziale presenta particolari problemi quando a variare in maniera incostante non è un singolo processo, ma lo sono due o più processi, e soprattutto allorché a variare in maniera incostante sono le relazioni tra i processi spaziali. Come già più volte osservato, l’eterogeneità ha le stesse conseguenze della dipendenza spaziale: l’inflazione di varianza, che aumenta l’incertezza dei risultati, diminuendo l’affidabilità delle stime dei modelli. La risposta metodologica a questo effetto spaziale è la *geographically weighted regression (GWR)*: ad oggi forse la metodologia analitica

³⁶ Si consideri anche l’Esempio 2.1.

³⁷ È importante osservare che la covarianza e la correlazione sono la medesima proprietà, a meno di un fattore di standardizzazione che consente di costruire un indice di correlazione racchiuso tra due estremi (-1 e +1), mentre la covarianza varia in funzione della grandezza della variabile.

spaziale più diffusa ed apprezzata³⁸. Prima di soffermarci sulla popolarità dell'analisi *GWR*, osserviamo che si tratta di una forma di analisi locale (*local analysis*). L'insieme di tecniche noto come analisi locale comprende una serie di metodologie, soprattutto di natura esplorativa e descrittiva, che furono sviluppate a partire dai primi anni '90 da Art Getis e Keith Ord (1992), che definirono alcuni indici locali di *clustering spaziale* e in seguito da Luc Anselin (1995) che definì una serie di indici locali di associazione spaziale³⁹.

L'analisi locale affronta il problema dell'instabilità spaziale delle relazioni stimando una miriade di indici o coefficienti locali, ciascuno dei quali corrisponde ad un piccolo intorno di ogni unità spaziale campionata: l'analisi è una sola, ma i coefficienti locali sono tanti quante le unità d'indagine (Anselin, 1995). Si consideri, ad esempio, un'analisi dei servizi sanitari in funzione della popolazione residente nel comune di Venezia, basata sui dati delle sezioni di censimento. Una regressione classica (globale) utilizzerebbe le 256 unità spaziali per produrre un coefficiente che lega le due variabili per l'intero comune. Se invece la stessa analisi fosse condotta usando la regressione *GWR*, si otterrebbero 256 coefficienti locali, ciascuno relativo alle due variabili in un intorno di ogni sezione di censimento.

Anche in questo caso il problema sarebbe numericamente risolto perché, riducendo l'estensione spaziale dell'analisi fino al paradosso della singola unità spaziale, si otterrebbe la stazionarietà dei fenomeni spaziali: ovviamente però al prezzo di un'estrema frammentazione dei risultati. Ne consegue che ad essere frammentati non sono solo i risultati, ma anche lo spazio, o territorio, cui essi si riferiscono. I risultati analitici sarebbero effettivamente affidabili, ossia non affetti da un eccesso di varianza, ma certamente non estrapolabili ad ambiti spaziali che si estendano oltre la singola unità. L'estrema frammentazione dello spazio operata dall'analisi quantitativa locale pone dunque anche delle questioni concettuali: qual è il senso di un'analisi che valga solo su atomi di territorio? È pur vero che i risultati dell'analisi locale si prestano a rappresentazioni cartografiche dettagliate, ma è anche vero che tali risultati possono sì essere mappati utilizzando tecniche di tipo

³⁸ Definita da Fotheringham, Charlton, Brunson (1998).

³⁹ In particolare, questo gruppo di tecniche comprende gli indici G_i e G_i^* di Getis e Ord e la versione locale dell'indice di autocorrelazione spaziale Moran's I .

GIS, ma tuttavia non possono essere efficacemente comunicati ai decisori, né tradotti in decisioni per il territorio, a causa della loro eccessiva frammentazione; tale comunicazione richiede quindi, comunque, una qualche forma di riagggregazione che magari tenga conto anche della dipendenza spaziale⁴⁰.

Ciascuna delle microregressioni *GWR* viene comunque calcolata su un intorno, per quanto piccolo, di un'osservazione, e richiede l'impiego di altre osservazioni⁴¹. La necessità di ridurre l'intorno di ogni osservazione per ridurre l'eterogeneità spaziale cozza con la necessità di avere un numero di osservazioni tale da garantire la robustezza del risultato statistico, e ciò conduce, nella pratica, alla definizione di intorni di dimensioni intermedie fra la singola osservazione e l'intera area di studio⁴². In tali intorni, come si è visto, è probabile che le osservazioni, che necessariamente si trovano a poca distanza l'una dall'altra, siano affette da autocorrelazione spaziale. Ciò conduce al rischio che almeno in alcune delle microregressioni *GWR* l'autocorrelazione spaziale sia anche più elevata che nell'intera area di studio, dove la distanza tra osservazioni è, complessivamente, maggiore, col risultato potenziale, già osservato, di esacerbare l'effetto dell'autocorrelazione spaziale nel tentativo di affrontarne l'eterogeneità. Il problema viene in qualche modo discusso da Stuart Fotheringham, ma rimane perlopiù nell'ombra, lasciando qualche insoddisfazione, in quanto l'autore stesso, pur argomentando che l'autocorrelazione

⁴⁰ Fotheringham (2009, p. 400): «Local statistics are calculated and local statistical models are calibrated by weighting observations according to their distance from a focal point such that the observations closer to the focal point are weighted more heavily in the calculation than are observations farther away. This procedure means that the statistic calculated is unique to a focal point, and that each time the location of the focal point changes, a new local statistic is calculated, generating a set of such values that can be displayed as a surface to show the extent of any spatial variation. In effect, this set of calculations creates a whole new geography—that of spatially varying relationships».

⁴¹ Quelle contenute all'interno del *kernel*, definito dalla *bandwidth*.

⁴² Come osservato nel capitolo 2, di fondamentale importanza è anche la selezione delle variabili esplicative del modello, che vengono generalmente selezionate a livello globale, senza verificarne la significatività nelle regressioni locali. Questo problema è stato affrontato di recente con la proposta di una regressione iperlocale (*Hyper-local GWR*), che ottimizza simultaneamente la dimensione degli intorni (*bandwidth*) selezionando le variabili esplicative localmente.

spaziale viene automaticamente risolta dal modello *GWR*, propone in termini generici un modello che integri le componenti *SAR* e *GWR*⁴³.

La regressione *GWR* è matematicamente molto meno complessa della regressione *SAR* e fin dall'inizio è stata propagandata come una tecnica che si presta a rappresentazioni visuali e cartografiche della regressione, particolarmente adatte ad applicazioni in ambiente GIS. Questa circostanza fa supporre che la relativa semplicità del metodo e la lusinga di una tecnica di tipo GIS e per sua natura visuale siano alla base del successo popolare di *GWR*⁴⁴. Questo potrebbe contribuire anche a spiegare la, relativamente scarsa, preoccupazione di Stuart Fotheringham e colleghi per gli aspetti del metodo che rimangono più problematici sotto i profili concettuale e computazionale.

A questo proposito potrei io stessa riferire non pochi aneddoti sulla tendenza manifestata da numerosi colleghi e studenti ad utilizzare *GWR* in maniera non sempre appropriata: generalmente si tratta di individui più entusiasti del GIS e delle tecnologie dell'informazione geografica che ferrati nel rigore dell'analisi spaziale. Viceversa, potrei citare opinioni scambiate informalmente da esperti che manifestano sostanziale scetticismo sull'effettiva capacità di *GWR* di risolvere il problema dell'autocorrelazione spaziale. Tra le tendenze più comunemente osservate vi è quella ad applicare *GWR* a dati affetti da autocorrelazione spaziale, senza chiarire (o comprendere) che il metodo non è necessariamente capace di risolvere il problema. Vi è poi la tendenza a calcolare *GWR* per poi applicarne i coefficienti (quali?) come strumento d'interpolazione spaziale; a parte la necessità di scegliere tra una miriade di coefficienti spaziali quelli da utilizzare, questa tendenza

⁴³ Fotheringham (2009, pp. 400-402): «Therefore, the calculation of local statistics and the calibration of local statistical models depend on a spatial weighting function by which observations at locations nearer to a focal point are given more weight than are observations farther away. That is, the construction of local statistics depends on spatial dependency — the property that observations closer to each other in geographic space are more likely to be similar than are observations farther apart. [...] Consequently, calibrating local models removes this problem of spatial dependency in residuals, as demonstrated by Fotheringham, Charlton, and Brunson (2002). [...] This end can be achieved, for example, in a geographically weighted version of a spatial regression model. For instance, a geographically weighted spatial lag model yields local measures of spatial autocorrelation through the estimated parameters on the spatially lagged y variable. Therefore, such models account for local spatial dependence as well as provide information about the degree of spatial nonstationarity in the processes being modelled».

⁴⁴ Comber *et Alii* (2018); Fotheringham (2009).

denota una contraddizione di fondo, poiché coefficienti determinati ad hoc per atomi di spazio vengono poi applicati all'intera area di studio – non si comprende quindi perché si sia condotta un'analisi localmente dettagliata per poi estenderne i risultati (quali?) all'intera area di studio⁴⁵. Vi è infine forse la più preoccupante delle tendenze, quella a svolgere un'intera analisi *GWR* in un pacchetto GIS, ciò che può avvenire senza che l'utente abbia condotto alcuna analisi statistica (anche solo sulla distribuzione o sulle proprietà dei dati) e che poi ne presenti i risultati in forma di mappa (che spesso è anche un buon prodotto cartografico) ma che è stata prodotta senza che l'utente abbia anche solo visto i coefficienti del modello. Soprattutto quest'ultima tendenza, su cui si tornerà nel seguito, evidenzia una nota e pesante tensione tra il bagaglio di conoscenze richiesto per condurre propriamente analisi spaziali e la superficiale semplicità con cui esse vengono proposte dai pacchetti GIS.

3.5. Lo spazio dell'analisi spaziale

Se un processo spaziale è una struttura matematica che concettualizza un fenomeno geografico, e la matrice dei pesi spaziali è un altro costruito matematico che definisce le relazioni tra unità spaziali, non ci si può nascondere che tutte queste astrazioni poggiano su una rappresentazione dello spazio. Tale rappresentazione implica comunque il passaggio da uno spazio empirico, popolato di esseri, oggetti ed eventi, ad uno spazio matematico, popolato di primitive geometriche – punti, linee e poligoni – correate di numeri che ne rappresentano gli attributi⁴⁶. La geografia teoretico-quantitativa ha tradizionalmente utilizzato una rappresentazione dello spazio di tipo assoluto, o newtoniano, concettualizzato come un contenitore neutro ed esterno ad esseri, oggetti ed eventi. Questa rappresentazione consente di contrassegnare oggetti ed eventi con coordinate cartesiane e di rappresentarli

⁴⁵ Non è impossibile utilizzare *GWR* a questo scopo, come ad esempio in Bertazzon *et Alii* (2015), in cui, tuttavia, si utilizzarono i diversi coefficienti locali per interpolazioni localizzate ed integrate in un modello *SAR*.

⁴⁶ In ogni caso questa trasformazione è mediata da uno dei cosiddetti 'modelli dei dati' (*data models*) di cui i principali sono il modello *raster* e quello vettoriale – questi aspetti saranno ripresi nel capitolo 5.

secondo la geometria euclidea. Al polo opposto della rappresentazione vi è lo spazio relativo, leibniziano, inteso come uno spazio che coesiste con gli oggetti e gli eventi, che da essi non è separabile, e che non può esistere senza di essi⁴⁷. Senza addentrarci in questioni filosofiche che trascendono lo scopo di questo libro, è importante chiederci se la rappresentazione tradizionale dello spazio che ha sorretto tanta geografia quantitativa sia la più adatta a sorreggere un'analisi spaziale relativizzata e frazionata, in cui le relazioni osservate – autocorrelazione spaziale – possono essere mediate da processi diversi da quelli osservati – processi interattivi. Come notava già Tony Gatrell, il concetto di spazio relativo è più generale ed empiricamente più utile del concetto di spazio assoluto e questo per due ragioni: la prima è che riferendosi ad una 'relazione' si prefigurano molti altri modi in cui sia possibile descrivere la separazione spaziale tra oggetti; la seconda è che non è necessario che la relazione sia un qualcosa che possiede le proprietà metriche della distanza (le cosiddette proprietà 'metriche')⁴⁸.

Di fondamentale importanza è la rappresentazione e, indirettamente, la misurazione della distanza, ossia la separazione tra due o più unità spaziali (più frequentemente punti o poligoni): rappresentazione e misurazione che sottendono anche le relazioni di vicinanza, prossimità, contiguità, codificate nella matrice dei pesi spaziali. Molti sono i limiti della rappresentazione/misurazione della distanza in questo tipo di configurazioni matematico-geometriche, anche ammesso che lo spazio goda delle proprietà della geometria euclidea, e che la distanza goda delle proprietà metriche che in esso sussistono. Basti pensare che la matrice dei pesi spaziali è generalmente simmetrica: ma chiunque abbia compiuto anche una semplice escursione in montagna avrà notato la differenza tra l'ascesa e la discesa, benché la lunghezza del percorso sia la medesima. Basti, a questo proposito, pensare alla potenziale considerazione di aspetti qualitativi della distanza, come la

⁴⁷ Si vedano in proposito i lavori di Gatrell (1983); Couclelis (1999); le posizioni più recenti di Goodchild (2018) e critiche di O'Sullivan, Bergmann, Thatcher (2018).

⁴⁸ Si veda Gatrell (1983, p. 3): «This leads me to elaborate the concept of *relative space*, which sees space as a relation defined on a set of objects. I emphasize this, since it is the concept of space illustrated in this book. As I hope subsequently to demonstrate, it is more general and empirically more useful than the concept of absolute space, for two reasons; first, by referring to a 'relation' we are envisaging many ways in which the spatial separation of objects can be described. Second, we do not require the relation to be something with the properties of distance (so called 'metric' properties)».

già notata percezione della distanza delle strutture sanitarie da parte delle persone più bisognose di quelle strutture – anziani e malati. Considerazioni analoghe valgono per la definizione di barriere culturali o sociali, che possono rendere quelle strutture sanitarie di fatto inaccessibili a porzioni della popolazione⁴⁹, o ancora di barriere fisiche, come corsi d'acqua in assenza di ponti, edifici di grandi dimensioni o aree recintate. Analogamente, fiumi, montagne e precipizi possono costituire barriere fisiche insuperabili per gli animali, ed assumono grande importanza in analisi ambientali ed ecologiche⁵⁰.

Aspetto ancora più delicato della rappresentazione dello spazio nel modello quantitativo è il posizionamento delle persone. Innanzitutto, vi è la delicata questione dell'etica e del rispetto della privacy, che tradizionalmente viene risolta con forme di aggregazione spaziale: il che ci riporta alla discussione sull'aggregazione spaziale e il *MAUP*. Parte del problema è la presunta fissità delle persone nello spazio, quando in realtà esse si muovono e i loro movimenti potrebbero essere particolarmente rilevanti per l'analisi. Si pensi alle analisi di geografia sanitaria e ambientale: molta di questa letteratura si occupa dell'esposizione a fattori inquinanti. Tipicamente si assume che le persone non si muovano dal proprio luogo di residenza, quando in realtà l'esposizione può non avvenire nel luogo di residenza – e neppure nel luogo di lavoro o a scuola –, ma nel tragitto che li separa⁵¹. Ancora più paradossale, ad esempio, l'analisi delle malattie zoonotiche causate dalle punture di zecca: gli studi più recenti impiegano analisi spaziali, utilizzando le coordinate geografiche di ogni morso⁵². Quali coordinate geografiche? Quelle della residenza della persona morsa? Dell'ambulatorio in cui è stata visitata? O ancora del luogo in cui verosimilmente è stata morsa? Di particolare rilievo e tragica attualità nella geografia sanitaria sono gli studi sul contagio e la diffusione delle malattie,

⁴⁹ In questo caso si può pensare a come una donna proveniente da certe culture tradizionali possa non essere in grado di coprire la distanza che la separa dal negozio o dal medico perché non le è permesso viaggiare da sola e non c'è un familiare maschio che la possa accompagnare durante l'orario di apertura.

⁵⁰ O, ancora, nelle analisi ambientali ed ecologiche, le difficoltà relative allo spostamento di animali ed insetti connesse a barriere naturali o artificiali come autostrade, superstrade o staccionate.

⁵¹ Si vedano ad esempio Bertazzon *et Alii* (2020) e Bernatsky *et Alii* (2015).

⁵² Namgyal *et Alii* (2021).

stimolati dal dilagare del nuovo coronavirus⁵³. Numerosi studi ambiscono a rappresentare i contatti interpersonali, tramite i quali verosimilmente avviene la maggior parte dei contagi. Al fine di rappresentare questi contatti, spesso utilizzando dati ad elevatissima risoluzione spaziale⁵⁴, si definiscono dei *network* di relazioni, anche di tipo probabilistico, innestati di norma in modelli di diffusione ed interazione spaziale oppure in modelli di reti neurali e intelligenza artificiale.

Per tutte queste situazioni appare più appropriata una rappresentazione dello spazio di tipo relazionale e probabilistico, riferita non tanto alla residenza, ma al probabile raggio di spostamento delle persone, definito a sua volta dalle relazioni sociali, i movimenti necessari (scuola, lavoro, spesa), e le mete di ricreazione o svago. Tutto questo prefigura uno spazio fatto di *luoghi* che sono ben più di mere porzioni di spazio definite da una coppia di coordinate geografiche, che interagiscono con le relazioni, che a loro volta avvengono su scale diverse in virtù di processi spaziali interattivi. Tutto ciò richiede una rappresentazione più complessa, che non si esaurisce nei numeri e nelle primitive geometriche, ma richiede l'apporto di modelli e schemi diversi da quelli espressi dalla mera geografia quantitativa. Anche la transizione, operativa, da un rigido modello di spazio assoluto ad un più ricco spazio relativo, flessibile e definito dalla questione di ricerca, costituisce un passaggio importante della transizione dalla classica geografia quantitativa verso la moderna analisi spaziale⁵⁵.

3.6. Dal relativismo la nuova analisi spaziale

L'analisi spaziale che emerge dai paragrafi precedenti è quindi molto più che un approccio meramente quantitativo, quale erano i metodi tradizionali usciti dalla vecchia rivoluzione teoretico-quantitativa. Innanzitutto, accettando la pluralità di risultati derivante dal problema dell'aggregazione spaziale (*modifiable areal unit problem – MAUP*) di Stan Openshaw (1977), l'analisi quantitativa diviene relativa alle sole unità territoriali su cui è calcolata. È poi un'analisi che non può

⁵³ Do Lee, Qian, Schwanen (2021); Sun, Chen, Vibout (2020); Murgante *et Alii* (2020).

⁵⁴ Spesso tratti dai movimenti telefonici, anche su questi aspetti si tornerà nel capitolo 5.

⁵⁵ Per alcuni sviluppi recenti si veda Bergmann e O'Sullivan (2018).

più accontentarsi di soli metodi matematici (deterministici), ma bisogna di metodi statistici (stocastici) per dar conto dell'incertezza. Incertezza che in parte è insita nei fenomeni spaziali, in parte è indotta dai limiti della loro misurazione, e in parte dalle limitazioni dei metodi deterministici, tuttora inadeguati a comprendere appieno tali fenomeni⁵⁶. I metodi statistici sembrano fornire una soluzione, grazie alla loro capacità di incorporare l'incertezza e dar conto dell'errore; tuttavia, la loro applicazione ai dati spaziali espone l'analisi all'inflazione di varianza indotta dagli effetti spaziali, dipendenza e eterogeneità spaziali, che ne riduce l'affidabilità. Così la dipendenza spaziale è risolta da un modello di relazioni territoriali definito sulla vicinanza delle unità spaziali, ove quello di vicinanza è un concetto quanto mai relativo e legato al contesto. Ed infine l'analisi statistica locale frammenta lo spazio già relativizzato, paradossalmente limitando la validità dell'analisi quantitativa ad un solo atomo di territorio. Non solo, ma i medesimi strumenti analitici sono stati via via estesi all'analisi di dati spazio-temporali⁵⁷, portando così ad una relativizzazione non solo dello spazio, ma anche dello spazio-tempo.

La centralità dei metodi quantitativi non ha impedito l'apertura a metodi non quantitativi, come dimostra la letteratura recente e crescente sull'uso dei metodi misti⁵⁸, che si propongono di affrontare una medesima questione di ricerca da angolazioni diverse, mediante l'utilizzo di una pluralità di metodi, quantitativi e qualitativi. La vocazione della geografia ad integrare lo studio dell'ambiente e del sociale, che ben si esprime nella geografia quantitativa e nel GIS, diviene un presupposto capace di far dialogare questa geografia sia al proprio interno che con altre discipline⁵⁹. Riprendendo l'esempio già usato più sopra, si consideri l'accessibilità dei servizi sanitari per la popolazione residente in una qualsiasi provincia italiana. Qui la risposta della geografia sanitaria classica (quantitativa e logico-deduttiva) è un'analisi basata sulla distanza tra i servizi localizzati e la residenza dei potenziali utenti. La risposta di una metodologia d'indagine mista affiancherebbe all'analisi della distanza un'indagine qualitativa condotta mediante

⁵⁶ Bauer e Rose (2015); Chen *et Alii* (2013).

⁵⁷ Si vedano: An *et Alii* (2015); Fotheringham, Crespo; Yao (2015) e Lin *et Alii* (2015).

⁵⁸ Alexander, Alfonso, Hansen (2015), Lafreniere e Gilliland (2015), Pocewicz *et Alii* (2012) e Preston e Wilson (2014).

⁵⁹ Goodchild (2013) e Kalbasi *et Alii* (2014).

interviste agli utenti e volta ad esaminare la percezione della distanza e dei tempi di percorrenza, dei tempi d'attesa per i servizi sanitari, o della disponibilità mezzi di trasporto pubblici. Non priva di difficoltà, l'applicazione di metodi misti è anzi capace di aprire questioni nuove, sia di natura applicativa che teorica. Infatti, metodi e questioni di ricerca non sono indipendenti l'uno dall'altro, e quindi l'uso di metodi misti modifica e moltiplica le questioni tipicamente affrontate con i soli metodi quantitativi. Vi sono poi difficoltà insite nei dati, poiché, mentre l'analisi quantitativa richiede scarni dati numerici su campioni vasti e rappresentativi, la ricerca qualitativa richiede dettaglio e ricchezza di informazioni, che si possono ottenere solo per campioni relativamente piccoli. Notevoli, in questo senso, sono il contributo di Dan Griffith (2013) volto a determinare la dimensione del campione per analisi *qualitative* in presenza di autocorrelazione spaziale, avvalendosi di tecniche sviluppate in contesto *quantitativo*, e quello di Richard Medina e George Hepner (2011) volto a valutare le dipendenze *socio-spaziali* nei network di attività terroristiche. Benché l'applicazione di metodi misti sia ancora agli inizi, essa costituisce un chiaro segnale di apertura ad affrontare le questioni con metodi che trascendono il puro quantitativismo.

Parte II

Definizione, identità, autocritica

4. Verso una definizione di analisi spaziale

To have command of definition is to have control of discourse.
Livingstone, *The Geographical Tradition*, 1992, p. 304.

4.1. Premessa

In apertura di questo volume abbiamo fornito una definizione operativa di analisi spaziale, ossia quella di un «processo che genera conoscenza utile per rispondere a domande sull'ordine, la disposizione e la struttura inerenti ai problemi spaziali». Essa costituirebbe «un dominio unico, che esiste all'incrocio di molte discipline e prospettive metodologiche diverse». Nei due capitoli successivi ne abbiamo poi analizzato i contenuti, prima in termini prettamente quantitativi, e poi cercando di comprenderne meglio il senso, esaminandoli anche da un punto di vista più qualitativo e concettuale. Su queste basi cerchiamo ora di approfondire la definizione di analisi spaziale, cercando soprattutto di capire cos'è e cosa non è, ossia in che modo possa essere caratterizzata, distinguendola da altre parti della geografia.

Ne esamineremo il rapporto con le discipline da cui in qualche misura discende, la geografia quantitativa e il GIS, quest'ultimo inteso nell'accezione di scienza dell'informazione geografica. Per arrivare a comprenderne le sovrapposizioni e le differenze, ma anche per cercare di intuire quali potrebbero essere gli sviluppi futuri dell'analisi spaziale, ci soffermeremo su alcuni aspetti della vicenda del GIS. In particolare, ci soffermeremo sulla posizione di quest'ultimo rispetto alla scienza anche in rapporto alla sua storia di *big business*, una caratteristica che nel corso della sua vicenda ne ha ostacolato, più che favorito, l'evoluzione verso una scienza. La dialettica tra scienza e *business* è molto importante al momento attuale: momento in cui l'esplosione dei *big data* sembra attirare sempre più l'attenzione sulle potenzialità

dell'analisi spaziale con la sua capacità nello studiare i dati spaziali. Situazione questa che potrebbe farle rivivere una vicenda simile a quella del GIS, anche se, è auspicabile, evitando di ripeterne, col senno di poi, gli errori.

Vi è poi un altro aspetto dell'analisi spaziale che sarà affrontato in questo capitolo: la sua potenzialità non tanto come *big business*, ma come strumento per la comprensione dei problemi aventi connotazione spaziale. Si tratta di riflessioni che condurranno all'apertura della questione sull'appetibilità dell'analisi spaziale, che verrà ripresa nei capitoli successivi.

La rinnovata attenzione verso l'analisi spaziale sembra, inoltre, essere accompagnata da una fioritura di nomi diversi, che vanno dal ritorno alla *scienza spaziale* fino al neologismo *spatial analytics*, passando per altri neologismi e ambiguità terminologiche. Arriviamo a contare 15 nomi diversi, di molti dei quali (se non di ciascuno) cerchiamo di comprendere il significato e ove possibile l'origine e l'accezione comune. Così, se tradizionalmente (fino a circa un decennio fa) era possibile situare l'analisi spaziale all'intersezione tra geografia quantitativa, GIS e statistica, oggi essa si trova all'intersezione di numerose discipline e "comunità di pratica" situate tra l'analisi geografica, la statistica e l'informatica. Alcune di esse portano il nome di scienza, altre sono centrate sui dati, altre ancora tendono a trascinarla verso l'ingegneria e l'informatica.

Certo siamo in un momento di grande fermento, stimolato e sollecitato dall'esplosione dei *big data* e da sviluppi tecnologici senza precedenti (come *supercomputing*, *parallel computing* e *quantum computing*). Ma siamo anche in un momento in cui si comincia a riflettere sul senso dell'analisi dei dati spaziali che va oltre i dati e oltre la tecnologia, ponendosi domande importanti sulla significatività, il valore e l'impatto di queste analisi.

4.2. Che cos'è la [moderna] geografia quantitativa?

Al volgere del secolo Stuart Fotheringham e colleghi definivano la geografia quantitativa come «una o più delle seguenti attività: analisi di dati spaziali numerici; sviluppo di teoria spaziale; costruzione e verifica (*testing*) di modelli matematici di processi spaziali. Lo scopo di

tutte queste attività è di aumentare la nostra comprensione dei processi spaziali»¹. Accettando questa definizione, Alan Murray vi aggiunge la creazione di informazione e conoscenza spaziale, e si chiede se i metodi sviluppati in altre discipline appartengano all'analisi spaziale. Per questo propone la seguente definizione: «Così si potrebbe dire che la geografia quantitativa è l'insieme dei metodi che sono applicati, o potrebbero/possono essere applicati, dai geografi e da altri per studiare fenomeni, questioni (*issues*) e problemi spaziali», precisando ulteriormente che la propria definizione «riconosce che un aspetto di primaria importanza della geografia quantitativa è lo sviluppo di una qualche forma di modello, che generalmente ha un contesto o una componente spaziale significativi»². La geografia quantitativa sarebbe dunque costituita da un insieme di metodi che, sempre secondo l'autore, possono essere classificati entro le seguenti ampie categorie: sistemi d'informazione geografica; rilevamento aereo (sistema di posizionamento globale, fotogrammetria e telerilevamento); statistica ed analisi esplorativa dei dati spaziali (ESDA – *exploratory spatial data analysis*); matematica e ricerca operativa [*optimization*]; analisi regionale; informatica e simulazione computerizzata³.

Nel delineare il proprio campo di applicazione, rivendicandone i progressi compiuti, la disciplina non si è sottratta ad una riflessione critica sui propri metodi e sui limiti della loro applicazione. A fronte dei progressi compiuti nell'affinare tali metodi e risolverne i principali

¹ Fotheringham, Brunson, Charton (2000, p. 4): «Quantitative geography consists of one or more of the following activities: analysis of numerical spatial data; the development of spatial theory; and construction and testing of mathematical models of spatial processes. The goal of all these activities is to add to our understanding of spatial processes».

² Murray (2010, p. 144): «This [Fotheringham et al (2000)'s] is a reasonable definition/characterization, though the creation of spatial information and knowledge may be missing. Further, there is some ambiguity about whether methods developed by other disciplines (mathematics, statistics, engineering, social sciences, etc.) for aspatial contexts would qualify as quantitative geography methods. [...] Thus, one might say that quantitative geography is the collection of methods that are applied, or could/can be applied, by geographers and others to study spatial phenomena, issues and problems. [...] What this overview will do is recognize that a major facet of quantitative geography is development of some sort of model, generally having a significant spatial context or component».

³ Murray (2010, p. 145): «The following broad categories are used to detail methods in quantitative geography: geographic information systems (GISs); airborne sensing (global positioning system, photogrammetry, and remote sensing); statistics and exploratory spatial data analysis (ESDA); mathematics and optimization; regional analysis; and computer science and simulation».

problemi, Alan Murray (2010)⁴ espone un elenco dei principali problemi tuttora irrisolti: incertezza dei dati spaziali, i problemi legati alla modellizzazione, ovvero astrazione, scala e unità di misura, indipendenza dalla cornice di riferimento, infine dinamiche spazio-temporali, ed integrazione tra categorie. L'incertezza è in parte una caratteristica intrinseca dei fenomeni spaziali, e quindi anche dei dati spaziali (si pensi alla mobilità degli individui rispetto al luogo di residenza, all'arbitrarietà delle unità spaziali di riferimento, o alla mobilità di fenomeni fisici come il vento o le correnti). I miglioramenti tecnici dei metodi di rilevazione possono ridurre in parte l'incertezza, ma rimane la necessità di rappresentare ed analizzare propriamente l'incertezza insita nei fenomeni, come quella indotta dall'imperfezione dei metodi di analisi quantitativa⁵. A ciò si aggiunge il problema della

⁴ Murray (2010, pp. 158-160): «An often discussed issue is spatial data uncertainty. For quantitative geography this is no doubt important and significant. [...] A model, whether data, statistical, or mathematical in orientation, is an abstraction of reality. As discussed previously, a model reflects intended (and unintended) relationships, and one's knowledge of relationships. If nothing else, we are now keenly aware that our models of reality have much potential to influence or bias analysis. [...] Tobler (1989) highlights that model specification, or rather incorrect model specification, is a likely source of error, resulting in an inability to statistically confirm certain spatial relationships. This is referred to as frame dependence, or rather that a particular statistical test is dependent on the spatial frame utilized and a different frame may produce differing results. [...] Perhaps not so obvious in the review of quantitative geography methods is work focusing on spatio-temporal patterns. It most certainly is implied, but admittedly not discussed in any detail. [...] An approach is developed to solve the model using GIS-based methods, taking advantage of spatial relationships and knowledge, thereby reflecting an intelligent approach to searching space. Although there is ample evidence of inter-category integration already occurring, much remains possible».

⁵ Anselin (2002, pp. 247-256, corsivi dell'autore): «While undoubtedly considerable progress has been made, most applications of spatial econometrics are rather limited in the way in which *spatial interaction* is incorporated in the model specifications. The typical approach is to distinguish between so-called *spatial lag* and *spatial error* models. [...] What is not always well understood in this process is that different spatial models induce sometimes radically different spatial correlation patterns, which do not necessarily match the underlying theoretical interaction model. [...] In practice, the motivation for applying a spatial econometric model is typically not driven by formal theoretical concerns, but instead is a result of data "problems". For example, the scale and location of the process under study does not necessarily match the available data, such as when agricultural land markets are studied with data at the county level. This mismatch will tend to result in model error structures that show a systematic spatial pattern. Also, explanatory variables are often "constructed" by spatial interpolation to make their scale compatible with that of the dependent variable. Again, this spatial prediction will tend to result in prediction "errors" that show systematic spatial variation. This problem is commonly encountered in models where economic outcomes are related to environmental or resource variables, such as air or water quality. Spatially aggregate measures of the latter are computed by interpolating measures obtained for a small set of

rappresentazione, che implica l'astrazione e l'indipendenza dal contesto di riferimento. Significativi progressi sono stati compiuti nella rappresentazione di fenomeni spazio-temporali dinamici⁶, soprattutto con i progressi compiuti dai GIS, ma le difficoltà sussistono nell'analisi di tali dinamiche, complicate da autocorrelazioni e eterogeneità che si compenetrano nello spazio e nel tempo. Infine, se la definizione di categorie può semplificare l'analisi degli attributi delle entità spaziali, rimane la necessità di integrare tali categorie sia in termini analitici che concettuali.

Vi è poi un altro aspetto, lamentato già da Stuart Fotheringham e colleghi: proprio nei momenti in cui esisterebbe una domanda di geografia quantitativa, stimolata anche dalla disponibilità di dati e mezzi tecnologici, la sua applicazione appare in qualche modo ostacolata da alcune circostanze. Da un lato si tratterebbe delle tensioni epistemiche all'interno della disciplina, che si possono riassumere in un atteggiamento negativo da parte di certa geografia nei confronti di quel positivismo che aveva caratterizzato le origini della geografia quantitativa. Dall'altro ci sarebbe una percezione, talvolta mascherata e non espressa apertamente, che l'analisi dei dati spaziali sia una materia "difficile", in quanto richiede una certa familiarità con strumenti matematici e statistici più che elementari, con quali non tutti i geografi si sentono a proprio agio⁷. La commistione di questi due elementi

monitoring stations, whose locations do not coincide with those of the economic agents. Another often encountered situation is when data on important variables are missing, and those variables show spatial structure, as is often the case in studies of tropical land use and deforestation. A common characteristic of these data problems is that the error term in a regression model will tend to be spatially correlated. [...] A fundamental problem in the analysis of spatial correlation in a pure cross-sectional setting is the lack of identification of the parameters of the complete covariance matrix».

⁶ Si veda il lavoro di Peuquet (1994).

⁷ Fotheringham, Brunson, Charlton (2000, p. XI): «One of the more puzzling paradoxes that will face those who come to review the development of geography will be why, at the end of the twentieth century, much of geography turned its back on quantitative spatial data analysis just as many other disciplines came to recognize its importance. At a time when geography should have been meeting the rapidly growing demand for spatial data analysts, the majority of its graduates were, at best, non-quantitative and, in quite a few cases, were actively anti-quantitative. A commonly expressed reason for the negative attitude by many geographers towards one of the discipline's basic elements is a disillusionment on their part with the positivist philosophical underpinnings of much of the early work in quantitative geography. Another, less frequently stated reason, is that spatial data analysis and spatial modelling are perceived to be relatively difficult, not only by students, but also by many academic geographers who typically have non-quantitative backgrounds. Unfortunately, this perception has

porterebbe a drastiche posizioni di rifiuto dei metodi quantitativi in geografia che ne ostacolano uno sviluppo sereno ed un'affermazione nell'ambito della disciplina.

Possiamo però dire che la geografia quantitativa non è più quella disciplina arida e rigidamente legata alla visione positivista degli anni '70 del secolo scorso, ma è piuttosto una disciplina vivace e caratterizzata non solo da progressi tecnici, ma anche da un dinamico dibattito intellettuale⁸. Quanto alla difficoltà percepita rispetto all'uso di strumenti quantitativi, la proposta di Stuart Fotheringham e colleghi sembra quella di evitare le preclusioni derivanti da atteggiamenti aprioristicamente negativi, assieme a quella di avvicinare la geografia quantitativa passando per l'uscio del GIS, come sempre più accattivante ed avvincente, grazie alle sue capacità di visualizzazione⁹.

deterred many researchers from appreciating the nature of the debates which have emerged and which will continue to emerge within modern quantitative geography. This becomes clear in continuing criticisms of quantitative geography which pertain to methodologies that have been surpassed by developments within the field». L'evoluzione dell'analisi spaziale e della geografia quantitativa dal positivismo originale sono state discusse anche altrove in questo volume.

⁸ Kwan e Schwanen (2009, p. XI): «We highlight some promising developments in modern quantitative geography and reflect on the ways in which the critical-quantitative binary can be at least partially eclipsed. We emphasize that knowledge in quantitative methods is essential for deciphering and challenging regressive political agendas, now often supported by numbers and quantitative analysis. Quantitative geography, when integrated with a critical sensibility and used appropriately, can be a powerful tool for fostering progressive social and political change».

⁹ Fotheringham, Brunson, Charlton (2000, pp. 3-4): «A second reason for writing the book is that we hope to demonstrate that because of the changes taking place and that have taken place within the subject, several of the well-oiled criticisms traditionally levelled at quantitative geography no longer apply. For instance, the overly simplistic depictions of many that quantitative geographers search for global laws, and that individuals' actions can be modelled without understanding their cognitive and behavioural processes, have rather limited applicability. For those who insist on 'pigeon-holing' everything, modern quantitative geography, with its emphasis on issues such as local relationships, exploratory analysis and individuals' spatial cognitive processes, must be a difficult area to classify. A third reason is the hope that some of the changes taking place in quantitative geography might make it more appealing to students and by advertising the existence of these developments, we might foster a greater interest in and appreciation for what modern quantitative geography has to offer. This is particularly the case, for example, in the subsequent discussions on topics such as visualization, exploratory data analysis, local forms of analysis, experimental significance testing and GIS-based forms of spatial analysis».

4.3. Che cos'è [e che cosa non è] il GIS?

Anche il GIS, come la geografia quantitativa, è cresciuto, si è evoluto, è mutato nel corso del suo mezzo secolo di storia. Ma lo sviluppo storico del primo è stato molto diverso da quello della seconda. La prima nacque, lo si è visto, col proposito di trasformare la geografia in una scienza, come geografia teoretico-quantitativa, e di lì crebbe, con alterne vicende, ma sempre come una specializzazione a pieno titolo della geografia. Al GIS invece mancò, sin dall'inizio, un'affiliazione netta con la geografia e, forse, persino “una casa e una famiglia” in cui riconoscersi: nato come strumento, solo in seguito, e non senza fatica, si conquistò il pedigree di scienza. Infine, non si deve dimenticare che la sua capacità di generare profitto e di essere, esso stesso, un *big business* ha da sempre giocato un ruolo fondamentale nel suo sviluppo.

La realizzazione del GIS negli anni '60 fu resa possibile da una serie di sviluppi in varie discipline che si accumularono ed integrarono in una proficua convergenza di idee stimolate da una molteplicità d'interessi e di applicazioni. Se, da un lato, tale multidisciplinarietà pose le basi per la ricchezza della disciplina, dall'altro essa fu l'elemento chiave che generò la progressiva marginalizzazione del geografo e della geografia nel GIS. Se, infatti, l'ampia gamma di partecipanti favorì lo sviluppo di una molteplicità di applicazioni, che stimolarono lo sviluppo di sistemi sempre più raffinati ed efficienti, essa contribuì anche alla frammentazione dell'oggetto e della comunità del GIS, il cui nucleo centrale si andò ramificando, stirando e slabbrando in risposta ad esigenze pratiche anche molto disparate. Inoltre, l'enfasi sugli aspetti applicativi e tecnologici del GIS pose i presupposti perché esso fosse percepito da più parti come una metodologia ancillare, orientandone quindi lo sviluppo nella direzione di “sistemi per la soluzione di problemi”, anziché in quella di una branca della geografia, che sarebbe stata forse più lenta nel produrre soluzioni tecniche, ma certo più preparata a riflettere sui propri presupposti, oggetti e metodi, così da maturare come disciplina scientifica¹⁰. L'enfasi sulle

¹⁰ Pickles (1997, p. 368): «success in developing new applications meant that the center of gravity in GIS shifted rapidly in the late 1980s from universities to private corporate developers with residual ties (if any) into universities through original algorithm developers. A new institutional configuration emerged to dominate both the development and the use of GIS as a tool, and in these corporate settings, the role of academic geographers was uncertain.

applicazioni e il desiderio di fornire risposte immediate al committente o finanziatore di turno contribuirono anche a rallentare la realizzazione dell'analisi spaziale: infatti, se i geografi accademici avessero potuto svolgere il proprio lavoro di ricerca senza pressioni (da committenti o finanziatori) si sarebbero probabilmente applicati all'analisi spaziale dedicandole tutto il tempo e la cura che essa richiedeva¹¹. Viceversa, l'importanza di questo lavoro sfuggiva, nei decenni d'oro del GIS, a quanti in esso cercavano soprattutto soluzioni operative veloci, efficienti e, se possibile, redditizie.

Questa situazione era destinata a produrre poderosi sviluppi tecnici che sarebbero diventati formidabili punti di forza dei *sistemi* d'informazione geografica, ma avrebbero allo stesso tempo costituito una delle maggiori debolezze di quella che sarebbe poi stata definita *scienza* dell'informazione geografica. L'elevato valore commerciale e gli interessi economici in gioco contribuirono allo sviluppo di sistemi sempre più raffinati ma proporzionalmente svincolati dalla ricerca accademica. L'appetibilità del GIS per molti settori dell'industria provocò anche un drenaggio di risorse dall'accademia¹² all'industria, con conseguenti sviluppi in direzioni applicative eterodirette, a discapito della ricerca accademica indipendente¹³. Se la produzione del GIS

Moreover, the broadening of the range of applications of GIS beyond geography to other research and applied fields fostered a gradual loss of the geography in GIS usage, or more accurately, a reduction of what the geographical entailed within GIS (now reduced to locational attribute, terrain feature, and landscape content). It was no longer the case that everyone in the field of GIS looked back to common origins, such as the Harvard Graphics Lab, the work of Tomlinson in Canada, and the Minnesota Land Management Information System».

¹¹ Si vedano, come approfondiremo in seguito, i lavori di Openshaw (1991; 1992; 1994).

¹² Seguendo la tradizione anglofona, usiamo i termini "accademia" e "accademico" come sinonimi di "università" e "universitario".

¹³ Pickles (1997, p. 367): «In the 1990s, GIS has been generalized in ways in which even strong advocates in the 1980s could not dream. It has been generalized throughout the geography curriculum either directly in terms of courses and textbooks or indirectly in terms of a more general familiarity with and use of languages and images of geographic information systems. But it has also been generalized in other ways: it has become the tool of choice among many nongeographers whose work demands that they handle spatially-referenced information (ranging from botanists to transport engineers), it has merged technically with other image processing and data-capture technologies (ranging from the theoretically simple GPS of the surveyor to the more intellectually challenging three and multidimensional virtual representations and simulations of geography, architecture, geology, atmospheric physics, fluid mechanics, and even the arts), and it has become the province of commercial software and hardware developers and vendors whose relations with researchers at universities have become, of late, increasingly tenuous».

apparve fin da subito dominata dagli informatici, in quella delle applicazioni emerse la figura professionale del *GISer*¹⁴ che l'accademia, e particolarmente la geografia, ebbero il compito di formare: si tratta di un'ulteriore dimensione del profitto (*big business*) insito nel GIS, anche per l'accademia¹⁵.

In altri termini, l'enorme sviluppo commerciale del GIS, avvenuto prevalentemente negli anni '80, comportò grosse distorsioni in quanto il suo fulcro venne spostato dalla "produzione di ricerca" al "soddisfamento del mercato"; inoltre la produzione di "pacchetti GIS", indirizzata sempre più dai grandi produttori, seguì ritmi elevatissimi spinti più dal marketing, che poggiava spesso su soluzioni "d'effetto", che su autentici progressi metodologici. In tale situazione è evidente come il ruolo principale sia stato svolto da coloro che materialmente realizzavano i prodotti (gli informatici) o i *GISers* assieme a coloro che ne curavano gli aspetti di marketing: tutto questo evidentemente a spese del geografo, sempre più emarginato.

È certo che l'elevato valore commerciale del GIS contribuì, da una parte, alla sua fortuna, ma allo stesso tempo costituì un vero e proprio ostacolo per lo sviluppo dell'analisi spaziale, che rimane sempre, secondo Michael Goodchild, "ciò che il GIS è disegnato per fare"¹⁶.

4.4. E la scienza dell'informazione geografica?

Se dunque, da un lato, l'enfasi sulle applicazioni e la capacità di generare profitto tendevano ad allontanare il GIS dalla geografia, allo stesso tempo la geografia accademica sentiva il bisogno di non lasciarlo sfuggire, e di affrontare le questioni fondamentali di ricerca che erano state troppo a lungo trascurate. Da questa tensione emerse

¹⁴ Il "*GISer*" secondo il linguaggio nordamericano è colui che lavora nelle applicazioni del GIS. Traducibile in italiano, probabilmente, col termine *GISista*.

¹⁵ Si veda l'imponente critica che si trova nel testo di Pickles (1995a).

¹⁶ Goodchild (1992, p. 38): «Despite widespread recognition that analysis is central to the purpose of a GIS, the lack of integration of GIS and spatial analysis, and the comparative simplicity of the analytical functionality of many systems continues to be a major concern. In the early days of the statistical package SAS, there was a very rapid increase in the range of tests and techniques implemented in the system. Unfortunately, the same has not been true of GIS, and remarkably little progress has been made in incorporating the range of known techniques of spatial analysis into current products».

l'appello per una *Scienza dell'Informazione Geografica*, lanciato per la prima volta da Michael Goodchild nel 1990¹⁷, che condusse al suo debutto sul palcoscenico accademico, sancito dal cambiamento del titolo della più autorevole rivista internazionale di GIS, da *International Journal of Geographical Information Systems* ad *International Journal of Geographical Information Science* nel 1997, motivato in un editoriale dell'allora direttore Peter Fisher¹⁸.

In quel periodo l'NCGIA (*National Center for Geographic Information and Analysis*) definiva la scienza dell'informazione geografica: «il campo di ricerca fondamentale che tenta di ridefinire i concetti geografici e il loro uso nel contesto dell'informazione geografica, e, più in generale, nell'era digitale. Essa riesamina alcuni tra i temi più fondamentali nei campi tradizionalmente orientati allo spazio, come geografia, cartografia e geodesia, mentre incorpora i più recenti sviluppi della scienza cognitiva e della scienza dell'informazione, e sta cominciando ad abbracciare temi di ricerca più specializzati in discipline consolidate come l'informatica, la statistica, la matematica e la psicologia»¹⁹.

¹⁷ La proposta fu esposta al *Fourth International Symposium on Spatial Data Handling* tenutosi a Zurigo, poi pubblicata dall'*International Journal of Geographical Information Science*, in un articolo dal titolo tanto sintetico quanto eloquente: *Geographical Information Science* di Goodchild (1992).

¹⁸ Fisher (1997, p. 1): «Over the 10 years of the publication of IJGIS the field to which it contributes has evolved considerably, and particularly the meaning of the 'S'. Goodchild (1992) presented a coherent and influential argument that it was the science underlying the system that was important. He documents the reasons for this, and highlights the rich diversity of the possible subject matter of Geographical Information Science. He even makes the point that in developing the Canadian Geographical Information System (CGIS), widely recognized as the first geographical information system, the fundamental issue was developing the science to the point where the system could be developed. Indeed, it is apparent that in the 10 volumes of IJGIS to date, articles have contributed not to the understanding or development of systems as such, but to the science which underpins and exploits the systems. That science has all the diversity of the contribution we have come to expect in IJGIS, including advances in cadastral systems, spatial visualization, diffusion of technology, human computer interaction, algorithm developments, diffusion of innovation, spatial indexing, topology, uncertainty modelling and handling, etc., with contributors from many and varied parent academic disciplines, particularly geography, computer and information science, and surveying (geomatics), as well as the application sciences, including ecology, geology, civil engineering, demography, land management, and the like».

¹⁹ (<https://escholarship.org/uc/item/7mh0r0bt>): «Central to our research plan are geographic concepts and the notion of a geographic information science. While geographic information systems (GISs) are ubiquitous and expertise in GIS is in great demand, we foresee that in the long term the development of GISs must rest on a strong scientific basis, as provided

Nell'ormai classico lavoro del 1992, Michael Goodchild²⁰ articola il contenuto della scienza dell'informazione geografica in otto punti che elenchiamo di seguito:

- a. Raccolta e misurazione dei dati;
- b. Acquisizione dei dati;
- c. Statistica spaziale;
- d. Modellizzazione dei dati e teorie dei dati spaziali;
- e. Strutture, algoritmi e processi di dati;
- f. Visualizzazione;
- g. Strumenti analitici;
- h. Questioni istituzionali, manageriali ed etiche.

Si tratta evidentemente di una serie di questioni importanti, di cui vale la pena osservare che per una buona parte sono quantomeno pertinenti, quando non addirittura centrali, all'analisi spaziale²¹. Pur senza

through the multi-disciplinary field of geographic information science. [...] The field of geographic information science is too broad to be studied in its entirety. Therefore, we have identified areas that we consider to provide the highest potential to advance geographic information science within the near future. These strategic areas of our research highlight the role of geographic information science in the era of information technology. [...] Scientific Motivation. Research in geographic information science serves the needs of science and scientists in two ways. First, our research will address areas where our understanding of key geographic notions and their appropriate representations is currently incomplete. We see such basic research as particularly important at this time in areas where human conceptualization and the digital implementation of corresponding concepts interact and conflict [...]. Second, our research will contribute to the conceptualizations, methods, and tools with which scientists approach geographically distributed phenomena. Thus it will contribute to the infrastructure of science, particularly for those disciplines whose subject matter is distributed over the Earth's surface and for which a geographic perspective is likely to prove useful [...]. Technological Motivation. Our second motivation derives from the technology push that we are currently experiencing. The proliferation of faster computer hardware and the emergence of an information infrastructure are quickly changing the way people think and work, creating digital worlds. Societal Motivation. Bullish predictions of a glowing future for GIS aside, we also believe it is important that such a fast-moving and groundbreaking technology be subjected to the kinds of dispassionate reflection at which the academic research sector excels. No other group is likely to take the kinds of long, hard looks at GIS that are needed if its benefits to society are to be maximized, and its potential abuses avoided or controlled».

²⁰ Goodchild (1992, pp. 34-40, corsivi miei: si riportano in corsivo solo i titoli dei sottoparagrafi): «Having established that geographical information has unique properties and problems, we can now review the set of generic questions which might make up a geographical information science. [...] *Data collection and measurement* [...] *Data capture* [...] *Spatial statistics* [...] *Data modeling and theories of spatial data* [...] *Data structures, algorithms and processes* [...] *Display* [...] *Analytical tools; Institutional, managerial and ethical issues*».

²¹ Su queste questioni si tornerà più avanti.

soffermarci su ciascuna, osserviamo in particolare il punto (c), talmente centrale da rendere superflua qualsiasi chiosa, il punto (d) che, benché si riferisca soprattutto al modello dei dati, spazia fino ad abbracciare contenuti di portata assai più ampia²², il punto (g), benché anch'esso si riferisca, in parte, alle cosiddette operazioni analitiche spaziali elementari²³, e infine il punto (h), che sembra delineare (già da allora) la possibilità di un dibattito critico anche all'interno della nascente scienza dell'informazione geografica.

Tuttavia, non si può trascurare che fu proprio la legittimazione accademica del GIS come scienza a far sorgere dei dubbi su quella legittimità, minata dalla vicenda storica caratterizzata dal successo, dal profitto, dall'urgenza, e dall'incertezza del suo posizionamento disciplinare. Per questo sorse un importante dibattito volto a stabilire se “fare GIS” sia davvero “fare scienza”, da cui emerse una pluralità di posizioni, comunque riconducibili a tre filoni trainanti: se “fare GIS” sia “fare scienza”, se sia “utilizzare uno strumento, o, meglio, un attrezzo”, o se sia “costruire un attrezzo”. La questione era chiaramente intrecciata alla necessità di (ri)definire il ruolo non solo del GIS nella geografia, ma anche della geografia nel GIS, poiché, paradossalmente, proprio la geografia rischiava di rimanerne tagliata fuori. Travolto dal proprio successo, il GIS avrebbe infatti forgiato, soprattutto all'inizio, una generazione di accademici che erano poco più che tecnici, eppure in posizione tale da poter facilmente divenire “scienziati dell'informazione geografica”. Si legge in tutto questo da un lato l'urgenza di garantire un futuro al GIS nelle Università, che potesse perdurare oltre il periodo caratterizzato dall'enfasi tecnologica, ma dall'altro anche la necessità di affrontare questioni fondamentali, di cui la comunità accademica era consapevole da lungo tempo, ma che non aveva saputo affrontare, costretta a far fronte agli eventi.

²² Goodchild (1992, p. 37): «These issues are precipitating lively discussion over the entire question of the degree to which we view, analyse, represent and model the world as discrete or continuous, as a collection of objects or a set of fields. Do we think in terms of variables with defined values everywhere in space, or of an empty space littered with possibly overlapping objects? In essence, these issues have brought the GIS debate from the comparative obscurity of internal data structures to the much more general issues of how we understand geographical variation».

²³ Tipiche sono le operazioni di sovrapposizione (*overlay*), *buffer*, e *point-in-polygon*, che saranno discusse nel seguito.

Secondo Dawn Wright e colleghi: «parte del dibattito sul GIS riflette il dibattito degli anni '50 sulla quantificazione in geografia, ma la natura interdisciplinare del GIS lo porta ad una maggiore visibilità. [...] il dibattito stesso sulla natura del GIS può essere un modo di far (ri)entrare nella disciplina non tanto i *GISystems* ma la *GIScience*. Vi sono poi le implicazioni sociali: il messaggio che manda, a chi dà potere, e le responsabilità di chi lo sviluppa per il suo uso futuro»²⁴. Fu un dibattito di notevole complessità ed importanza, in cui la geografia accademica appariva consapevole del grande potenziale scientifico del GIS: una disciplina dalle molteplici implicazioni, i cui aspetti scientifici erano stati a lungo trascurati. È pur vero che permaneva l'innegabile necessità di difendere posizioni non ancora consolidate, e sempre più a rischio man mano che il GIS tendeva a sfuggire dalle mani della geografia, ma vi era anche un desiderio genuino di affrontare le questioni rimaste aperte nel momento in cui le pressioni tecnologiche e commerciali si fossero allentate.

Era quindi importante comprendere se il GIS fosse un attrezzo o una scienza, e se meritasse un posto nell'accademia e nella geografia: se si trattava di scienza, sarebbe stata la geografia come disciplina accademica a trarne beneficio, anche in termini di cattedre e finanziamenti, ma soprattutto del nuovo equilibrio di poteri che tutto ciò avrebbe determinato. Se, invece, si trattava di attrezzo, essa avrebbe dovuto cedere il passo ai fabbricanti di attrezzi, i *toolmakers*, lasciando che cattedre e finanziamenti fossero ripartiti tra coloro che costruivano l'attrezzo [prevalentemente gli informatici] e i numerosi quanto variegati settori in cui l'attrezzo era applicato²⁵.

²⁴ Wright, Goodchild, Proctor (1997, p. 346): «Of course geography is no stranger to methodological debates, and some of the arguments over GIS echo the arguments over quantification in the 1950s. We cannot be certain whether the GIS debate is more or less intense than its precursors, but we can be certain that it is important given the interdisciplinary nature of GIS, albeit a nature in which geography is widely accepted as having a unique role (Morrison 1991, Kennedy 1994). Geography's debates over GIS are thus unusually exposed to general view. [...] More interesting are the social implications of GIS — the message it sends, whom it empowers, and the responsibility its developers should bear for its eventual use (Smith 1992, Pickles 1995, Harvey e Chrisman (2004)». Si veda anche Pickles (1997).

²⁵ Si vedano ancora Wright, Goodchild, Proctor (1997, p. 358): «Surely the desired end from all perspectives is the building of an intellectual foundation for GIS by geographers and members of allied disciplines alike, which will ensure its survival long after the novelty of the technology has worn off» e Pickles (1997, p. 369): «In effect, the authors seem to be calling for Geographic Information Science to give theoretical and analytical legitimacy to GIS, in

Per molti che si riconoscevano nella comunità del GIS²⁶ era anche importante comprendere a fondo il significato dell'espressione "fare GIS": se fuori dall'accademia l'attività di GIS è definita dagli obiettivi dell'attività stessa e di chi la svolge, più complessa è la sua definizione in ambito accademico²⁷. Secondo David Maguire, in quest'ultimo caso le difficoltà sorgono nel definire l'obiettivo centrale dell'attività di GIS, che potrebbe consistere nello sviluppo di *hardware* e *software*, fino all'elaborazione di informazioni o addirittura allo sviluppo di applicazioni specifiche²⁸. Per Michael Goodchild anche solo manipolare l'informazione spaziale con la tecnologia GIS presenta una gamma di sfide intellettuali e scientifiche di assai più ampio respiro di quanto implicato dal "manipolare dati spaziali"²⁹. In ogni caso Dawn Wright e colleghi arrivano a definire tre posizioni base, suggerendo però l'esistenza di una continuità di posizioni intermedie, cosicché "fare GIS" è un'attività spesso difficile da classificare entro una sola di queste definizioni:

much the same way that Regional Science emerged in the 1960s to provide an intellectual setting within which spatial analysis and spatial science could develop and to serve as a support to regional economics, regional planning, and geographical analysis. The GIS in which technical rationality dominates so thoroughly now must give way to a science of GIS: the scholarly investigation of its origins, logics, systems, new capacities, and new uses. Geographical Information Science is to emerge to deal with the demands of a booming field of geographic information systems and their uses».

²⁶ Il già citato articolo di Wright, Goodchild, Proctor (1997) prende le mosse da un dibattito svoltosi nei mesi di ottobre e novembre 1993 sulla lista elettronica di discussione "GIS-L" sulla questione strumento vs. scienza; alla discussione presero parte 40 individui di sei Paesi, che produssero in totale 64 posizioni sulla lista (Wright, Goodchild, Proctor, 1997). Benché la discussione non avesse prodotto alcun argomento conclusivo, da essa emersero una serie di posizioni significative.

²⁷ Per una ditta produttrice di software, fare GIS significa produrre e sviluppare software per lo svolgimento di funzioni specifiche (creazione e utilizzo di database, rappresentazioni cartografiche, funzioni analitiche più o meno complesse). Per un ente locale, fare GIS significa generalmente catalogare informazione geografica e svolgere funzioni di database e analisi per la soluzione di problemi specifici e per la presentazione al pubblico e al decisore di situazioni, proposte o simulazioni dotate di dimensione spaziale. Società private o di consulenza avranno obiettivi definiti specificamente dai rispettivi progetti, che comunque si collocano nel rango d'attività appena definito.

²⁸ Maguire (1991, p. 10): «The final reason for definitional difficulties stems from genuine academic debate about the central focus of current GIS activity. As the discussion below demonstrates, some people believe that hardware and software are the central focus, others argue that the key element is information processing or even applications».

²⁹ Si veda in proposito la lista di problemi definita da Goodchild (1992) e riportata sopra.

- 1) se il GIS è una “cassetta degli attrezzi” per la scienza, allora il GIS è un *utensile* e “fare GIS” significa fare uso di una particolare attrezzatura che, pur complessa, rimane comunque distinta e neutrale rispetto all’obiettivo perseguito;
- 2) se il “fare GIS” corrisponde all’attività di un artigiano o meglio di un ingegnere che costruisce degli *attrezzi*, allora “fare GIS” coincide con la professione del “fabbricante di attrezzi”³⁰ e l’attività del GIS riceve scopo e il significato dal suo stesso oggetto, ossia lo *strumento*³¹ prodotto;
- 3) se infine “fare GIS” è fare scienza, allora sono i suoi elementi concettuali che consentono allo *strumento* GIS di reclamare un posto come scienza³².

La posizione secondo cui il GIS è un *utensile* identifica dunque la pratica del GIS con l’uso di una particolare classe di *software*, i relativi strumenti *hardware* assieme ai dati geografici digitali necessari per perseguire un proposito specifico: in questo caso lo strumento in sé stesso è per sua natura neutrale e quello che conta è solamente la capacità del ricercatore³³. O meglio, gli strumenti, i metodi o le tecniche non hanno, ai fini della conoscenza, un fondamento obiettivo, cioè, per dirla con Karl Popper “i metodi sono privi di importanza, qualsiasi metodo è legittimo se conduce a risultati suscettibili di una discussione razionale. Quello che conta non sono i metodi o le tecniche, ma una certa sensibilità ai problemi, e un’ardente passione per essi; o, come

³⁰ “*Toolmaker*” nella letteratura nordamericana.

³¹ Il termine inglese “tool” può essere tradotto in italiano con “attrezzo”, “utensile” o “strumento”. Si è qui utilizzato il termine “attrezzo” (nei due enunciati *cassetta degli attrezzi* e *fabbricante di attrezzi*) in modo neutro, senza cioè fare nessuna considerazione alla possibile interpretazione che sottostà alla parola “attrezzo”. Al contrario i due termini “utensile” e “strumento” sono stati qui utilizzati secondo l’interpretazione di Koyré (2000, p. 101): *utensile* «qualcosa che prolunga e rinforza l’azione delle nostre membra, dei nostri organi sensibili», mentre *strumento* «non è un prolungamento dei sensi, ma nell’accezione più forte e più letterale del termine, incarnazione dello spirito, materializzazione del pensiero»; per una migliore analisi dei due termini su veda l’intero volume di Koyré (2000). In ogni caso, nel prosieguo del testo, i due termini si utilizzeranno sempre con questa accezione.

³² Si veda al riguardo Wright, Goodchild, Proctor (1997).

³³ Sulla “neutralità” dello strumento si veda Lando (1985, p. 253): «se [...] vi sono stretti rapporti tra i tre elementi dell’analisi (dati-strumento-risultati) [...] essi non sono diretti bensì indiretti in quanto fra essi si pone sempre un portatore di ideologia: cioè il ricercatore che è colui che immette “ideologia” sia nei dati che nello strumento e che interpreta, dal suo punto di vista, i risultati».

dicevano i Greci, la dote naturale di provare meraviglia”³⁴. Se il GIS è impiegato come attrezzo per la ricerca scientifica, anche fare GIS può essere fare scienza, ma l’esistenza e l’uso dello strumento restano comunque separabili dal problema scientifico³⁵. Secondo questa posizione, dunque, il GIS è accettato ed usato solamente per ciò che consente di fare: ne consegue che chiunque lo usa fa GIS, e chiunque può fare GIS. In questo modo, fare GIS ha ben poco a che fare sia con la geografia sia con la scienza.

Secondo la posizione del “fabbricante di attrezzi” colui che “fa GIS” è impegnato a migliorare le capacità e la facilità di uso dell’*attrezzo*, promuove l’adozione del GIS, gioca un ruolo nell’istruire gli utenti, e s’impegna per assicurarne un uso responsabile. In assenza di particolari istanze o situazioni in cui appaia inconfutabile che particolari e straordinarie illuminazioni scientifiche siano unicamente attribuibili all’uso del GIS, l’attività del *fabbricante di attrezzi* sarà definita più propriamente come ingegneria che come scienza. Qual è allora la posizione e il ruolo del geografo in tale attività di ingegneria? Due aspetti appaiono particolarmente importanti. Il primo concerne le difficoltà del geografo rispetto al livello tecnologico del GIS: ben si sa che pochi geografi possiedono le capacità tecniche necessarie a costruire grandi sistemi di *software* o forse a scrivere codici “a tenuta industriale”³⁶. L’altro aspetto riguarda le implicazioni ideologiche, sociali, culturali dello sviluppo dell’attrezzo e che, se affrontate appieno, spostano quest’attività su un livello ben diverso dalla semplice ingegneria. A tutto questo gli autori aggiungono un’osservazione che in un certo senso conclude il dibattito sul ruolo imprescindibile del geografo anche nella *fabbricazione degli attrezzi*. I geografi, infatti, possiederebbero «un’eccellente comprensione dei concetti geografici che formano gli elementi primitivi dei database e dell’elaborazione dei GIS e i modi in cui questi concetti sono incorporati nelle teorie, nei metodi, nei modelli e nell’analisi»; sempre i geografi sarebbero poi «esperti in una disciplina che integra la comprensione di una vasta gamma di

³⁴ Popper (1994, p. 127).

³⁵ Wright, Goodchild, Proctor (1997, p. 355): «If the investigation merits the label “research,” then “doing GIS” is probably “doing science” as well, but the existence and use of the tool are separable from the substantive problem».

³⁶ Questo problema però è comune a quasi tutti gli accademici salvo, forse, gli ingegneri e gli informatici.

processi che influenzano i fenomeni sulla superficie della terra». Entrambe queste abilità sono così essenziali per “fare GIS” nel senso di *fabbricarne gli attrezzi*, che «un *fabbricante di attrezzi GIS* necessita un’istruzione di base in geografia assieme a [...]»³⁷. In altri termini, non è possibile *fabbricare gli attrezzi GIS* senza un geografo. In questo senso, si può pensare al ruolo del geografo nell’attività di costruzione dello strumento come a quello di un ponte che collega l’attività d’ingegneria tra due rive: a monte con la scienza geografica, e a valle con la comunità degli utenti. In questo caso, il geografo non svolge l’attività d’ingegneria, ma si occupa dei suoi presupposti scientifici e delle sue implicazioni (sociali, culturali, ambientali, economiche, ecc.). Così il ruolo del geografo, nell’attività di fabbricazione degli attrezzi, è caratterizzato proprio dalle sue capacità specifiche: egli è “confinato” ad occuparsi degli aspetti concettuali che sono alla base e che emergono dalla fabbricazione e dall’uso degli attrezzi GIS. In questo caso, il geografo se “fa GIS” non produce né *hardware* né *software*, ma “fa scienza”: così, in questo caso, lo strumento non può più essere considerato neutrale nei termini popperiani, detti prima.

Affascinante è il parallelo proposto da John Pickles (1997)³⁸ con il telescopio di Galileo nello schema “attrezzo/fabbricante/scienza”. «L’astronomia non sarebbe possibile nella sua forma post-galileiana

³⁷ Wright, Goodchild, Proctor (1997, p. 356): «Geographers possess two unique and powerful abilities as GIS toolmakers. The first of these is an excellent understanding of the geographic concepts that form the primitive elements of GIS databases and processing and the ways that these concepts are embedded in theories, methods of analysis, and models. Second is that geographers are trained in a discipline that integrates understanding of a wide range of processes influencing phenomena on the earth’s surface. Both of these abilities are essential to “doing GIS” if one adopts the position that “doing GIS” is toolmaking. A GIS toolmaker thus requires a basic education in geography together with technical courses that emphasize critical analysis of the technology’s capabilities».

³⁸ Pickles (1997, pp. 365-370) « Clearly, astronomy would not be possible in its post-Galilean form without the telescope, which allowed us to see more and to see differently, and as a result transformed our view of the universe and our place in it. But it would also not have been possible without the broader context of need and demand, and legitimization of its practices, generated by so mundane a matter as sea travel and commercial expansion. Moreover, its success would have been unlikely without a host of parallel and related developments in mathematics, physics, geology, and philosophy, themselves tied, at least indirectly, to specific social contexts of production. Moreover, what notions of world, earth, and space have been sustained and enhanced by the emergence of an abstract science of stellar mapping and a metaphysics of a God’s eye view? [...] it is now high time that the practitioners of GIS accept the call to the hard theoretical work that will change their thinking, the forms and uses of their geographic information systems, and, we hope, the worlds in which they are deployed».

senza il telescopio, ma non sarebbe possibile neppure senza la domanda di conoscenza astronomica e la legittimazione delle sue pratiche generata da un problema estremamente materiale, quale la navigazione marittima e l'espansione del commercio. Analogamente la riuscita del telescopio è legata ad una miriade di sviluppi paralleli e collegati in matematica, fisica, geologia e filosofia, a loro volta collegati, almeno indirettamente, a specifici contesti sociali di produzione. Inoltre, quali nozioni del mondo, della terra e dello spazio sono state sostenute e migliorate dall'emergere di una scienza astratta della mappatura delle stelle e di una metafisica di una visione dell'occhio di Dio? [...] è proprio ora che i praticanti del GIS accettino la chiamata al duro lavoro teorico che cambierà il loro modo di pensare, le forme e gli usi dei loro sistemi d'informazione geografica e, speriamo, i mondi in cui essi sono impiegati».

Se Wright, Goodchild, Proctor (1997) vedevano una gamma di posizioni tra l'uso di un attrezzo, la sua fabbricazione, e la scienza di e con l'attrezzo, la metafora di John Pickles indicava inequivocabilmente una scienza che non poteva più attendere di essere sviluppata, ma in cui la geografia si trovava caricata di un gravosissimo fardello di responsabilità: che ne sarebbe stato del mondo e dell'umanità se Galileo non avesse accettato di affrontare e sviluppare la scienza che era in potenza nel suo telescopio? Così spettava alla geografia saper far sbocciare dalla tecnologia dei GIS la Scienza dell'Informazione Geografica.

E allora, quali sono il posto e il ruolo dell'analisi spaziale in questa scienza?

4.5. Che cos'è dunque l'analisi spaziale?

Secondo Alan Murray (2010) l'analisi spaziale è una delle attività che costituiscono la geografia quantitativa, ovvero è parte di una articolazione della geografia che risale alla metà del XX secolo. Secondo Michael Goodchild l'analisi spaziale è proprio “ciò che il GIS è stato creato per fare”³⁹, ed egli stesso situa la sua data di nascita negli anni

³⁹ Goodchild (1992, pp. 31-32): «Geographical information systems are sometimes accused of being technology driven, a technology in search of applications. That seems to be

che precedettero la realizzazione del CGIS di Roger Tomlinson nel 1967. Eppure, ancora all'inizio degli anni '90 David Maguire⁴⁰ affermava che la fase dell'analisi del GIS doveva ancora compiersi, e ancora quasi un decennio dopo Stan Openshaw lamentava la gravità e la diffusione di ciò che egli stesso definiva come il "crimine dell'analisi spaziale", commesso da quanti, in possesso di dati spaziali, si limitano a visualizzarli, classificarli, aggiornarli, ma non li analizzano⁴¹.

Stan Openshaw esprime con lucida chiarezza, pur se con caratteristici toni provocatori, non solo l'importanza, ma addirittura la necessità e l'impellenza dell'*analisi* dei dati spaziali e, benché di essa non dia una precisa definizione, tuttavia fornisce una serie di chiarimenti importanti. L'affermazione secondo cui "*no one bothers to properly analyse*" contiene due elementi chiave; il primo è l'uso del verbo *bother*, che indica un particolare atteggiamento, e che si potrebbe tradurre come "nessuno si prende la briga di ...", quindi non è una semplice rinuncia a compiere un'analisi, ma piuttosto un rifiuto di affrontare le complicazioni dell'analisi, come espresso, nella stessa

more true of some periods of the 25-year history of GIS than of others. For example, it is difficult to suggest that Tomlinson and the developers of the Canadian Geographical Information System (CGIS) (Tomlinson *et Alii* (1976)) were driven by the appallingly primitive hardware capabilities of 1965. On the other hand the prospect of a menu driven, full colour, pull-down menu raster GIS in the 386-based personal computer on one's desk has clearly sold many systems in the past few years. Technological development comes in distinct bursts, and so does the technology drive behind GIS. It may be the motivation behind the desire to handle spatial data, but it fails to explain many of the diverse research efforts being reported at meetings and in the literature».

⁴⁰ Maguire (1991, p. 17): «Thus in the early 1990s, while there are many well developed GIS at the inventory stage, there is a much smaller number which are used for analytical purposes and only a select few which are employed in a management context».

⁴¹ Openshaw (1998, p. 7): «Openshaw (1993) coined the term "spatial analysis crime" to describe users of GIS who have successfully created databases relating to all kinds of potentially useful information but who then fail to analyse it [...]. Vendor GIS slogans such as "just map it" reflect this neglect of more in depth analysis to explore the patterns, relationships, and processes hidden in the mappable data. The vendors often claim (wrongly) that users are not interested in geographical analysis, or if they are, then the vendors (either accidentally or deliberately) confuse the end-users by having spatial analysis modules that equate spatial analysis with little more than polygon manipulation, overlay, and buffering. By way of example note how MapInfo defines spatial analysis: "An operation that examines data with the intent to extract or create new data that fulfils some required condition or conditions. It includes such GIS functions as polygon overlay or buffer generation and concepts of contains, intersects, within or adjacent." (Page 396, MapInfo professional User Guide, 1995). Yet this is not a definition of spatial analysis any quantitative geographer or spatial statistician would accept as being useful».

proposizione, dall'avverbio *properly*, che suggerisce che una qualche forma di analisi venga sì condotta, ma che non sia condotta propriamente – potremmo dire che si tratta di analisi che non sono fatte “bene”, oppure “come si deve”.

Interessante è poi la critica all'atteggiamento dei venditori di *software* GIS, questa ancor più articolata, innanzitutto con lo slogan “*just map it*”⁴², che effettivamente tende ad attirare l'attenzione sull'importanza (o la necessità) di far comparire checchessia sulla mappa, ma non quella di comprendere il senso di quelli che definisce “dati mappabili”. Si direbbe, leggendo il passo di Stan Openshaw, che il problema non sia la mancanza di interesse per l'analisi da parte dei potenziali utenti, ma da parte dei venditori. Infatti, benché egli non ne spieghi le ragioni, si può supporre che per loro sia più semplice o economico mettere in commercio sistemi per la manipolazione grafica dei dati piuttosto che sistemi per la loro analisi. Egli continua poi enfatizzando la differenza tra quella che i venditori chiamano *spatial analysis*, ma che in realtà si riduce alle classiche operazioni grafiche di manipolazione di poligoni, sovrapposizione (*overlay*) e *buffering*. L'autore poi conclude affermando che si tratta di definizioni per cui «nessun geografo quantitativo accetterebbe questa definizione come qualcosa di utile».

Questo tipo di operazioni elementari, tipicamente svolte in GIS, sono esemplificate in Fig. 4.1, che illustra, rispettivamente, le operazioni di selezione per attributo [*selection by attribute*] (a), *buffering* (b), e selezione per posizione [*selection by location*] (c).

⁴² Di resa meno efficace in italiano, si potrebbe dire “basta metterlo sulla mappa”.

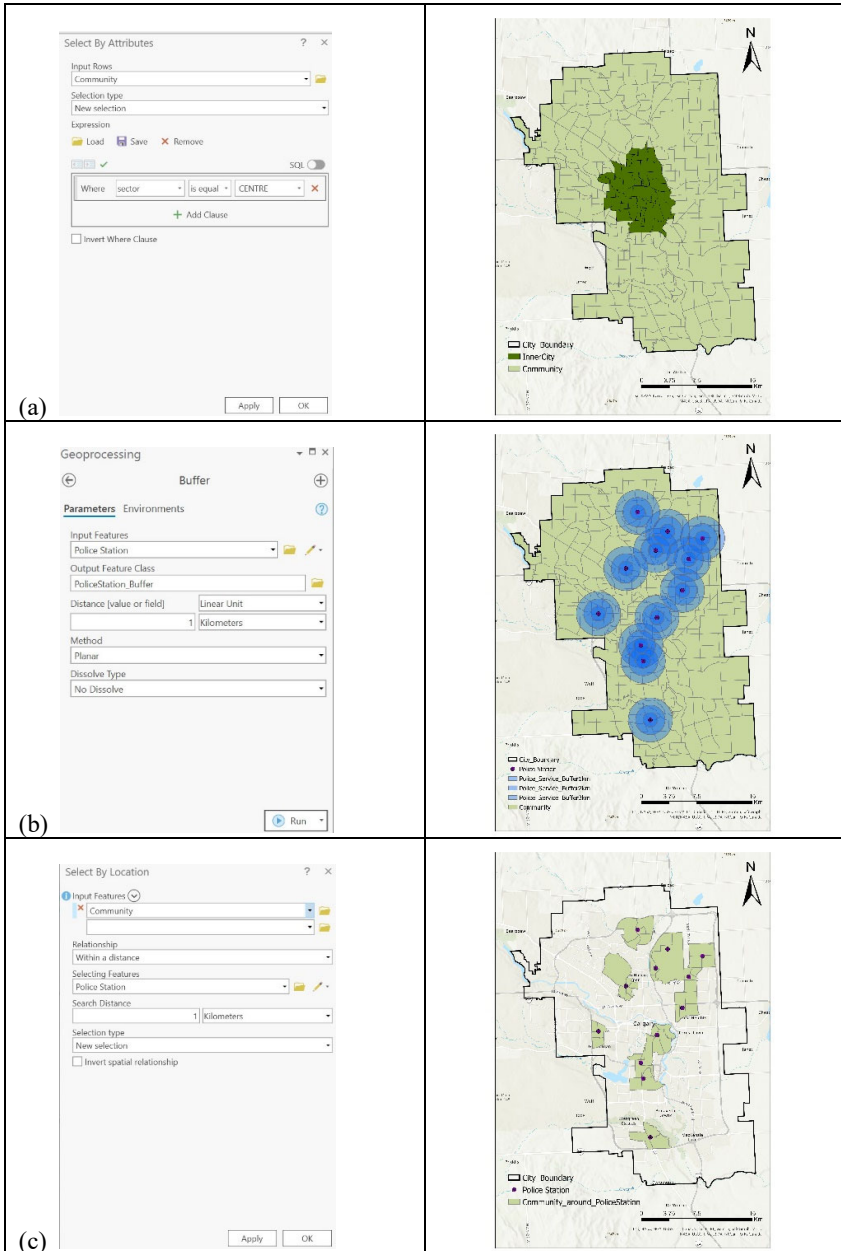


Fig. 4.1 - Operazioni di selection by attribute (a), buffering (b), selection by location (c).

Costruire un modello analitico significa porre delle ipotesi, verificarle, valutare l'errore insito nel modello, verificare le proprietà statistiche dell'errore e delle stime. Definire un *buffer* attorno ad una figura geometrica è un altro mestiere: non richiede ipotesi e verifiche, ma solo l'identificazione della figura e la definizione dell'ampiezza del *buffer*, e analogamente le altre operazioni analitiche elementari sono operazioni geometriche e deterministiche. E anche in questo sta la differenza fondamentale tra operazioni deterministiche oppure stocastiche, poiché le ultime, ma non le prime, richiedono *necessariamente* l'impiego della statistica, ciò che, tra l'altro, conduce all'uso del termine "analisi statistica spaziale", o anche semplicemente "statistica spaziale" per differenziare quel tipo di analisi da quella elementare, geometrica e deterministica, alla quale riserviamo la definizione di "analisi spaziale elementare".

Osserviamo infine che la demarcazione tra i due tipi di analisi spaziale può non essere sempre così netta, poiché esistono analisi spaziali che possiamo dire intermedie, e questa distinzione va vieppiù sfumandosi con l'aumentare della sofisticazione delle analisi possibili nei pacchetti GIS. Su questo aspetto torneremo nei prossimi capitoli, perché apre un'altra questione importante, che non attiene tanto all'analisi in sé, né alle potenzialità dello strumento, quanto alla preparazione e all'atteggiamento mentale del cosiddetto "utente" e di come questi si pone nei confronti dell'analisi e delle eventuali ipotesi e questioni concettuali che l'operazione può richiedere.

Benché il termine "analisi spaziale" sia frequentemente usato in contesto GIS per indicare tanto i modelli che le analisi elementari, proprio la distinzione tra questi due significati è uno degli elementi chiave della definizione di analisi spaziale⁴³.

Il ragionamento di Stan Openshaw appare perfettamente allineato con l'argomento di Michael Goodchild in favore di una scienza dell'informazione geografica, di cui una gran parte degli elementi chiave sono direttamente legati all'analisi spaziale e che qui vale la pena qui approfondire. L'importanza della statistica spaziale⁴⁴ deriva

⁴³ Come si è osservato, tale ambiguità è presente anche in Goodchild (1992).

⁴⁴ Goodchild (1992, pp. 35-36): «*Spatial statistics*. As spatial data are always an approximation or generalization of reality, they are full of uncertainty and inaccuracy. [...] We need much better methods of measuring and describing uncertainty, particularly in the complex spatial objects common in GIS. We need better methods for dealing with the world as a set of

dal fatto che «i dati spaziali sono sempre un'approssimazione o generalizzazione della realtà, sono pieni di incertezza ed inaccuratezza. Per questo c'è bisogno di metodi migliori per misurare e descrivere l'incertezza, e la gran parte delle risposte a questi problemi verranno dalla statistica spaziale». È anche interessante notare come l'autore chiuda il paragrafo dedicato alla statistica spaziale affermando che quest'ultima «non è un campo semplice, e molte di queste tecniche vanno ben oltre la statistica elementare nella loro sofisticazione concettuale».

Anche Michael Goodchild lamenta la carenza di analisi spaziale in GIS, nonostante proprio questa ne sia lo scopo ultimo⁴⁵, e si spinge ad auspicare una maggiore integrazione tra GIS ed analisi spaziale, analizzando le cause del ritardo⁴⁶, che identifica essenzialmente in tre aspetti: i primi due riguardano l'enfasi sull'informazione anziché sull'analisi geografica, e gli elevati profitti, per cui i venditori spingono per immettere sul mercato prodotti anche semplici ma in tempi rapidi, riecheggiando in questo senso l'argomentazione di Stan Openshaw. Il terzo aspetto, invece, va al cuore della scienza, poiché si riferisce ai problemi della rappresentazione dei dati spaziali e alla

overlapping continua, instead of forcing the world into the mould of rigidly bounded objects. Most of the answers to these questions will have to come from spatial statistics, [...]. The consequences of uncertainty for GIS products are never estimated. [...] However, spatial statistics are not an easy field, and many of these techniques go well beyond elementary statistics in their conceptual sophistication».

⁴⁵ Goodchild (1992, p. 38): «Despite widespread recognition that analysis is central to the purpose of a GIS, the lack of integration of GIS and spatial analysis, and the comparative simplicity of the analytical functionality of many systems continues to be a major concern».

⁴⁶ Goodchild (1992, p. 38): «[...] remarkably little progress has been made in incorporating the range of known techniques of spatial analysis into current products. There are many reasons for this. One obvious reason is the heavy emphasis in the GIS marketplace on information management rather than analysis. The lucrative markets for GIS technology have comparatively unsophisticated needs, emphasizing simple queries and tabulations. Another is the relative obscurity of spatial analysis, a set of techniques developed in a variety of disciplines, without any clear system of codification or strong conceptual or theoretical framework. [...] At this stage, integration of GIS and spatial analysis is proceeding slowly, in at least three different modes. Some analytical capabilities are being added directly to GIS, for example in the recent expansion of functionality in several modules for network analysis. Some progress is being made in loosely coupled analysis, where an independent analysis module relies on a GIS for its input data, and for such functions as display. However, still missing is an effective form of tight coupling, in which data could be passed between a GIS and a spatial analysis module without loss of higher structures, such as topology, object identity, metadata, or various kinds of relationships. At present this is impossible, to a large extent because of a lack of standards for data models. Instead, coupling has to occur at a lower level, and higher structures have to be rebuilt on an arbitrary basis».

conseguente difficoltà di eseguire su queste rappresentazioni delle operazioni analitiche. Benché non del tutto esplicita, l'argomentazione di Goodchild pare forse talvolta considerare anche analisi spaziali che potremmo dire complesse, ma perlopiù pare comprendere anche le operazioni cosiddette geometriche, o elementari.

La definizione di Michael Goodchild sancisce la centralità dell'analisi spaziale in GIS, il cui scopo non è visualizzare dati o creare mappe, ma, appunto, analizzare i dati spaziali. Più precisamente, sempre secondo lui, analisi spaziale è qualsiasi metodo di analisi i cui risultati dipendono dalle localizzazioni degli oggetti analizzati, ossia qualsiasi metodo analitico i cui risultati non sono invarianti rispetto alla localizzazione⁴⁷.

4.6. Il nome dell'analisi spaziale

Le difficoltà di dare una definizione precisa dell'*analisi spaziale* e di comprenderne l'essenza si traducono anche nella difficoltà di darle un nome. Sin dal capitolo 1 abbiamo indicato tre nuclei principali da cui l'*analisi spaziale* prese forma: la branca della geografia centrata sull'impiego di metodi quantitativi che fu sviluppata a partire dagli anni '50 e fu chiamata *geografia teoretico-quantitativa*; quella sviluppata a partire dai sistemi d'informazione geografica e che dagli anni '90 fu chiamata *scienza dell'informazione geografica*; e infine una serie di sviluppi intrecciati nella ricerca statistica, che ebbero il nome di *statistica spaziale*. Per la materia che si pone all'intersezione di questi tre filoni, in questo testo utilizziamo un nome, *analisi spaziale*, che tuttavia non è l'unico che in qualche modo la definisce, e che per questo non riesce a qualificarla senza ambiguità.

Forse per le ragioni che si sono discusse nel paragrafo precedente, forse per la parentela stretta con i filoni da cui, in parte, discende, e

⁴⁷ De Smith, Goodchild, Longley (2020): «The more difficult of the two questions above is the first — what should be considered as “geospatial analysis”? In conceptual terms, the phrase identifies the subset of techniques that are applicable when, as a minimum, data can be referenced on a two-dimensional frame and relate to terrestrial activities. The results of geospatial analysis will change if the location or extent of the frame changes, or if objects are repositioned within it: if they do not, then “everywhere is nowhere”, location is unimportant, and it is simpler and more appropriate to use conventional, *aspatial*, techniques».

forse anche per il grande fermento di idee che la circonda, si assiste quotidianamente all'uso di una molteplicità di termini, ciò che contribuisce a renderne sfuggente il contenuto. Abbiamo provato a raccogliere tutti questi termini, identificandone ben 16 nella letteratura anglofona, di cui proponiamo un elenco in Tab. 4.1.

Data la numerosità dei nomi elencati nella tabella, ne proponiamo dei raggruppamenti, che sono impliciti anche nell'ordine dell'elenco. Essi sono: i nuclei originari (1-4); i nomi centrati sul termine *analisi* (4-7 e 16); quelli centrati sul termine *scienza* (8-12, oltre a 2); quelli centrati sul termine *dati* (11-13, oltre a 5); i neologismi a sé stanti (14-15); e infine il grande assente, ossia il nome “fantasma” che, pur dando il titolo ad una delle più autorevoli riviste in materia⁴⁸, non compare nella letteratura, e che discutiamo assieme al gruppo *analisi*.

Tab. 4.1 - Terminologia relativa all'analisi spaziale

1.	Quantitative Geography
2.	Geographical Information Science
3.	Spatial Statistics
4.	Spatial Statistical Analysis
5.	Spatial Data Analysis
6.	Spatial Analysis
7.	Geospatial Analysis
8.	Geospatial Science
9.	Spatial Science
10.	Spatial Information Science
11.	Spatial Data Science
12.	Geographic Data Science
13.	Spatial Data Analytics
14.	Geostatistics
15.	Geocomputation
16.	Geographical Analysis

⁴⁸ Si tratta del nome *analisi geografica* (*geographical analysis*) che, se scritto con le iniziali maiuscole, costituisce il titolo della rivista *Geographical Analysis* (<https://onlinelibrary.wiley.com/journal/15384632>).

Nuclei originari. I primi nomi si riferiscono alle discipline consolidate in cui s'innestano le origini dell'analisi spaziale, di cui abbiamo già parlato diffusamente. Secondo Alan Murray, Michael Goodchild e Luc Anselin, non vi sono dubbi che l'analisi spaziale sia parte della geografia quantitativa⁴⁹ (*quantitative geography*), della scienza dell'informazione geografica (*geographical information science*), e della statistica spaziale (*spatial statistics*). Per quest'ultima consideriamo il termine analisi statistica spaziale (*spatial statistical analysis*) sostanzialmente come sinonimo di statistica spaziale (*spatial statistics*).

Analisi. Questo gruppo di nomi potrebbe essere quello che meglio caratterizza la nostra materia: analisi dei dati spaziali (*spatial data analysis*), analisi spaziale (*spatial analysis*) o analisi geospaziale (*geospatial analysis*). Dei primi due c'è poco da dire, salvo osservare la presenza di *data* nel primo, che l'accomuna anche ad un altro gruppo. Ci soffermiamo invece sul terzo nome, *geospatial analysis*, che differisce dal precedente, *spatial analysis*, solo per il prefisso *geo*. Etimologicamente il termine *geospatial* appare come un tentativo di integrare i termini geografico e spaziale, ossia di rivendicare lo spazio alla geografia⁵⁰. Non è chiaro tuttavia perché, volendo affermare l'affiliazione con la geografia, non si usi semplicemente *geographic* oppure *geographical*. In questo potrebbe giocare il desiderio di includere altre scienze, come geologia e geofisica, o, più probabilmente, quello di evocare la geografia in maniera indiretta passando, proprio per il GIS, la scienza e le tecnologie dell'informazione geografica.

In ogni caso, il nome *geospatial analysis* è usato frequentemente come sinonimo di *spatial analysis*⁵¹, sia nel senso di “analisi statistica spaziale”, sia in quello di “analisi spaziale elementare in GIS”. Sono proprio Michael De Smith e colleghi a fornire la seguente definizione: «In questa Guida affrontiamo l'intero spettro dell'analisi spaziale e delle tecniche di modellizzazione ad essa associate che sono fornite all'interno dei sistemi d'informazione geografica attualmente

⁴⁹ Benché in altri capitoli abbiamo privilegiato il termine “geografia teoretico-quantitativa”, qui preferiamo utilizzare il termine più recente utilizzato da Murray (2010) e Fotheringham, Brunson, Charlton (2000); in realtà quest'ultimo testo è intitolato *Quantitative geography. Perspectives on Spatial Data Analysis*.

⁵⁰ Potrebbe anche essere un banale espediente per chiarire che il termine spaziale non si riferisce ad agenzie spaziali e imprese extraterrestri.

⁵¹ Si veda De Smith, Goodchild, Longley (2020).

disponibili e comunemente usati e del relativo *software*. Collettivamente tali tecniche e strumenti sono spesso descritti come analisi geospaziale, anche se noi usiamo la forma più comune, analisi spaziale, nella maggior parte delle nostre discussioni»⁵². La definizione non spiega l'origine del termine, ma dà per assodata la sua accettazione nell'uso comune. Alla luce di quanto detto poco sopra e di quanto affermato da De Smith, Goodchild, Longley (2000), ci pare che il termine *geospatial analysis*, per quanto adeguato, sia più legato all'idea di GIS che a quella di *spatial analysis*. In altri termini, si ha la sensazione che non si possa avere *geospatial analysis* al di fuori del GIS, benché al suo esterno possa aversi *spatial analysis*. È difficile non richiamare, in questo contesto, il termine *geographical analysis*, che paradossalmente suona un po' vecchio e legato alla geografia quantitativa più tradizionale, e che tuttavia da 50 anni compare sulle pagine di *Geographical Analysis*⁵³ prescindendo, sostanzialmente, dal GIS.

Scienza. Il sostantivo *science* compare già nel nome *geographical information science*, tra i nuclei originari dell'analisi spaziale. La ritroviamo poi, qualificata dall'aggettivo appena discusso, nella forma di scienza geospaziale (*geospatial science*), termine dal suono superficialmente molto simile a scienza spaziale (*spatial science*).

Spatial science è un termine vecchio, usato un tempo per definire quella scienza spaziale che fu l'essenza della proposta della "rivoluzione quantitativa", quasi un sinonimo di geografia-teoretico-quantitativa. Che il termine scienza (*science*) definisca a buon diritto questa geografia ci pare fuori discussione⁵⁴. La vicenda della geografia a partire dagli anni '80 fu, come è noto, attraversata da alcune matrici disciplinari, spesso in contrasto con la geografia teoretico-quantitativa, che contribuirono anche alla parabola discendente del termine *spatial science*. «Tale termine si è invece nuovamente imposto in epoche

⁵² De Smith, Goodchild, Longley (2000, p. 19): «In this Guide we address the full spectrum of spatial analysis and associated modeling techniques that are provided within currently available and widely used geographic information systems (GIS) and associated software. Collectively such techniques and tools are often now described as geospatial analysis, although we use the more common form, spatial analysis, in most of our discussions» www.springer.com/journal/10707 (visitato il 01/04/2021).

⁵³ Franklin (2021).

⁵⁴ Sheppard (2001, p.540) « Quantitative geography is a form of 'science and technology': geography as science combined with all the technologies incorporated into its practices (most prominently, spatial statistical analysis), [...]».

recenti, soprattutto in risposta all'ascesa dei *big data*. Questa rinnovata attenzione rende giustizia alla scienza spaziale, rivalutandone i risultati [*achievements*]]⁵⁵, come afferma Michiel van Meeteren in un articolo pubblicato per il venticinquesimo anniversario del volume di David Livingstone, significativamente intitolato *Statistics Do Sweat: Situated Messiness and Spatial Science*, in cui evidenzia anche il valore pratico e la portata emotiva della *spatial science*, in contrasto con le posizioni di David Livingstone, secondo il quale *Statistics Don't Bleed* poiché si tratta di una geografia arida, priva di vita ed emozioni.

Anche in questo caso si ha la sensazione che il termine *geospatial science* possa contribuire a quello svecchiamento e rivalutazione del termine suggerito da Michiel van Meeteren, ma anche che esso contenga un tentativo di inglobare la *spatial science* nell'orbita del GIS, perché, ancora una volta, il prefisso *geo* sancisce – come già detto – l'affiliazione non con la geografia ma col GIS, e non si potrà fare *geospatial science* al di fuori del GIS.

In questo gruppo di nomi compare poi *spatial information science*⁵⁶, posta, come precisato in apertura della rivista *GeoInformatica*⁵⁷, all'intersezione tra le scienze informatiche e gli studi spaziali. Per questo suo legame con l'informatica, essa ricorda molto il mestiere del fabbricante di attrezzi, se non del GIS, dell'analisi spaziale.

A cavaliere tra il gruppo delle *scienze* e quello dei dati si trova la *spatial data science*, di cui riportiamo la definizione data da Luc Anselin, secondo cui essa «è un sottoinsieme della *data science* che considera le caratteristiche speciali dei dati spaziali nei metodi analitici e negli strumenti *software*. Il suo emergere fu stimolato dall'avvento dei *big data*, molti dei quali sono georeferenziati (e corredati di riferimento temporale). Ciò richiese un ripensamento dei tradizionali metodi analitici spaziali e geocomputazionali, degli algoritmi e degli strumenti *software* per immagazzinare, richiamare, esplorare, analizzare,

⁵⁵ van Meeteren (2019, p. 455): «Moreover, there is resurgent interest in practices from the spatial science era (Johnston *et Alii* 2019; Wylie 2009), particularly prompted by the rise of big data (van Meeteren and Poorthuis, 2018). This renewed attention vindicates re-evaluating spatial science's achievements».

⁵⁶ Che appare anche nella forma *geospatial information science*.

⁵⁷ «Located at the confluence of two rapidly advancing domains, GeoInformatica provides an effective forum for disseminating original and fundamental research and experience in the expanding uses of computer science for spatial studies». (www.springer.com/journal/10707).

visualizzare e imparare da tali dati»⁵⁸. Sul fatto che si tratti davvero di una scienza sono stati espressi dei dubbi⁵⁹, in ogni caso è importante notare come proprio questo nome sia uno dei pochi in cui compare l'aggettivo *geographic*.

Dati. Oltre a *spatial data science* e *geographic data science*, di cui si è detto, vi è ancora un nome centrato sui dati, *spatial data analytics*, che volutamente non abbiamo incluso nel gruppo dei nomi centrati su *analysis*, perché *analytics* ha una connotazione diversa. Dal suono moderno e tecnologico, quasi paradossalmente il nome evoca aspetti tecnologici più che analitici. L'origine del termine non è accademica, ed indica una serie di analisi di tipo statistico, definite in modo piuttosto generico. Wikipedia in inglese fornisce una definizione di *analytics* di cui non troviamo una corrispondenza esatta nel corrispondente sito in italiano: «*Analytics* è la scoperta, identificazione e comunicazione di dinamiche sensate nei dati e il loro utilizzo per migliorare l'efficacia dei processi decisionali. In altri termini, *analytics* può essere inteso come il tessuto connettivo tra i dati e l'efficacia delle decisioni, all'interno di un'organizzazione. Particolarmente prezioso in aree ricche di informazioni registrate, l'*analytics* poggia sull'applicazione simultanea di statistica, programmazione informatica e ricerca operativa per quantificare i risultati»⁶⁰.

Secondo il *blog* di Geospatial World, «*geospatial data analytics* è la direzione del futuro. Le imprese dipendono dal *data analytics* per una gamma di servizi e prodotti che aiutano l'azienda a raggiungere i propri obiettivi più velocemente e a soddisfare i propri clienti in tempi

⁵⁸ Anselin (2020, p. 1): «Spatial data science is a subset of data science that takes into account the special characteristics of spatial data in analytical methods and software tools. Its emergence was stimulated by the advent of big data, much of which is georeferenced (and time stamped). This required a rethinking of traditional spatial analytical and geocomputation methods, algorithms, and software tools to store, retrieve, explore, analyze, visualize, and learn from such data».

⁵⁹ Si veda in particolare l'articolo di Scheider *et Alii* (2020) dal titolo «*Why Geographic Data Science is not a Science*».

⁶⁰ «Analytics is the discovery, interpretation, and communication of meaningful patterns in data; and the process of applying those patterns towards effective decision making. In other words, analytics can be understood as the connective tissue between data and effective decision making, within an organization. Especially valuable in areas rich with recorded information, analytics relies on the simultaneous application of statistics, computer programming and operations research to quantify performance». (<https://en.wikipedia.org/wiki/Analytics>). Ricerca effettuata il 20/04/21.

brevi. Quando il *data analytics* convenzionale viene potenziato con localizzazione e geospaziale [sic], risponde olisticamente alle esigenze del cliente. Le tecnologie geospaziali migliorano l'utilità, la funzionalità e la rilevanza del *data analytics*»⁶¹. Non si può fare a meno di notare la potenzialità di *big business* che traspare da queste osservazioni.

La letteratura accademica relativa allo *spatial analytics* si trova per lo più negli atti di conferenze di carattere ingegneristico e informatico⁶². Nella letteratura tradizionalmente legata all'analisi spaziale il termine è stato finora quasi assente, seppure con qualche eccezione degna di nota⁶³. Soltanto in periodi molto recenti sono comparsi articoli che affrontano l'argomento anche da angolature teoriche e critiche⁶⁴. Tra questi ricordiamo brevemente⁶⁵ Chris Brunsdon e Alexis Comber che s'interrogano sulle proprietà statistiche, ma anche la rilevanza delle analitiche nelle scienze sociali⁶⁶. Dello stesso tenore è l'argomentazione di Alan Murray, pur centrata sulla significatività statistica, secondo il quale «la difficoltà è invariabilmente quella di stabilire che i risultati ottenuti abbiano un senso e forniscano illuminazioni [*insights*] e conoscenza»⁶⁷. È invece degno di nota che proprio Alan Murray sembra usare il termine *spatial analytics* come sinonimo di

⁶¹ «Data analytics is one of the hottest trends in enterprises today. Companies are dependent on data analytics for a wide array of services and features which help the company achieve its goals faster and deliver utility to its customer in a timely manner. When conventional data analytics is powered with location and geospatial, it holistically takes care of the customer requirements. Geospatial technologies enhance the utility, functionality as well as the relevance of data analytics». www.geospatialworld.net/blogs/geospatial-data-analytics-the-way-forward-in-analytics/. Visitato il 31/03/2021.

⁶² Come ad esempio IEEE; si vedano Klein *et Alii* (2015), Chen *et Alii* (2014) e Pandey *et Alii* (2018).

⁶³ L'eccezione più notevole è uno studioso del calibro di Alan Murray, che da anni utilizza questo termine come sinonimo di analisi spaziale, si vedano: Murray (2018); Feng e Murray (2018); Murray e Grubestic (2019).

⁶⁴ Murray (2020); Brunsdon e Comber (2020).

⁶⁵ Su questi argomenti si tornerà nei Capitoli 5 e 6.

⁶⁶ Brunsdon e Comber (2020, p. 89): «We focus on a set of challenges underlying the collection and analysis of big data. In particular, we consider 1) inference [...] 2) the need for analyses that answer questions of practical significance [...] by adopting critical approaches to spatial data science».

⁶⁷ Murray (2020, p. 1): «The challenge is invariably establishing meaning in obtained findings, providing insights and knowledge».

*spatial analysis*⁶⁸. Nonostante questa autorevole eccezione, concludiamo che *spatial data analytics* rimanga distinto da *spatial (data) analysis*.

Neologismi. Un breve discorso a parte meritano i termini che non rientrano né nel gruppo *analisi* né nel gruppo *scienza*: la geostatistica (*geostatistics*) e il geocalcolo (*geocomputation*). Il neologismo *geostatistics* risale al lavoro di George Matheron (1963)⁶⁹ e definisce una branca molto specifica della statistica spaziale, sviluppata all'interno della ricerca mineraria. Tra i concetti fondamentali della geostatistica sono le *variabili regionalizzate*, il *variogramma* e il *kriging*. Il concetto di *variabile regionalizzata*, mutuato dalla lingua francese, è molto vicino a quello di processo stocastico spaziale, mentre il *variogramma*, o *semivariogramma*, è una funzione analoga all'autocorrelazione spaziale; il *kriging* invece è una metodologia intesa a stimare la quantità di minerale presente nel sottosuolo sulla base di un campionamento parsimonioso, utilizzato anche per varie forme di interpolazione spaziale⁷⁰. Anche la geostatistica poggia sulla dipendenza spaziale, e dalle proprie origini in geologia mutua il suffisso *geo*. Gli strumenti della geostatistica sono molto simili a quelli dell'analisi statistica spaziale⁷¹,

⁶⁸ Murray (2020, p. 1): «Anselin (2020) characterized spatial data science as seeking to extract meaningful information from data involving location, proximity, and interaction, requiring specialized expertise to appropriately deal with the complexities of geography».

⁶⁹ Matheron fondò il Centre de Géostatistique et de Morphologie Mathématique all'Ecole de Mines di Parigi a Fontainebleau.

⁷⁰ Per le definizioni di processo stocastico spaziale e di autocorrelazione spaziale si vedano le definizioni fornite nel capitolo 2.

⁷¹ Cressie (1989, p. 197): «We have not yet been able to escape the three-dimensional world in which we live, nor the unidirectional flow of time through which we live. The notion that data close together, in time or space, are likely to be correlated (i.e., cannot be modeled as independent) is natural and has been used successfully by statisticians to model the processes generating the data. Pure temporal models, or time series models as they have come to be known, are usually based on identically distributed errors that are dependent and occur at equally spaced time points [e.g., the autoregressive moving average models of Box and Jenkins (1970)]. The equal spacing and unidirectional flow of time underlie the construction of these models. Spatial models are a more recent addition to the statistics literature. Geology, soil science, crop science, forestry, astronomy, or simply any discipline that works with data collected from different spatial locations needs to develop (not necessarily statistical) models that indicate when there is dependence between measurements at different locations. The models need to be more flexible than their temporal counterparts; for example, it is not reasonable to assume that spatial locations of data occur regularly, and “past”, “present”, and “future” make way for (possibly anisotropic) dependence in a multitude of directions. [...] Geostatistics is usually concerned with spatial prediction, but there are other important areas

e possiamo quindi concludere che la geostatistica coincida quasi perfettamente solo con una porzione dell'analisi spaziale.

L'altro neologismo, *geocomputation*, che potremmo rendere come "geocalcolo", compare sul finire del secolo scorso e, come suggerisce il nome, è centrato sulla potenza di calcolo applicato ai dati spaziali, anzi, geografici, o geospaziali⁷². Fino ad oggi la *geocomputation* si è dedicata soprattutto allo sviluppo di metodi radicati nell'intelligenza artificiale (*artificial intelligence - AI*) e le reti neurali (*neural networks*), oggi spesso riassunti con il termine *machine learning*, e di crescente popolarità. La fede nella possibilità di risolvere i problemi spaziali grazie alla potenza di calcolo, implicita nell'idea stessa di *geocomputation*, non è priva di elementi critici nei confronti dell'analisi statistica spaziale, che non sarebbe preparata ad affrontare problemi numerici di vaste proporzioni⁷³. Tuttavia, l'affermazione secondo cui «il problema è essenzialmente il non aver, per lungo tempo, sviluppato strumenti e stili di analisi e modellizzazione dei dati geografici che siano appropriati per i dati geografici – benché con poche eccezioni» echeggia quelle di Stuart Fotheringham e colleghi sulle passate carenze e future potenzialità della geografia quantitativa nello sviluppare metodi analitici propriamente spaziali⁷⁴. La *geocomputation*, dunque,

(such as stationary-distribution estimation, effect of aggregation, and spatial design) that offer fruitful open problems».

⁷² Si veda in proposito il capitolo di Openshaw e Alvanides (1999).

⁷³ Openshaw e Alvanides (1999, pp. 267-268): «One might argue that it is scandalous that so many key databases are not currently being properly analysed [...]. Yet perhaps the users cannot be blamed for not using tools that are unavailable! The problem is essentially a longstanding failure to evolve distinctly geographical-data-appropriate tools and styles of analysis and modelling – albeit with a small number of exceptions. [...]. In essence, no amount of apparent statistical sophistication should be allowed to hide the fact that much of spatial statistics is very limited in what it can do, and is even more limiting in its view of spatial information and the handling of their special properties».

⁷⁴ Fotheringham, Brunson, Charlton (2000, pp. XI-XII): «Perhaps the most important role of this book is to provide examples of recent research in quantitative geography where the emphasis has been on the development of techniques explicitly for *spatial* data analysis. What makes the methods of modern quantitative geography different from many of their predecessors is that they have been developed with the recognition that spatial data have unique properties and that these properties make the use of methods borrowed from aspatial disciplines highly questionable. As such, this book acknowledges what the authors see as a turning point in the development of quantitative geography. It is written at a period when quantitative geography has reached a stage of maturity in which its practitioners are no longer primarily importers of other disciplines' techniques but are mainly exporters of novel ideas about the analysis of spatial data».

ha solo qualcosa in comune con l'analisi spaziale, perché fondamentalmente diverso è l'approccio delle due all'analisi statistica dei dati. A conclusione di questa breve rassegna di nomi notiamo che l'elenco presentato in Tab. 4.1 ne costituisce ovviamente solo uno dei possibili ordinamenti. In un approccio assai quantitativo all'argomento, abbiamo contato quante volte compaiono nella tabella sia gli aggettivi elencati nelle caselle di sinistra, sia i sostantivi elencati in quelle di destra. Memori, tuttavia, della maggiore popolarità del GIS rispetto all'analisi quantitativa, abbiamo scelto di visualizzare questo conteggio⁷⁵, rappresentando le rispettive presenze tramite una nuvola di parole (*word cloud*)⁷⁶ in Fig. 4.2.

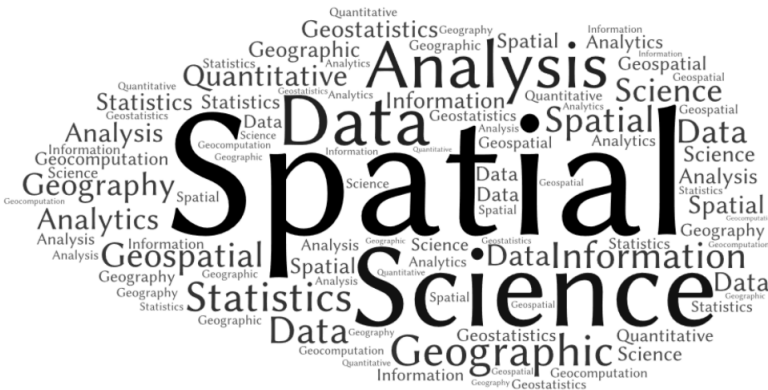


Fig. 4.2 - Terminologia comunemente usata per definire l'analisi spaziale.

⁷⁵ Tra gli aggettivi, *quantitative* compare una volta, *geographic/geographical* 2 volte, *geo-spatial* 2 volte e *spatial* 8 volte; tra i sostantivi, *geography* compare una volta, *information* 2 volte, *statistics/statistical* 2 volte, *data* 4 volte, *science* 6 volte, *analysis* 5 volte e *analytics* una volta (se invece contiamo assieme *analysis* e *analytics* arriviamo a 6 presenze, tante quante *science*). È importante notare che il prefisso *geo-* compare complessivamente 8 volte (5 volte in aggettivi e 3 in sostantivi). Se quindi lo avessimo immesso come termine separato (ad esempio *geo-spatial* e *geo-computation*) esso avrebbe assunto quasi le stesse dimensioni di *spatial* nell'immagine, poiché questa operazione avrebbe anche portato a 10 il numero di presenze di *spatial*.

⁷⁶ Secondo la definizione dell'Oxford Dictionary (lexico online): «A word cloud is an image composed of words used in a particular text or subject, in which the size of each word indicates its frequency or importance». www.lexico.com/definition/word_cloud. I 16 nomi sono stati inseriti nel software online *WordArt* (<https://wordart.com/>), dove, oltre alla lista di parole, si sono definiti solo la forma del diagramma, il *font*, e la direzione delle parole.

L'immagine, una semplice visualizzazione, evidenzia come il termine che più spicca è *spatial*, seguito da *science*, quindi da *analysis e data*, mentre tutte le altre parole compaiono con dimensioni assai più modeste. Le considerazioni qui esposte si limitano alla frequenza di alcuni termini identificati in letteratura, ma trascurano molti altri fattori, tra cui la frequenza con cui quei vocaboli sono utilizzati. Pur con tutti i limiti di questa analisi, osservando il diagramma viene spontaneo chiedersi se ciò che stiamo dibattendo altro non è che la naturale evoluzione di quella che fu la vecchia *Spatial Science*. Tuttavia, la storia è complicata e non è così semplice per la materia riconoscersi direttamente in quel nome. Forse ciò cui stiamo assistendo è una tensione tra il riconoscersi nel vecchio nome e l'affermare la propria diversità da ciò che esso un tempo definì, pur nella continuità di quella tradizione.

4.7. Il cammino dell'analisi spaziale

Alla luce delle sfaccettate definizioni e della molteplice terminologia esaminati in questo capitolo, vorremmo tentare una conclusione su cosa sia e dove si collochi l'*analisi spaziale*. Siamo partiti da una definizione che la vedeva all'intersezione delle discipline, più consolidate e meglio delimitate, da cui trae origine. È questa la visione tradizionale, corrispondente alle posizioni espresse da Stuart Fotheringham e Alan Murray tra il 2000 e il 2010⁷⁷, e rappresentata schematicamente in Fig. 4.3.

La situazione attuale, invece, appare più complessa, articolata in una molteplicità di discipline e comunità di pratica (*communities of practice*), per cui l'analisi spaziale sembra trovarsi all'intersezione di numerose tra queste "entità", come mostrato in Fig. 4.4⁷⁸.

L'emergere di comunità di pratica sempre nuove e il proliferare di nomi diversi in riferimento all'analisi di dati spaziali suggerisce certo che la materia sia dinamica e in rapida evoluzione, ma anche, e soprattutto, che essa stia suscitando un notevole interesse, non solo nella

⁷⁷ Si fa riferimento ai già citati articoli di Murray (2010) e Fotheringham, Brunson, Charlton (2000) relativi alla geografia quantitativa.

⁷⁸ Lo schema proposto sia nella Fig. 4.3 che nella Fig. 4.4 intende solo rappresentare sovrapposizioni e intersezioni, senza specificarne la porzione.

geografia, sia all'interno che all'esterno della comunità accademica. Si tratta di un fermento di non poco conto, quello che già alcuni anni orsono aveva fatto supporre ad alcuni che una nuova rivoluzione quantitativa stesse bussando alla porta⁷⁹.

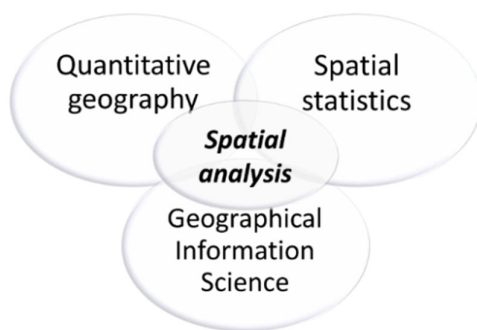


Fig. 4.3 - L'analisi spaziale all'intersezione delle discipline da cui ha avuto origine.

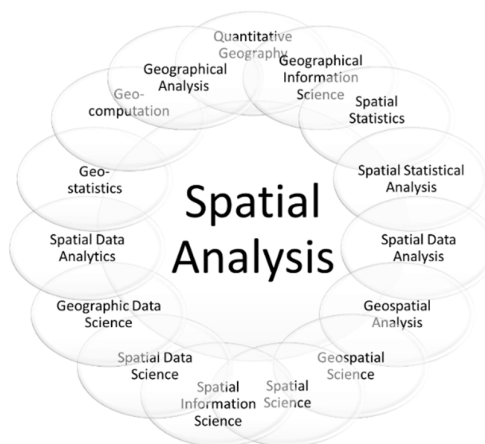


Fig. 4.4 - L'analisi spaziale all'intersezione delle nuove discipline e comunità.

⁷⁹ Cresswell (2014, p. 54): «The big data phenomena appear to provide opportunities for a new quantitative revolution due to its focus on the individual level and simultaneous geocoding of everything. We need to be cautious about embracing this process due to the many ethical and political dilemmas that are implicated».

Realisticamente, molto di questo entusiasmo deriva non solo dalla crescente disponibilità di dati e di tecnologia, ma anche, riteniamo, da altri due fattori. Da un lato crediamo ci sia il sentore, la lusinga, che l'analisi, o meglio le analitiche (*analytics*), dei dati spaziali, possano celare il prossimo *big business*, pronto a sbocciare dagli *analytics* che poggiano non solo sulla disponibilità, ma anche sulla raccolta e il commercio dei dati spaziali sul cui valore commerciale ed impatto politico e sociale rimangono ormai pochi dubbi. È quindi probabile che questa rivoluzione, trainata dai dati geolocalizzati, apra delle prospettive all'analisi spaziale e alla geografia. La questione più importante sarebbe comprendere di quali prospettive si tratti, e per questo sarebbe importante imparare dal proprio passato, osservare e comprendere come il *big business* abbia modellato la vicenda del GIS, per non trovarsi impreparati ad affrontare questa nuova realtà.

Il secondo fattore crediamo poggi sulla consapevolezza delle potenzialità dell'analisi spaziale o geografica nel comprendere ed interpretare i dati geolocalizzati e, in questo modo, nel fornire un contributo fondamentale alla comprensione dei grandi problemi del nostro tempo: dal cambiamento climatico alle migrazioni, alle pandemie. In questa fase storica, come vedremo nei prossimi capitoli, la geografia ha in buona parte superato le tensioni tra fautori e oppositori del quantitativismo ed è quindi più preparata ad accogliere senza pregiudizi le illuminazioni [*insights*] che l'analisi quantitativa dei dati geografici può consentire di ottenere.

Di fronte a questi segnali positivi è tuttavia importante non sottovalutare le considerazioni espresse da Stuart Fotheringham e colleghi rispetto alla riluttanza di molti (geografi) ad abbracciare una disciplina ritenuta difficile, ostica, e la sensazione, espressa con forza da Stan Openshaw, che la materia sia meno piacevole e affascinante di quanto viene offerto dal GIS⁸⁰. Auspichiamo che l'uso di nomi diversi da quelli consolidati non celi un tentativo di mascherare le difficoltà tradizionalmente percepite nell'analisi statistica dei dati spaziali, allo scopo di rendere la materia più accattivante. È invece importante comprendere che per fare *buona* analisi spaziale bisogna prendersi la briga di impararla e svolgerla [come si deve], affrontando difficoltà ben maggiori di quelle richieste per la semplice visualizzazione e analisi

⁸⁰ Fotheringham, Brunson, Charlton (2000); Openshaw (1998).

elementare dei dati spaziali. Se l'analisi dei dati geolocalizzati serba delle promesse per il futuro, è importante anche preparare dei professionisti che siano in grado di svolgere questa *buona* analisi spaziale, e per formare queste persone serve, ma non basta, insegnare loro la matematica e la statistica, oltre che la geografia. Come, ancora una volta, ci insegna l'esperienza del GIS, la sopravvivenza della disciplina richiede anche il consolidamento delle sue basi concettuali, filosofiche o, con un aggettivo solo, scientifiche.

Forse non è un caso che di fronte al cosiddetto diluvio di dati (*data deluge*), spaziali e non, si assista anche ad una proliferazione di termini, vecchi e nuovi, che contengono l'idea di scienza. Condividendo le perplessità circa l'accoppiata terminologica "*data science*", troviamo tuttavia incoraggianti i contenuti più meditati e le spinte ad una riflessione critica che emergono dai lavori più recenti. Proprio questa riflessione potrebbe traghettare l'odierna analisi spaziale oltre l'entusiasmo suscitato dalla grande disponibilità di dati e dagli sviluppi della tecnologia. In fondo, è l'analisi spaziale che si occupa di analisi complesse, di modelli con una loro ontologia, epistemologia, con implicazioni etiche e sorrette da presupposti concettuali e teorici. È sempre quell'analisi spaziale che, criticata per lo scarso spessore teorico, seppe rispondere alle critiche avviando un serio discorso su se stessa.

Il potenziale *big business* dell'analisi dei dati spaziali è allettante, e si potrà trarre profitto da analisi e *analitiche*, ma senza dimenticare l'esperienza del GIS, né gli errori che ne avrebbero decretato la morte accademica, se non avesse saputo trasformarsi in *scienza*.

Forse, dopo una lunga vicenda e tanto fermento, torneremo a chiamarla *scienza spaziale*, o forse ne svecchieremo il nome in *scienza geospaziale*, e forse sapremo sentire, e far sentire, in quel prefisso il sapore della geografia, la vera casa e la famiglia dell'analisi spaziale.

5. La critica e le sfide del presente e del futuro

“Well, in our country,” said Alice, still panting a little, “you’d generally get to somewhere else - if you run very fast for a long time, as we’ve been doing.”

“A slow sort of country!” said the Queen. “Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!”

Carroll, *Through the Looking Glass*, 2010, p. 17.

5.1. Premessa

Sin dall’ultimo decennio del secolo scorso l’analisi spaziale, assieme all’intera *GIScience*, ha intrapreso un discorso autocritico che ne segna una nuova, importante fase di crescita e maturazione. In larga misura questo discorso procede in risposta alle critiche mosse al GIS e alla geografia quantitativa negli anni 1990. È quindi l’intera *GIScience* ad affrontare costruttivamente il discorso, internalizzando le critiche e rispondendovi in maniera creativa. Di conseguenza, questo stesso fenomeno segna una nuova fase della *GIScience* e della geografia quantitativa consentendo anche il superamento di posizioni tradizionalmente antagonistiche e il distanziamento dall’eccessivo rigore delle vecchie posizioni quantitative. La *GIScience*, e con essa l’analisi spaziale, si apre così a un dialogo proficuo con altre geografie, si affaccia su discorsi e istanze sociali, superando il paradigma positivista che l’ha tradizionalmente contraddistinta.

In questo percorso s’innestano prepotentemente anche alcuni fenomeni recenti: il diluvio di dati, che forse potrebbe essere il volano per una rinascita del quantitativismo in geografia, ma anche una serie di fenomeni geospaziali di grande attualità che richiedono non solo analisi quantitativa competente e trasparente, ma anche consapevolezza delle conseguenze di tali analisi sull’ambiente e sulla società. Ci riferiamo al cambiamento climatico, alle migrazioni, alla pandemia da COVID-19, che hanno portato prepotentemente alla ribalta l’analisi e la rappresentazione dei dati geografici. Probabilmente saranno questi, più che il diluvio di dati, a costringere la geografia quantitativa, il GIS

e l'analisi spaziale, ad affrontare sempre più consapevolmente quei discorsi spesso trascurati, dall'etica all'autocritica, che da lungo tempo attendevano una risposta.

In questa particolare fase storica si coglie, infatti, una tensione in seno alla quantitativa e della *GIScience*, che è proteso da un lato verso la necessità di una profonda riflessione interna, e dall'altro verso l'urgenza di rispondere ai problemi concreti sul tappeto: gli stimoli dei dati, la tecnologia assieme ai grandi fenomeni in atto. Si tratta forse di due tendenze disgiunte, che ancora debbono integrarsi in un discorso armonico: da un lato vi è l'apertura al dialogo e la riflessione sui propri paradigmi e metodi, dall'altro vi è l'esigenza di sviluppare tecniche e analisi, che spaziano dall'ormai tradizionale *geocomputation* alle più moderne *analytics*, forse meno mature sul piano metodologico ed epistemologico, ma più che mai appetibili e redditizie.

Questa apparente dicotomia lascia intravedere un'ulteriore questione: il coinvolgimento profondo, la consapevolezza di sé, dei propri metodi e del proprio oggetto da parte di chi *fa* analisi spaziale. Tra lo sviluppo delle tecniche e la riflessione sui contenuti sta il lavoro concreto svolto ogni giorno da chi conduce analisi spaziale, ed è importante comprendere come avvenga questo lavoro. Il vecchio dibattito sul GIS tra scienza, attrezzo e fabbricante di attrezzi evidenziò come al geografo non potesse essere chiesto di fabbricare i propri attrezzi, perché quelle competenze andavano oltre quelle del geografo. Diverso appare invece il discorso odierno dell'analisi spaziale che, proprio dal dibattito critico che emerge dai *big data*, chiede all'analista spaziale di saper costruire i propri attrezzi e di saper controllare dall'interno i meccanismi che guidano l'analisi.

Si tratta di una questione di grandissima importanza, che riporta l'intero processo analitico all'interno della disciplina, affermando la centralità di chi conduce quotidianamente l'analisi. Per comprendere la profondità della questione, in questo capitolo ci si sofferma anche sugli aspetti concreti, sulle caratteristiche degli attrezzi dell'analisi spaziale, e sull'evoluzione più recente degli strumenti, dal *software* commerciale per l'analisi spaziale elementare alla compilazione dei comandi (*scripting*) per l'analisi spaziale statistica. Il discorso sul GIS non può rimanere disgiunto da quello dell'analisi spaziale, nell'intreccio costante che ancora una volta mostra come *colui che* fa analisi spaziale non è *colui che* fa GIS. Il cerchio, che conduce dallo sviluppo di

tecniche e che conduce all'autocritica ed all'esecuzione consapevole dell'analisi, si chiude imponendo un grande fardello a chi *fa* analisi spaziale, che deve essere, insieme, geografo, statistico, critico, tecnico e programmatore.

5.2. Dalla critica all'autocritica

La moderna analisi spaziale, come i suoi precursori, la geografia teoretico-quantitativa e il GIS, è da sempre impegnata nella ricerca e nella realizzazione di soluzioni applicative per quelli che sono percepiti come i problemi importanti che richiedono un'immediata soluzione. E i problemi da risolvere, si sa, non mancano mai. Ci sono poi gli sviluppi della tecnologia, sempre più veloci e incalzanti, da cui una disciplina applicativa non può prescindere, proprio perché in quella veloce tecnologia – e non altrove – si trovano le soluzioni ai problemi pressanti. Infine, ci sono altri fenomeni, intrecciati ai precedenti, che, per le medesime ragioni, la disciplina non può ignorare: tra i più eclatanti, l'esplosione dei dati, che, come si è visto, è un fenomeno ormai vecchio di decenni, ma che ultimamente ha assunto addirittura le proporzioni di un diluvio.

L'analisi spaziale sembra talvolta trovarsi, come la piccola Alice di Lewis Carroll, nel regno della Regina Rossa, dove bisogna correre a perdifiato soltanto per rimanere nello stesso luogo, ma se si vuole anche andare da qualche parte, bisogna correre ad una velocità doppia. Eppure, proprio come Alice, nessuno di noi può correre ad una velocità doppia rispetto al massimo delle proprie capacità. E così l'analisi spaziale, come già il GIS, ha a lungo tralasciato quelle questioni, pure importanti, che spaziano dall'etica, ai metodi e all'epistemologia che pure la sorreggono, impiegando tutte le proprie energie nelle soluzioni applicative.

In larga parte questa vicenda non riguarda solo l'analisi spaziale, ma, quasi in un unico intreccio, anche il GIS e la geografia quantitativa. La spinta, si potrebbe dire l'obbligo, ad iniziare un discorso su se stessa provenne dalle pulsioni critiche di altre geografie e il discorso critico fu avviato relativamente tardi, quando certe caratteristiche si erano orai cristallizzate. Ci si riferisce in particolare alla frammentazione, alla mancanza di coesione e di autocoscienza, che fa da

contraltare all'enfasi dell'analisi su applicazioni anche molto disparate. Per questo le mancò non solo una seria riflessione su se stessa, ma anche sui propri presupposti, metodi e protagonisti. Tutto questo contribuisce anche a spiegare la scarsa chiarezza, all'interno della disciplina stessa, sulla propria tradizione, i propri presupposti teorici, i propri obiettivi.

Il dibattito critico prese le mosse proprio in un momento di particolare intensità e fermento per il GIS: quell'inizio degli anni '90 in cui si assisteva alla prima esplosione di dati¹, usciva il primo autorevole testo che finalmente dava forma alla materia GIS² e, contemporaneamente, veniva espressa per la prima volta l'idea di una scienza dell'informazione geografica³. Forse anche in relazione a tutti questi fenomeni, che davano ampia visibilità all'informazione e analisi geografica, il dibattito, iniziato al di fuori delle comunità dalla geografia quantitativa e dal GIS, ebbe ampia risonanza assumendo talvolta anche toni accesi⁴, benché toni non meno accesi risuonassero anche all'interno di quelle comunità⁵. Tuttavia, negli anni successivi, questa condizione iniziò a cambiare, allorché anche la geografia quantitativa ha iniziato a compiere uno sforzo per rispondere costruttivamente e in prima persona alle questioni poste da voci critiche nella geografia,

¹ Su questo si veda il lavoro, scritto pochi anni dopo, di Openshaw e Alvanides (1999).

² Si tratta dei due volumi "the big books" curati da Maguire, Goodchild, Rhind (1991). Mentre prima della loro uscita circolava in rete il "Core Curriculum" curato dall'NCGIA: un corposo insieme di lezioni tenute ciascuna da un esperto, che si potevano legittimamente stampare e, come era allora la prassi, rilegate in grosse spirali per essere lette e studiate. Uscito nel 1990, il primo Core Curriculum era una serie di 75 lezioni, suddivise in 3 cicli, destinate a costituire una sorta di libro di testo, di fatto il primo libro di testo sul GIS, di cui allora non esistevano altri esemplari, per i corsi universitari in GIS, ai rispettivi livelli. Ciascuna lezione, o serie di lezioni, affrontava un argomento diverso, ed era stata curata da uno specialista di quell'argomento; ciascuna era corredata da esempi pratici, immagini da proiettare in classe, e domande d'esame. L'intera serie poteva essere acquistata su floppy disc per un importo molto modesto, ed ottenne un enorme successo: fu adottata da numerose istituzioni e tradotta in molte lingue. Oggi il Core Curriculum, ormai datato, è disponibile su Internet e lo si può scaricare gratuitamente dalla pagina dell'NCGIA (National Center for Geographic Information and Analysis): www.ncgia.ucsb.edu/pubs/core.html. Si vedano in proposito Kemp e Goodchild (1991) e Goodchild e Kemp (1992).

³ Il riferimento è all'ormai noto articolo di Goodchild (1992), intitolato appunto *Geographical Information Science*.

⁴ Del dibattito sulla geografia teoretico-quantitativa si è detto; quello più recente sul GIS si fa risalire al testo di Pickles (1995a) titolato appunto *Ground Truth*, senza con questo tralasciare altri protagonisti del dibattito critico, come Smith (1992).

⁵ Si veda fra tutti Openshaw (1991).

affrontando dall'interno un discorso sui propri presupposti e metodi⁶. Anche in questo caso il discorso autocritico dell'analisi spaziale e della geografia quantitativa non è disgiunto da quello del GIS, anzi, anche in questo caso quello sul GIS tende a produrre e ricevere più risonanza, proprio a causa della maggiore visibilità di quest'ultimo, dentro e fuori l'accademia.

5.2.1. *Critical GIScience*

Il discorso sui fondamenti e presupposti del GIS assunse visibilità in seguito alle posizioni espresse verso la metà degli anni '90 da alcuni geografi critici⁷. Tuttavia, il dibattito ha radici precedenti, e secondo la cronologia proposta da Nadine Schuurman (2000), prese le mosse da un editoriale di Peter Taylor del 1990⁸. L'autrice delinea tre fasi del dibattito, che si snodano a partire dai primi anni '90⁹. La prima fase, che definisce dell'animosità, «fu contraddistinta da aperta ostilità disciplinare, ciò nonostante, vi parteciparono attivamente anche un certo numero di ricercatori GIS e ne emerse una qualche forma di dialogo. Alla metà del decennio, tuttavia, le file dei critici erano cresciute, mentre quelle dei difensori del GIS si erano assottigliate. Solo pochi studiosi di GIS accolsero positivamente quella che poteva sembrare un'intrusione e seppero comprendere che quelle critiche effettivamente puntavano il dito su carenze fondamentali del GIS. Altri invece si infastidirono perché quelle posizioni critiche manifestavano anche

⁶ Si vedano, tra i primi, Sheppard (2005) e Schuurman (2006).

⁷ Tra questi si ricordano in particolare Pickles (1995b) e Smith (1992).

⁸ Taylor, 1990.

⁹ Come nota Schuurman (2000, p. 570): «I have organized these critiques and their responses from GIS practitioners into three waves or periods, each distinguished by shifts in strategies and motivation on the part of critics. The first wave, from 1990 to 1994, was distinguished by the animosity of debate. Despite overt disciplinary hostility, a number of GIS researchers did engage in these discussions and a dialog of sorts resulted. By the middle of the decade, however, the ranks of critics had increased while defenses from GIS had tapered off. A few GIS scholars welcomed the intervention and felt the critiques expressed critical shortcomings of GIS (Poiker, pers. interview, 1997; Harvey and Chrisman, 1998). Others were disturbed by the clear lack of understanding of GIS exhibited by critics (Openshaw, 1991; Clarke, pers. interview, 1998; Marble, pers. interview, 1998). At the same time, a number of critics and GIS researchers had begun to work together on NCGIA Initiative 19. Critics were drawn into the upper echelons of GIS research, and the present era of increased co-operation between critics and the GIS community was initiated».

chiaramente un difetto, o almeno una superficialità di comprensione del GIS. Proprio in quel periodo numerosi furono sia i critici che i fautori del GIS che avevano iniziato a lavorare all'iniziativa 19 dell'NCGIA. I critici furono trasportati nelle alte sfere della ricerca GIS, ed ebbe inizio l'era corrente, di maggiore co-operazione tra i critici e la comunità del GIS».

L'iniziativa 19 dell'NCGIA fu avviata nel marzo del 1996, sotto la direzione di Trevor Harris e Daniel Weiner, il suo «obiettivo principale era proprio quello di sviluppare e sostenere ricerche in cui sia i costruttori e applicatori del GIS (*developers and practitioners*), sia i critici e teorizzatori sociali, interessati alla natura dell'uso del GIS, potessero lavorare insieme. È in questo contesto che l'iniziativa si propone tre ampi temi di ricerca che riguardano la produzione e l'uso del GIS: l'amministrazione e il controllo delle popolazioni; i conflitti locali che coinvolgono popolazioni svantaggiate e l'ecologia politica relativa all'accesso e all'uso delle risorse naturali. Gli scopi centrali dell'iniziativa sono di: esaminare come la disponibilità di dati e le tecniche di visualizzazione influenzano i modi in cui le risorse naturali e la società sono rappresentati in GIS; esaminare quali limiti alla rappresentazione possono essere intrinseci alla logica del GIS; determinare come le rappresentazioni dell'ambiente e della società nel GIS influenzino le questioni poste e le soluzioni proposte nelle applicazioni pratiche; determinare se e come la conoscenza, le opinioni (*views*) e le necessità di coloro che sono impattati (*affected*) dall'applicazione del GIS possono essere rappresentate adeguatamente in situazioni sociali conflittuali dove il GIS è usato come strumento decisionale (*decision-making tool*); esaminare fino a che punto le nuove funzionalità del GIS possono consentire di allargare i limiti delle rappresentazioni correnti; investigare il grado in cui le applicazioni del GIS possono essere democratizzate mettendo la tecnologia nelle mani di uno spettro più ampio della società; e investigare [...] le implicazioni etiche e legali delle attività collegate»¹⁰.

¹⁰ «Initiative 19: The Social Implications of How People, Space, and Environment are Represented in GIS. [Co-leaders: Trevor Harris and Daniel Weiner (both of West Virginia University) - begun March, 1996]. The primary objective of this initiative is to develop and support research in which GIS developers and practitioners on the one hand, and critics and social theorists concerned with the nature of GIS use on the other hand, can work together. It is in this context that this initiative is concerned with three broad research themes involving

Si tratta evidentemente di un'agenda fitta e importante, con obiettivi ambiziosi e non facili da raggiungere. I contributi degli ultimi due decenni si snodano lungo diverse linee, che in parte cerchiamo di ripercorrere nei paragrafi successivi. L'oggetto del *critical GIS* rimane difficile da definire e disseminato in diversi filoni di ricerca. Ancora secondo la Schuurman, in un interessante articolo uscito nel 2009, il *critical GIS (CGIS)*, sarebbe «un approccio relativo alla valutazione di ciò che la tecnologia GIS trae (*draws*) da molteplici cassette degli attrezzi (*tool-kits*) intellettuali – dalla geografia, dalla teoria sociale e dalla scienza computazionale. Mentre le sue radici risiedono, dagli anni '90 in poi, nelle battaglie tra geografi umani e *GIScientisti*. Il CGIS è emerso come un approccio indipendente e costruttivo per migliorare il potere e il fascino del GIS. Il CGIS sta anche cominciando a guadagnare accettazione come componente legittima della grande tenda (*sic*) che è la *GIScience*». Proprio la grande varietà dei suoi contenuti rende difficile caratterizzare più precisamente il CGIS, che si occupa di femminismo, GIS partecipativo, marxismo ed epistemologia. «Tuttavia, c'è un implicito intendimento che quelle questioni disparate ricadono entro la stessa categoria. Il filo invisibile che lega tali questioni costituisce il *critical GIS*»¹¹.

GIS production and use. These are: the administration and control of populations; locational conflict involving disadvantaged populations; and, the political ecology of natural resource access and use. The central goals of the initiative are to: examine how data availability and visualization techniques influence the ways in which natural resources and society are represented in GIS; examine what limits to representation may be intrinsic to the logic of GIS; determine how the representations of environment and society in GIS influence the questions posed, and solutions proposed in practical applications; determine whether and how the knowledge, views, and needs of those affected by the application of GIS can be represented adequately in conflictual social situations where GIS is used as a decision-making tool; examine to what degree new functionalities of GIS may allow the limits of current representations to be extended; investigate the degree to which the application of GIS can be democratized by placing the technology in the hands of a broader spectrum of society; and, investigate (drawing on Initiative 16: Law, Public Policy and Spatial Data Bases) the ethical and legal implications of related activities». www.ncgia.ucsb.edu/research/initiatives.html (visitato il 5/7/2021).

¹¹ Le due citazioni sono tratte da Schuurman (2009, p. 139): «Critical GIS (CGIS) is an approach to evaluating GIS technology that draws upon multiple intellectual tool kits – from geography, social theory and computing science. While its roots are in the battles between human geographers and GIScientists in the 1990s, CGIS has emerged as an independent, constructive approach to enhancing the power and appeal of GIS. CGIS is also beginning to gain acceptance as a legitimate component of the broad tent that is GIScience. [...] It is difficult to characterize a group as diverse as critical GIScientists. A typical critical GIS session at a major

A distanza di dieci anni *The Canadian Geographer* (diretto dalla stessa Schuurman) dedicò un numero speciale al dibattito sul *Critical GIS*. Nell'editoriale che apre il fascicolo, Jim Thatcher, Luke Bergmann e David O'Sullivan, ragionando su vent'anni di *critical GIS*, ancora non ne stabiliscono con certezza la definizione e l'oggetto: «sia esso un approccio teorico, un'ideologia, una sub-disciplina, di una strategia di ricerca, di una tribù, o di una spina nel fianco della scienza seria»¹². Gli autori propendono però per una visione ottimistica, che poggia sulla vivacità del dibattito recente e dall'emergere di collaborazioni inter- e intra-disciplinari con alleati che spaziano dalle *digital humanities* alla *critical physical geography*. L'immagine di *critical GIS* che emerge da diversi punti di vista, e da articoli di vario tipo è quella di un campo (*field*) che si è chiaramente portato ben oltre la scrivania (*sic*) (benché non sia mai stato entro un ambito così ristretto), mentre continua a cimentarsi con questioni perduranti sulla "criticalità", sul ruolo degli educatori in questo ambiente tecno-sociale, e su come portarsi oltre la critica astratta verso una *prassi critica*.

Nel lungo dibattito sul *critical GIS*, come nella "grande tenda" della *GIScience* è molto difficile però isolare le riflessioni di pertinenza più diretta dell'analisi spaziale, che, ci pare, si possono ricondurre a pochi, ma ricchi e articolati, temi ricorrenti che, sostanzialmente, riguardano gli strumenti e le problematiche relative ai dati. Nei paragrafi seguenti riprendiamo questi aspetti, articolandoli, soggettivamente, in una sequenza che comprende *spatial thinking*, *critical mapping*, *geoetica*, *FOSS*, e si conclude con le questioni statistiche relative ai dati.

conference might include papers on topics as diverse as feminism, participatory GIS, Marxism and epistemology (Kwan 2002b; Crampton 2005; Harvey *et Alii* 2005; Sheppard 2005; Elwood 2006; Sullivan 2006; Ghose 2007). Yet somehow there is an implicit understanding that these disparate issues fall under the same category. The invisible thread that links these issues constitutes critical GIS».

¹² Thatcher, Bergmann, O'Sullivan (2018, p. 4): «Critical GIS, whether seen as a theoretical approach, ideology, sub-discipline, research strategy, tribe, or thorn in the side of serious science, is now over 20 years old (Haklay2012). [...] Our optimism resides in the recent vibrancy of the field (Thatcher *et Alii* 2016) and in emerging inter-, intra-, and sub-disciplinary collaborations with allies from the digital humanities (Bodenhamer *et Alii* 2010; Knowles *et Alii* 2015) to critical physical geography (see Lave *et Alii* 2014; Lane 2017). The picture of critical GIS that emerges from these viewpoints, research papers, and review essay is of a field that has clearly moved well beyond the desktop (although it was never so narrowly bound; see, among others, Obermeyer 1995; Warren 2004; Kaplan 2006) - while continuing to grapple with enduring questions of criticality, the role of educators in this techno-social setting, and how to move beyond critique to critical praxis».

5.2.2. *Critical spatial data science, critical spatial data analytics*

Nella letteratura più recente prende piede una riflessione sulla cosiddetta “economia politica del GIS” che, partendo da un’insofferenza crescente per il monopolio commerciale di Esri, si estende al discorso, molto importante per l’analisi spaziale, sul cosiddetto *free and open source software*, o *FOSS*, per arrivare fino a considerare gli aspetti legati alle decisioni, gestione e pianificazione, e il suo rapporto con i produttori di *software*, aprendosi anche ai problemi legati ad un insegnamento critico del GIS¹³. Le argomentazioni di Chris Brunsdon e Alexis Comber fanno riferimento ad una *critical spatial data science*, e ad una *critical spatial analytics*¹⁴: su questa terminologia si è discusso molto nel capitolo 4, e nonostante l’oscillazione terminologica che si osserva in questa letteratura recente, le questioni affrontate spingono il discorso critico al cuore dell’analisi spaziale.

La moderna analisi spaziale, come del resto l’intera *GIScience*, viene svolta, sia dentro che fuori l’accademia, per mezzo di *software*, da cui dipende la gran parte del lavoro che essa svolge. Tutto questo lavoro, soprattutto pratico e applicativo, ma anche teorico e di ricerca, si trova quindi a dipendere, in larga misura, dai produttori di *software*, le cui decisioni, tra cui scelte di produzione e politiche di prezzo, ne condizionano la stessa esistenza, modalità e risultati. Come si è già osservato nel capitolo 1, il mercato del *software* GIS è di fatto un monopolio, dominato da Esri¹⁵, e molto spesso università e grosse organizzazioni hanno contratti che regolano l’uso della gamma di prodotti, primi fra tutti ArcGIS e ArcGIS pro¹⁶. In contrasto si osserva, come si è visto, la crescita in parallelo degli *open data* assieme alle pubblicazioni *open access*, al *software* gratuito e *open source* (*free open source software* - *FOSS*). Elementi questi che, pur non avendo costi monetari diretti, scontano, generalmente, una maggiore difficoltà di apprendimento, una gamma di funzioni non sempre all’altezza di quella del prodotto commerciale, e, soprattutto nel caso dell’analisi spaziale, la

¹³ Thatcher, Bergmann, O’Sullivan (2018).

¹⁴ Brunsdon e Comber (2020; 2021).

¹⁵ www.esri.com/en-us/home

¹⁶ ArcGIS comprende le versioni desktop (www.esri.com/en-us/arcgis/products/arcgis-desktop/overview), online (www.esri.com/index.html) e professional – pro (www.esri.com/en-us/arcgis/products/arcgis-pro/overview).

necessità di apprendere linguaggi di coding diversi. Questi *software* generalmente non forniscono interfacce particolarmente *friendly*, richiedendo invece all'utente la capacità di programmare (*coding* o *scripting*) le varie funzioni. Oltre a ciò, occorre anche considerare gli aspetti economici, politici, e di potere connessi a chi controlla il prodotto e trae profitto dalla sua vendita.

Oltre a tutto questo occorre tener presente il tema caro a Chris Brunson¹⁷ relativo alla riproducibilità dell'analisi spaziale. Per loro sarebbe importante che l'analisi fosse pubblicata in maniera trasparente, ossia corredata di appendici costituite dai dati e dal programma (o codice) utilizzati. Ciò consentirebbe a chiunque di *criticare* l'analisi, verificandone la qualità, e potenzialmente suggerendo modifiche, elementi particolarmente importanti se i risultati di quell'analisi supportano decisioni di gestione e pianificazione del territorio che hanno un impatto su una popolazione o su segmenti vulnerabili di popolazione¹⁸. Tutto ciò ovviamente dipende dalla volontà degli autori, o dall'imposizione da parte delle riviste, di pubblicare dati e codici (*code*, o programma). Se le analisi sono svolte utilizzando un *software* commerciale, solo chi ha acquistato il prodotto ha la possibilità di riprodurlo, e anche questa stessa possibilità potrebbe essere di breve durata, nel caso – non raro – in cui il venditore decidesse di modificare la funzione descritta nella pubblicazione. I due autori, quindi, propongono un modello veramente trasparente di conduzione e pubblicazione dell'analisi spaziale, nella direzione che essi stessi definiscono “*open spatial analysis*”.

Come già visto nel capitolo 1, la comunità dell'analisi spaziale da lungo tempo si è orientata all'uso di *FOSS*, in particolare R ed RStudio¹⁹, il cui uso consente applicazioni ‘su misura’ delle funzioni analitiche. La capacità di programmare (*coding*), che tende a contraddistinguere la comunità dell'analisi spaziale, denota una consapevolezza e conoscenza degli strumenti (attrezzi) di molto superiore a quella di chi si limita a “premere tasti automaticamente”, costituendo una

¹⁷ Si veda, ad esempio, Brunson (2016).

¹⁸ Brunson e Comber (2021): «In essence, a reproducible research philosophy is one which allows all aspects of the *answer* generated by any given analysis to be tested. This is especially the case as our lives are increasingly lived and shaped by a vast amount of data – often termed *Big Data* (upper case intentional)».

¹⁹ Per R si veda: www.r-project.org/ e per RStudio: www.rstudio.com/

condizione necessaria per un approccio critico all'analisi. Riproducibilità, trasparenza, e applicazione ad *hoc* dell'analisi fanno parte di quell'approccio critico che Michael Goodchild e Donald Janelle (2010) definiscono *critical spatial thinking*.

Chris Brunsdon e Alexis Comber sostengono inoltre che la cosiddetta rivoluzione dei *big data* ha fatto aumentare particolarmente i dati spaziali, ciò che pone l'analisi (e gli analisti) spaziali sulla ribalta, forse anche loro malgrado. In effetti, la grande quantità di dati fa sì che le analisi si ripercuotano su un numero di persone molto superiore al passato, impattandone persino la stessa vita, ciò che, in sé, ne richiede un'attenzione critica. A questo si aggiunge la necessità di considerare in maniera critica scala e *MAUP*, rivisitati in chiave dei *big data*, soprattutto in considerazione del [mutato] rapporto tra campione e intera popolazione²⁰ assieme alle conseguenze sulla definizione e valutazione degli intervalli di confidenza nei test statistici inferenziali. Queste ultime considerazioni²¹ sono elaborate dai medesimi autori anche nel loro articolo su *JOSIS*²², in cui affermano che una *critical spatial data analysis* è quella in cui chi la svolge è consapevole delle limitazioni sia dei dati che del modo in cui i risultati delle analisi e i modelli vengono impiegati. Gli autori giungono ad evocare una sorta di giuramento ippocratico in cui si dichiara di essere consapevoli delle

²⁰ Si tratta della dubbia ipotesi che Brunsdon e Comber (2021), sintetizzano nella formula $n=all$, ovvero il campione coincide con l'intera popolazione.

²¹ Su questo si tornerà più avanti in questo capitolo.

²² Brunsdon e Comber (2020, p. 91): «A critical data analysis is one in which the practitioners are aware of both the limitations of the data and the way that analysis results and models are deployed. As well as considering techniques of data analysis, many argue (e.g., [14]) for critical reflection on the assumptions often adopted by the media and technical literature, that challenge the notion of big data as objective, and somehow 'better' [13]. This has led to the field of critical data studies [8] which argues for the need 'to explore the ways in which [data] are never simply neutral, objective, independent, raw representations of the world, but are situated, contingent, relational, contextual' [15, p. 5]. It emphasises the need to think about how data analyses are deployed — the cooking of data in collection and analysis — and to consider 'the technological, political, social and economic apparatuses and elements that constitutes and frames the generation, circulation and deployment of data' [13, p. 2]. We should 'think about big data science in terms of the common good and social contexts' [13, p. 5]). This criticality can be extended to embrace critical views of the more technical aspects of the data analysis».

conseguenze dell'uso dei modelli che si sviluppano, con particolare attenzione alle delicate questioni legate all'inferenza²³.

5.2.3. Critical spatial thinking e critical mapping

Lo *spatial thinking*, il pensare in maniera spaziale, non disgiunto da *spatial reasoning* e *spatial intelligence*, è un atteggiamento mentale che si attua sia nella ricerca, teorica o applicata, sia nella vita quotidiana. Secondo Michael Goodchild e Donald Janelle (2010), un modo per definire il pensare spaziale critico (*critical spatial thinking*) è in riferimento ai processi mentali che accompagnano l'uso delle tecnologie e dei dati spaziali. Il *critical spatial thinking* è in netto contrasto con il premere tasti meccanicamente, ed implica che i processi di manipolazione, analisi, *mining*, e modellizzazione dei dati provochino e richiedano un pensare critico riferito a concetti relativamente profondi quali accuratezza della scala, incertezza, ontologia, rappresentazione, complessità, proiezione ed etica²⁴. Come si vede, il pensiero spaziale critico è intriso di molti elementi della *critical GIScience*, ma anche specifici dell'analisi spaziale. L'articolo di Michael Goodchild e Donald Janelle, infatti, continua con una approfondita discussione nei confronti di una selezione dei più rappresentativi concetti spaziali che richiedono un pensare critico: non a caso la scelta ricade su eterogeneità e dipendenza spaziale, scala, inferenza su dati spaziali²⁵.

²³ Brunson e Comber (2020, p. 92): «And they advocate taking on board Derman's Hippocratic Oath of Modeling [9], in particular the declaration "I understand that my work may have enormous effects on society and the economy, many of them beyond my comprehension". One could easily add 'the environment' to 'society and the economy' in that statement. A typical concern may then be the wider implications of misinterpreting the inferential aspects of a particular analysis».

²⁴ Goodchild e Janelle (2010, p. 9): «One way to define critical spatial thinking is in relation to the use of spatial tools and data — as the mental processes that accompany the use of these technologies. Critical spatial thinking is in sharp contrast to rote button-pushing, and implies that the processes of data manipulation, analysis, data mining, and modeling provoke and require critical thinking, about such comparatively profound issues as scale accuracy, uncertainty, ontology, representation, complexity, projection, and ethics».

²⁵ Goodchild e Janelle (2010, pp. 9-10): «The remainder of this section discusses a selection of these spatial concepts, focusing on three that are of an advanced nature, and are typically acquired during senior undergraduate or graduate education, if ever. We discuss them here as examples of the concepts that are needed to underpin the critical spatial thinking skills that we might expect of spatially aware scholars».

L'approccio critico all'analisi spaziale e alla *GIScience* non può prescindere da un accenno, seppur breve, alla cartografia, o quantomeno al *mapping*, che spesso è il tramite per comunicare i risultati analitici, ma è anche un passaggio obbligato per l'analista, che, senza la visualizzazione, faticherebbe a comprendere l'analisi medesima e i suoi risultati. Per quanto riguarda la cartografia critica si rimanda all'ormai classico saggio di David Harley *Deconstructing the Map*²⁶, e poi fino al dibattito contemporaneo, che si arricchisce di contributi intesi a dar voce a quanti, tradizionalmente, lungi dal costruire la cartografia, l'hanno invece subita. Le *contro-mappe* producono così un discorso che è non solo critico, ma anche in aperto contrasto con le cartografie precedenti e ufficiali, spesso nelle mani del "potere costituito": si tratta delle cosiddette "3 C" della *contro-cartografia critica* (*Critical Counter-Cartography*)²⁷.

Tra i contributi più recenti ci soffermiamo su quello di Matthew Wilson²⁸, che si pone a cavaliere tra il critico GIS e la cartografia critica. Le sue *New Lines*, infatti, si riferiscono alle linee della mappa (*the Trouble of the Map*), le linee lungo le quali si sviluppa il *critical GIS*, e le linee che traccia chi, facendo GIS, produce mappe. Questione fondamentale rimane quella dell'identità tra ciò che si fa e ciò che si studia, perché non sempre chi fa il mestiere di criticare [sia l'analisi che la mappa] fa anche il mestiere di fare [analisi e mappa]. Perciò non sempre ciò che si fa è anche ciò che si studia. La disciplina del *critical GIS* è relativamente giovane, e forse per questo la sua breve storia è stata complessivamente pacifica e serena, di fatto un superamento costruttivo delle contrapposizioni, anche aspre, che hanno contraddistinto il confronto tra fautori e critici del GIS e del quantitativismo sul finire del secolo scorso. Si tratta di un confronto complesso, i cui personaggi si muovono su un palcoscenico in continua trasformazione, lungo linee che continuano a cambiare. Attraverso la pluralità di significati della parola *linee* (*lines*), il discorso di Wilson s'incentra sulla mappa, la carta geografica, cui si riferisce il provocatorio sottotitolo *The Trouble of the Map* (il guaio della carta geografica), che ci costringe a riflettere sul perché la mappa sia nei guai. Se la mappa sembra

²⁶ Harley (1989).

²⁷ Si veda ad esempio la collezione di critica e contro-cartografia *This is not an Atlas* curata da Halder e Michel (2019).

²⁸ Wilson (2015; 2017).

vivere oggi un periodo di grande visibilità e popolarità, se possibile ulteriormente aumentato dalla pandemia, i guai della mappa si manifestano ad un livello più profondo, ed essa diventa un po' vittima del suo stesso successo.

L'espressione di Matthew Wilson *The Trouble of the Map* può quindi riferirsi anche ai guai causati dalla mappa, quelli in cui rischia di incorrere sia chi fa, sia chi legge, chi usa la mappa, rincorrendo il significato delle sue linee, che è ad un tempo profondo ed effimero. La mappa, ci ricorda Wilson, ha un fascino e una capacità di attrazione quasi sensuale, il potere e la forza di piegarci su quelle linee irresistibili. È questo il guaio della mappa, la presunzione di affermare una verità che è molto più complessa delle linee che pretendono di dirla. Nei guai è, allora, chi non sa avvicinarsi alla mappa con le dovute conoscenze tecniche e capacità critiche, sia che la voglia fare o solo guardare. Ai guai della mappa si aggiunge anche il guaio che ormai troppi, non sapendo né teoria né critica e spesso nemmeno le tecniche, si improvvisano mappatori – non cartografi: dilettanti ingenui con l'illusione che la sola tecnologia li renda capaci di fare. E nei guai sono quanti, attratti dal fascino della mappa, inciampano nelle linee confuse dei mappatori della domenica o cadono nelle maglie formate dalle linee tracciate da mappatori in malafede. In un mondo in cui neppure la verità si distingue da ciò che è falso (*fake*) e ciò che è post-verità, occorrono davvero nuove e mappe e bussole per trovare la rotta, che altro non è se non la linea immaginata che sola ci può salvare dal naufragio in un mare di guai.

5.3. L'analisi spaziale e la sfida dei dati

Di fronte alla crescente proliferazione di dati spaziali si è addirittura ipotizzato, già da qualche anno, e seppure con toni critici, che una seconda rivoluzione quantitativa stia bussando alla porta della geografia²⁹: un'ipotesi affascinante ma tutta da verificare.

²⁹ Ci riferiamo al titolo dell'articolo di Wyly (2014) "The New Quantitative Revolution", che si conclude con l'affermazione (p. 36): «We have a new quantitative revolution, and this time the most dangerous revolutionaries are not people brandishing mathematical models while dreaming up provocative titles about the Augean stables (Gould, 1977) or declaring 'Never question the assumptions!' (Hanson, 1993: 555) and 'By our theories you shall know

La proliferazione di dati è un fenomeno complesso che si è sviluppato al di fuori dell'ambito geografico. Tuttavia, come si è visto, la gran parte di questi dati sono riferiti a precisi luoghi geografici, e sono (o possono essere) georeferenziati, ed è questa la ragione per cui la presunta seconda rivoluzione quantitativa potrebbe interessare la geografia, forse più di altre discipline. La quantità enorme di dati archiviati, siano essi captati dai satelliti, generati dalle telecomunicazioni o prodotti dai social media, offrirebbe grandi opportunità analitiche, imponendo, nel contempo, questioni nuove, che attraversano diversi piani. Il *piano metodologico*: come analizzare quantità vastissime di dati; quale sia la rappresentatività di quei dati rispetto alla popolazione cui appartengono; quali siano le loro proprietà statistiche e come esse possano essere verificate. Il *piano etico*: come archiviare, analizzare, sintetizzare e rappresentare tali dati, nel rispetto di chi li ha prodotti, forse anche inconsciamente, e della loro *privacy*, e nel rispetto di chi invece da tale produzione è rimasto fuori. Il *piano ontologico*: quale sia la sostanza dei dati e la loro relazione con il mondo che si pretende rappresentino. Il *piano epistemologico*: quale conoscenza del mondo si possa trarre da questi dati.

Le tecnologie che trattano questi *big data* sono, per loro natura, pervasive, invasive e gravide di conseguenze solo in parte note o prevedibili. Senza neppure considerarne le implicazioni sociali, politiche, economiche, non si può prescindere dalle questioni etiche, dai paradigmi e le teorie che dovrebbero puntellare l'analisi dei dati e che dovrebbero guidare l'interpretazione delle analisi e definire la loro relazione con il mondo che quei dati pretendono di rappresentare.

In questo contesto, non potendo occuparci di tutti i problemi posti dai dati e dalla loro abbondanza, si sceglie di occuparci dei due aspetti, che investono più direttamente l'analisi spaziale: quello dell'etica o, come si vedrà, *geoetica*, e quello delle loro proprietà stocastiche e del loro trattamento con metodi statistici.

us' (Harvey, 1969: 486). Now, the models arrive brandishing people – collections of thousands and millions of socially networked digital individuals in an expanding neoliberal noosphere». Interessante è comunque l'intero volume di *Dialogues in Human Geography* 4, no. 1 (2014), che raccoglie un importante dibattito originato dalla discussione dei contenuti dei programmi di geografia nel Regno Unito.

5.3.1. La geoetica

Vorrei aprire questo paragrafo presentando alcune esperienze personali che illustrano le questioni etiche legate all'analisi spaziale.

All'università di Calgary come, credo, nella gran parte degli atenei, qualsiasi ricerca che utilizzi dati relativi a soggetti umani deve ottenere l'approvazione di una commissione etica prima di poter iniziare³⁰. Commissioni diverse afferiscono alle facoltà sociali e umanistiche da un lato e a quelle mediche dall'altro, come diverse sono le restrizioni all'uso dei dati e alla pubblicazione dei risultati. Per molti anni ho fatto parte della commissione socio-umanistica in rappresentanza del dipartimento di geografia, e in tale veste ho esaminato e valutato numerose proposte di ricerca e più volte ho sottoposto proposte di ricerca al vaglio di entrambe le commissioni. Da entrambe le esperienze posso dedurre che le commissioni svolgono un lavoro serio e rigoroso.

Alcuni anni fa collaboravo con colleghi nella facoltà di medicina ad un progetto sull'incidenza del morbo di Chron nei bambini: una malattia rara in quella fascia d'età. La ricerca aveva ovviamente ottenuto l'approvazione della commissione etica di pertinenza. Uno studente del Master in GIS ebbe il compito di svolgere un'analisi dei *clusters* e di mapparne i risultati. Quando vidi quelle mappe gli impedii di pubblicarle, perché la rappresentazione cartografica rivelava con precisione spaventosa l'indirizzo di residenza di quei bambini, consentendo facilmente di risalire alla loro identità.

Quando mio figlio era in età prescolare, in ben due casi negai il consenso all'utilizzo dei suoi dati in progetti di ricerca che pure ritenevo molto rilevanti³¹. In entrambi i casi, ciò che mi indusse a negare il consenso fu la dimensione delle unità spaziali di indagine: si trattava di unità relativamente piccole³² soprattutto in certi quartieri; inoltre,

³⁰ Si veda: <https://research.ucalgary.ca/conduct-research/additional-resources/iriss-human-ethics-and-animal-care> (visitato il 18/06/2021).

³¹ Nel primo caso la ricerca indagava sui benefici della frequenza di un anno preliminare alla prima elementare (*kindergarten*) sul rendimento scolastico negli anni successivi. Nel secondo caso si indagavano i potenziali danni sulla salute dentale della sospensione dell'immissione di fluoro nella rete idrica pubblica. Entrambe le ricerche avevano ottenuto l'approvazione etica, tuttavia la prima l'aveva ottenuta da un'altra università.

³² Le unità spaziali in questione sono i *postal codes*, definiti per il Canada e analoghi agli *ZIP codes* degli Stati Uniti. Per la definizione si veda: www.canadapost-

nella mia zona di residenza l'età media della popolazione è piuttosto elevata, per cui sapevo che se il dato conteneva anno di nascita e sesso sarebbe stato molto facile identificare il bambino. Nel primo caso mi spinsi a parlarne con la dirigenza della scuola materna, cercando di spiegare che la ricerca metteva a repentaglio anche la privacy di bambini i cui genitori non fossero a conoscenza degli aspetti etici in questione, ma temo di non essere stata compresa.

Fin qui le mie esperienze personali, ma le potenziali questioni non si esauriscono qui. Quando utilizziamo dati satellitari, ci siamo mai chiesti chi li ha finanziati? Per quale scopo sono stati prodotti? Quali tecnologie sono state sviluppate ed impiegate per ottenerli? Chi trae beneficio e chi danno dal loro commercio? Domande analoghe, anche se di segno opposto, sorgono con riferimento agli *open data*: la domanda rimane sempre *cui prosit*? Banalmente queste domande somigliano a quelle relative al cosiddetto "consumo critico", tuttavia quando i dati vengono impiegati a fini di ricerca, i risultati che si ottengono dalla loro analisi vengono poi pubblicati e riecheggiano in un discorso scientifico o di gestione territoriale: il peso di queste domande è quindi notevolmente maggiore di quando ci si limita ad indossare un indumento o a consumare una bevanda.

Come abbiamo ormai più volte ripetuto, sulla scia di colleghi più autorevoli, "spaziale è speciale". Anche le questioni relative all'etica dell'acquisizione, gestione, analisi, protezione, pubblicazione dei dati spaziali necessitano di definizioni speciali, che si occupino di scala, aggregazione, dipendenza spaziale, cartografia. Eppure, come tendono a mostrare gli esempi, le norme etiche cui si fa comunemente riferimento nella ricerca, per l'utilizzo dei dati, la loro protezione, e quelle sulla protezione della proprietà intellettuale provengono da un approccio alla ricerca che possiamo chiamare tradizionale, nel senso di non-spaziale: la gestione di dati con metodi statistici a-spaziali e la loro rappresentazione in forma di tabelle e narrazioni testuali, senza l'uso di strumenti cartografici. Si ha la sensazione che, nonostante la proliferazione, con lo sviluppo del GIS, di tecnologie analitiche geospaziali e cartografiche sempre più accessibili, le loro implicazioni etiche abbiano ricevuto relativamente poca attenzione. Non è facile reperire,

postescanada.ca/cpc/en/support/articles/addressing-guidelines/postal-codes.page? (visitato il 19/06/2021).

nella letteratura accademica, riferimenti espliciti all'etica nell'uso e nella rappresentazione dei dati geospaziali in GIS, e i contenuti sono relativamente scarsi anche nella cosiddetta letteratura grigia³³. Una notevole eccezione a questa affermazione è l'articolo, datato, di Jeremy Crampton³⁴, che però sembra aver trovato ben poco seguito nella letteratura successiva. Analogamente, la questione della proprietà intellettuale in cartografia pare trovare al massimo brevi spazi nei manuali di cartografia, anche in questo caso con rare, anche se notevoli eccezioni³⁵. Almeno fino a poco tempo fa.

Il termine *geoethics* è emerso recentemente nel contesto delle scienze geologiche (*geosciences*), e si riferisce a diversi aspetti del rapporto tra le geoscienze e la società³⁶. Si tratta di un termine che, per il momento, sembra presente più nella letteratura grigia che in quella propriamente accademica, e non esattamente in quella geografica o geospaziale. Analogamente, riferimenti a norme di condotta compaiono, anche questi di rado, sui siti web di associazioni a cavallo tra il professionale e l'accademico. Ci riferiamo in particolare ad URISA (Urban and Regional Information Systems Association), il cui sito presenta un codice etico in GIS approvato dall'associazione nel 2003³⁷ e articolato in quattro aspetti etici della pratica del GIS: 1) gli obblighi nei confronti della società; 2) gli obblighi nei confronti dei datori di lavoro e dei finanziatori; 3) gli obblighi nei confronti dei colleghi e della corporazione (*profession*), 4) gli obblighi nei confronti degli individui nella società.

Di particolare interesse, invece, è l'iniziativa che l'AAG (American Association of Geographers) ha lanciato a partire dalla primavera del 2021: l'organizzazione di una serie di seminari (*webinars*) in materia di geoetica, dedicati sia a temi generali che specifici, e aperti anche ai non-membri dell'AAG³⁸. Si tratta di un'iniziativa importante, che comprende, tra gli altri argomenti, una discussione esplicita sull'etica

³³ In inglese *grey literature*, ovvero pagine web, relazioni di convegni, note tecniche, etc.

³⁴ Crampton (1995).

³⁵ Ci riferiamo alla trattazione, relativamente esaustiva, del copyright nel manuale di Kraak e Ormeling (2009, pp. 182-187).

³⁶ Si veda ad esempio la pagina dedicata dall'Università di Carleton (Canada) all'insegnamento della geoetica: https://serc.carleton.edu/geoethics/what_geoethics.html (visitato il 18/06/2021).

³⁷ Si veda: www.urisa.org/about-us/gis-code-of-ethics/ (visitato il 18/06/2021).

³⁸ Si veda: <https://aag-geoethics-series.secure-platform.com/a> (visitato il 18/06/2021).

nell'*analisi* geospaziale dei dati. Probabilmente è la prima volta, dopo l'articolo di Cramptom nel 1995, che le questioni etiche legate all'uso e alla rappresentazione dei dati geospaziali vengono affrontate organicamente, portando il termine *geoetica* all'interno della geografia e del GIS. L'iniziativa è in corso mentre scriviamo e, come insegnano esperienze passate, è prevedibile che produrrà o, quantomeno, stimolerà non solo un dibattito, ma anche letteratura scientifica. Al momento, l'iniziativa comprende, tra l'altro, la stesura di un documento, in fase di preparazione, fatto circolare tra i partecipanti ai seminari.

Perché ora? La sensazione, non troppo velata, è che questa recente attenzione agli aspetti etici dell'uso e rappresentazione dei dati geospaziali abbia la propria ragion d'essere in un fenomeno, probabilmente non previsto, che ha seguito, fin dalle sue prime fasi, la pandemia da COVID-19. Forse per la prima volta nella storia, in seguito alla pandemia è stata raccolta una quantità senza precedenti di dati epidemiologici relativi ad incidenza, mortalità, diagnostiche (tamponi), e utilizzo delle strutture sanitarie (ricoveri e occupazione delle terapie intensive). Questi dati, in quantità enorme e ad elevata risoluzione spaziale e temporale, sono stati quotidianamente raccolti, organizzati, e in molti casi pubblicati gratuitamente in forma di *open data*. In Italia, ad esempio, i dati sono pubblicati a partire da febbraio 2020 per ogni giorno a livello regionale e provinciale dalla Protezione Civile, e più recentemente l'ISTAT ha iniziato la pubblicazione, a livello comunale e cadenza mensile, dei dati di mortalità durante la pandemia, rapportati a quelli dei 5 anni precedenti³⁹. Oltre che *open data*, ossia accessibili pubblicamente e gratuitamente, tutti questi dati sono corredati di files GIS (*shape files*) che ne permettono l'attribuzione, analisi e rappresentazione per le rispettive unità geografiche. Come l'Italia, molti altri Paesi pubblicano i propri dati con analoghe modalità e risoluzione spazio-temporale.

Questa fioritura di dati consente a molti, che lo desiderino, di accedere ai dati epidemiologici e, anche semplicemente utilizzando tecnologie come la Web2.0, di analizzarli e rappresentarli, costruendo un proprio discorso sull'andamento dell'epidemia. Come tutti hanno

³⁹ I diversi dati epidemiologici sono pubblicati a livello regionale, mentre a livello provinciale sono pubblicati solo quelli sul numero quotidiano di casi, come si può vedere in: <https://github.com/pcm-dpc/COVID-19> (visitato il 19/06/2021). Per i dati ISTAT si veda: www.istat.it/it/archivio/240401 (visitato il 18/06/2021).

potuto vedere, la disponibilità dei dati ha condotto ad un’analoga fioritura di cartografia, mappe e *dashboards* che quotidianamente hanno informato il pubblico, lungo le diverse fasi della pandemia, utilizzando anche la comunicazione cartografica⁴⁰, utilizzando tecnologia GIS.

La pandemia, e la grandissima disponibilità di dati geospaziali, assieme all’interesse o alla necessità di comunicare un fenomeno d’interesse globale, ha così prepotentemente portato alla ribalta l’analisi e la rappresentazione cartografica dei dati geospaziali, sollevando questioni che spaziano dal controllo politico sulla mappa come strumento di comunicazione, fino agli effetti sui viaggi e i flussi turistici indotti, più o meno volutamente, da rappresentazioni cartografiche grossolane, aggregate o approssimative della situazione epidemiologica in diverse regioni e momenti della pandemia.

In tutto questo non si possono trascurare gli aspetti geoetici dell’utilizzo di dati relativi alle comunicazioni personali, in particolare quelli delle comunicazioni telefoniche, la cui collezione e utilizzo sono stati caldeggiati in diverse fasi della pandemia da numerosi governi. Dal punto di vista dell’analisi spaziale vorremmo concludere questa breve panoramica osservando che i dati epidemiologici discussi in queste righe possono ancora essere trattati con i metodi tradizionali dell’analisi spaziale, in quanto la dimensione ne consente un’analisi statistica che prenda in considerazione distribuzione e significatività. Quando invece si considerino i dati dedotti dalle comunicazioni telefoniche si entra pienamente nella dimensione dei *big data*, la cui analisi non può avvenire con gli strumenti della statistica spaziale tradizionale, ma deve necessariamente avvalersi di tecniche afferenti ad *analytics*, *geo-computation*, *data mining* o *machine learning*.

5.3.2. *Big data, proprietà stocastiche e statistica spaziale*

Parte delle perplessità sollevate dall’uso dei *big data*, come si è visto anche poco sopra, sono legate alle difficoltà di verificarne le proprietà stocastiche e condurvi analisi statistiche adeguate. Come si vedrà nel prossimo capitolo, i dati spaziali sono generalmente

⁴⁰ Tra i più noti è quello dell’università John Hopkins: <https://coronavirus.jhu.edu/map.html> (visitato il 19/06/2021).

caratterizzati da varie forme di incertezza, che tende a propagarsi attraverso le operazioni analitiche spaziali⁴¹. Non vi è ragione per escludere che queste forme di incertezza caratterizzino anche i grandi insiemi di dati, i *big data*. Anzi, è abbastanza intuitivo ritenere che, maggiori sono le quantità di dati, minore sia il controllo qualitativo che su di essi si può effettuare, e che quindi l'incertezza li caratterizzi in maniera analoga, se non superiore, agli insiemi più piccoli di dati.

Si è poi discusso, soprattutto nel capitolo 2, dell'errore, solo in parte legato all'incertezza, che caratterizza non solo i dati spaziali ma anche le operazioni analitiche, e della necessità di effettuare tali analisi con metodi statistici che, appunto, ne tengano conto. Come si è visto, infatti, la statistica spaziale è proprio l'approccio capace di valutare, controllare, e in certi casi minimizzare tale errore. Anche a questo proposito, è evidente che anche i grandi dati sono soggetti all'errore e che quindi necessitano di un apparato statistico adeguato. Tuttavia, come si è visto in questo capitolo e nel capitolo 4, benché le varie *analitiche* e le diverse forme di *machine learning* contengano analisi statistiche, raramente esse si prestano ad una trattazione statistica accurata, in larga misura proprio a causa della dimensione dei dati. Cominciano infatti a sorgere dubbi sull'efficacia ed utilità di queste forme, diciamo limitate, di analisi⁴².

Vi è poi l'aspetto più specifico o, come si è detto, *speciale*, dei dati spaziali. Si è visto in dettaglio nel capitolo 2 come tali dati siano generalmente affetti da dipendenza ed eterogeneità spaziale e, in particolare, come la dipendenza spaziale contraddica l'ipotesi, alla base di ogni statistica, di indipendenza delle osservazioni. È evidente che la dipendenza spaziale caratterizza non solo gli insiemi di dati di dimensioni medie o piccole, ma anche quelli di dimensioni grandi e molto grandi, i *big data*. Anzi, è molto probabile che li caratterizzi con maggiore intensità, poiché la dipendenza spaziale è una proprietà legata alla distanza, ed è tanto maggiore quanto minore è la distanza. Quanto più "grandi" sono gli insiemi dati, tanto più è probabile che le

⁴¹ Il termine qui comprende sia le operazioni analitiche elementari che quelle di analisi statistica spaziale.

⁴² Si veda ad esempio Brunson e Comber (2020, p. 90): «However, as with data analytics and statistics, there is a danger that the emergent discipline ignores some important lessons learned from its predecessors. If GIScience has become a key player here, it is vital that it ensures such lessons are heeded».

osservazioni siano rilevate a distanze relativamente piccole, ciò che tenderà ad acuirne la dipendenza spaziale. È quindi evidente che qualsiasi analisi statistica condotta sui *big data* tende a violare quella presunta indipendenza, e in questo sono comprese, ancora una volta, anche le *analitiche* e le varie forme di *machine learning*.

Vi è poi un aspetto ulteriore, di cui ci siamo occupati nel capitolo 2, ed è il problema dell'inferenza, di crescente attualità in seguito alla pandemia da COVID-19. I cosiddetti effetti spaziali, ossia l'effetto di dipendenza ed eterogeneità spaziale sui risultati analitici è quello di aumentare la varianza dei modelli e con essa l'incertezza delle stime. L'aumento della varianza, come si è visto, ha anche un'importante ripercussione sull'inferenza nei modelli. Aumentando la varianza, infatti, aumenta la probabilità di commettere il cosiddetto errore di tipo I, ossia quello di accettare come vera un'ipotesi falsa. Benché esistano dei metodi statistici per controllare questa circostanza, essi tendono ad agire sulla dipendenza spaziale locale, ossia tra unità spaziali contigue, ma non su quella globale, che quindi continua ad incrementare l'errore di tipo I⁴³. Come è noto vi è un *trade-off* tra errore di tipo I (il rischio di accettare falsi positivi) ed errore di tipo II (il rischio di rifiutare falsi negativi) e la preoccupazione per i falsi negativi determina le soglie dei test inferenziali. Anche questo problema statistico tende ad acuirsi al crescere della dimensione dei dati, e benché rimanga importante minimizzare il rischio di commettere errore di tipo II, la probabilità che ciò incrementi l'errore di tipo I in maniera incontrollata, a causa della grande dimensione dei dati spaziali può essere molto serio.

Si pensi, ancora, alla pandemia da COVID-19: a causa delle complessità delle analisi statistiche, non è facile, o nemmeno possibile, valutare le conseguenze della sovrastima del rischio di incidenza del virus in una certa regione; non è chiaro se/o di quanto sia opportuno eccedere nella sovrastima per evitare possibili sottostime; ancora, non è chiaro quali siano le conseguenze, nel breve e nel lungo periodo, di queste stime. A tutto questo si aggiunge il problema del *MAUP*, per cui, anche ammettendo che ogni Paese utilizzi metodologie statistiche uniformi e le medesime soglie di rischio, non è facile (ammesso che

⁴³ Parte di queste considerazioni sono state espresse da Peter Rogerson nel suo intervento al seminario dell'AAG "*Ethical Spatial Analytics*" (<https://aag-geoethics-series.secure-platform.com/a/solicitations/10/sessiongallery/206>). Si vedano in proposito anche i recenti articoli di Rogerson (2019) e Jordan *et Alii* (2019).

sia possibile) scorporare da questi calcoli l'effetto delle unità spaziali utilizzate da ciascun Paese o regione per le stime. E ciò implica che le stime del rischio in diversi Paesi e regioni non siano confrontabili.

5.3.3. *Data driven, analytics, data science: quale analisi?*

Come già visto nel capitolo 1, l'allontanamento dell'analisi spaziale dall'econometria spaziale (che cercava una verifica empirica di relazioni teoriche) ne ha in parte comportato un graduale allontanamento dalla teoria [economica], trasformandola sempre più in un'analisi *data driven*, ossia un'analisi che cerca di cogliere relazioni empiricamente verificabili, benché non necessariamente supportate da una solida base teorica. Questa transizione segna anche l'inizio di uno slittamento che porterà i dati ad assumere un ruolo sempre più centrale nell'analisi, facilitato dalla crescente disponibilità di dati e di strumenti computazionali per gestirli ed analizzarli.

In questa transizione verso la centralità dei dati si inserisce la comparsa, sul finire del secolo scorso, del termine *geocomputation*⁴⁴ che, come si comprende intuitivamente, s'incentra sulla potenza di calcolo applicato ai dati spaziali⁴⁵. L'affermarsi della *geocomputation*, o “geocalcolo”, fu consentito dai progressi compiuti dalle tecnologie dell'informazione geografica (*geographical information technologies* - GIT) e dal *supercomputing* o *high performance computing* (HPC), che, congiuntamente, consentivano, già negli anni '90, di gestire quantità crescenti di dati spaziali e di effettuare calcoli complessi su queste vaste collezioni di dati. D'altro canto, erano ancora le tecnologie dell'informazione geografica, prima fra tutte proprio il GIS, che consentivano di acquisire e gestire quelle quantità di dati su cui svolgere i calcoli. Già più di vent'anni fa Stan Openshaw e Seraphim Alvanides (1999) prevedevano che la crescente disponibilità sia di potenza computazionale sia di dati geografici avrebbe dovuto, in seguito, stimolare lo sviluppo di metodi analitici radicalmente nuovi⁴⁶.

⁴⁴ Il termine è stato introdotto già nel capitolo 4.6.

⁴⁵ Si veda in proposito Openshaw e Alvanides (1999).

⁴⁶ Openshaw e Alvanides (1999, p. 268): «[This chapter] argues that large-scale computation can now be used as a paradigm for solving some of the major spatial analysis problems that are relevant to GIS. However, if computational power is to be useful, then there also has

Se si può ritenere che effettivamente il GIS abbia consentito di computerizzare molte procedure che prima venivano svolte a mano, per l'analisi spaziale la sfida è assai maggiore, perché proprio quella disponibilità di dati abbinata alla maggiore potenza di calcolo dovrà condurre a porsi domande nuove e quindi a sviluppare metodologie analitiche computazionali nuove che sappiano affrontare quelle nuove domande con gli ultimi e più potenti mezzi a disposizione⁴⁷.

Se dunque l'allontanamento dall'econometria spaziale segnò un distacco dall'obiettivo della verifica empirica di relazioni teoriche, mantenendo tuttavia il rigore metodologico e il legame stretto con la statistica, nell'emergere della *geocomputation* si configura un'importante separazione tra quest'ultima e l'analisi quantitativa statistica spaziale di cui ci siamo occupati finora in questo volume.

Come infatti affermavano già Stan Openshaw e Seraphim Alvanides (1999), fu evidente sin dall'inizio che l'analisi spaziale non era preparata ad affrontare problemi numerici di vaste proporzioni⁴⁸, proprio per la difficoltà, se non addirittura l'impossibilità, di analizzare le proprietà statistiche di quantità enormi di dati. È proprio in questa difficoltà che si comprende la contrapposizione tra la logica dei metodi

to be a clear understanding of what the requirements and the user needs are. This leads us to a brief typology of alternative approaches and a brief illustrative case study based on one method in which high performance computing is combined with GIS data and artificial intelligence (AI) tools to develop better ways of engineering zoning systems as a decision support, analysis, modelling, and data management tool».

⁴⁷ Openshaw e Alvanides (1999, p. 267): «The GIS revolution has created an immense wealth of spatial information in a large number of different application areas [...]. The development of GIS can be largely regarded as the computerisation of pre-existing manual procedures and established technologies that were already fairly mature in research terms [...]. However, when the focus of attention switches to spatial analysis and modelling then it is a very different story. Frequently there have been no existing useful manual procedures to computerise, and over the last 20 years there has been very little relevant new research focused specifically on the special needs of GIS. Moreover, an increasing number of the emerging analysis tasks are novel and have not previously attracted much or any attention – for example, the exploration of very large spatial datasets for completely unknown patterns and relationships».

⁴⁸ Openshaw e Alvanides (1999, pp. 267-268): «One might argue that it is scandalous that so many key databases are not currently being properly analysed [...]. Yet perhaps the users cannot be blamed for not using tools that are unavailable! The problem is essentially a longstanding failure to evolve distinctly geographical-data-appropriate tools and styles of analysis and modelling – albeit with a small number of exceptions [...]. In essence, no amount of apparent statistical sophistication should be allowed to hide the fact that much of spatial statistics is very limited in what it can do, and is even more limiting in its view of spatial information and the handling of their special properties».

geocomputazionali, che analizzano i problemi spaziali basandosi soprattutto sulla capacità di calcolo, e la logica che regge l'analisi *statistica* spaziale.

Come si è discusso nel capitolo 2, infatti, un approccio di tipo statistico all'analisi dei dati poggia proprio sullo studio della loro distribuzione e delle loro proprietà statistiche, da cui l'analisi non prescinde nemmeno quando si applicano statistiche non parametriche. È dunque evidente che l'analisi della distribuzione e delle proprietà statistiche dei dati diventa eccessivamente onerosa (ammesso che sia comunque possibile) quando si vogliono sottoporre quantità molto grandi [*big*] di dati ad analisi quantitative numerose e computazionalmente intense. I *big data* sono grandi in almeno due dimensioni: la numerosità del campione e il numero di variabili. Si ricorderà che il processo stocastico spaziale e la dipendenza spaziale sono stati definiti, nel capitolo 2, in relazione ad una funzione di probabilità congiunta. Analizzare tale funzione per i *big data* sarebbe quanto meno oneroso, se non impossibile, ed è anche per questa ragione che la *geocomputation* applica metodologie che non necessitano di un approccio statistico.

Con i *big data* si impiegano frequentemente *AI*, *data mining* e *machine learning*, che utilizzano un apparato computazionale molto complesso, e proprio per questo vengono spesso applicati e percepiti come “*black box*” dove in sostanza l'utente immette dei dati e ottiene dei risultati anche interessanti senza tuttavia comprendere come siano stati ottenuti e senza essere intervenuto in aspetti di dettaglio ma qualificanti del metodo utilizzato, come la specificazione di parametri o forme funzionali. Questi metodi si differenziano quindi sia dalla geografia teoretico-quantitativa sia dall'analisi *data driven*: in essi l'elemento di controllo non è più la teoria e paradossalmente non lo sono più nemmeno i dati, ma sono le stesse tecniche esplorative che cercano di far emergere da quei dati dei *patterns* che, ovviamente, sono fortemente influenzati dalla tecnica impiegata e che non avrebbero ragion d'essere senza i *big data*.

Se dunque i dati e la grande capacità di calcolo sembrano farla sempre più da padroni nella gamma di funzioni analitiche che potrebbero diventare dominanti, nella letteratura più recente cominciano a comparire anche importanti elementi di critica nei loro confronti. Nell'articolo dall'interessante titolo «*Big Issues for Big Data: Challenges for Critical Spatial Data Analytics*», Chris Brunsdon e Alexis Comber

(2020) si interrogano sul ruolo dei *big data* nel contesto della *spatial data science*. Essi pongono l'accento proprio sui problemi di un'analisi di dati che non ne consideri distribuzione e significatività statistica. L'importanza di tali problemi e della loro mancata risoluzione pone, effettivamente, in discussione la rilevanza delle questioni di ricerca – e delle relative risposte – poste da queste metodologie, soprattutto dal punto di vista delle scienze sociali. Gli autori arrivano così ad evocare la necessità di un approccio critico allo *spatial data analytics*⁴⁹. Nella stessa direzione si pone Alan Murray (2020) che, centrando il suo articolo sul problema della significatività statistica, osserva come «la difficoltà è invariabilmente quella di stabilire che i risultati ottenuti abbiano un significato e forniscano illuminazioni [insights] e conoscenza»⁵⁰. È infine importante notare come Alan Murray sembri usare il termine “*spatial analytics*” come sinonimo di “*spatial analysis*”, ciò che si evince dai riferimenti sia alla dimensione del campione, che può essere anche molto piccolo, sia al suo stesso articolo del 2010 sulla geografia quantitativa, sia, infine, al richiamo alla citazione di Luc Anselin nei riguardi della *spatial data science*⁵¹. Le riflessioni critiche di autori come Murray, Anselin e Brunsdon, profondamente radicati nella tradizione dell'analisi statistica spaziale, fanno supporre che alla stagione del predominio dei dati e della potenza di calcolo si stia già

⁴⁹ Brunsdon e Comber (2020, p. 89): «In this paper we consider some of the issues of working with big data and big spatial data and highlight the need for an open and critical framework. We focus on a set of challenges underlying the collection and analysis of big data. In particular, we consider 1) inference when working with usually biased big data, challenging the assumed inferential superiority of data with observations, n , approaching N , the population ($n \neq N$). We also emphasise 2) the need for analyses that answer questions of practical significance or with greater emphasis on the size of the effect, rather than the truth or falsehood of a statistical statement; 3) the need to accept messiness in your data and to document all operations undertaken on the data because of this, in support of openness and reproducibility paradigms; and 4) the need to explicitly seek to understand the causes of bias, messiness etc in the data and the inferential consequences of using such data in analyses, by adopting critical approaches to spatial data science. In particular we consider the need to place individual data science studies in a wider social and economic contexts, along with the role of inferential theory».

⁵⁰ Murray (2020, p. 1): «The challenge is invariably establishing meaning in obtained findings, providing insights and knowledge». Si osservi che abbiamo volutamente tradotto l'inglese “meaning” con “significato” proprio per sottolineare la relazione tra significatività [statistica]e significato [senso].

⁵¹ Murray (2020, p. 1): «Anselin (2020) characterized spatial data science as seeking to extract meaningful information from data involving location, proximity, and interaction, requiring specialized expertise to appropriately deal with the complexities of geography».

apprestando a succedere quella di un approccio analitico statistico anche all'analisi delle grandi quantità di dati.

5.3.4 *Molteplicità di paradigmi e metodi*

Già nell'ormai datato dibattito sul GIS come scienza furono poste importanti questioni sulla legittimazione, l'epistemologia e la metodologia di quella potenziale scienza, e, trattandosi in larga misura di una scienza applicata, sul suo impatto sulla società⁵². Di alcune di queste questioni si è discusso nel contesto dell'etica e della scienza critica, ma un altro aspetto importante è la crescente messa in discussione del paradigma neopositivista, dominante nella geografia teoretico-quantitativa e nelle fasi iniziali del GIS⁵³. La questione coinvolge la stessa definizione di scienza⁵⁴ tradizionalmente concepita come un approccio logico e sistematico a problemi che richiedono risposte generalizzabili, da cui risulta evidente l'affiliazione ad un particolare paradigma scientifico – quello neopositivista – che di fatto ignora l'intera gamma di quelli alternativi, con le rispettive ontologie, epistemologie e metodologie⁵⁵. La questione non è molto diversa da quella che negli anni '60 portò all'emergere di una Scienza Regionale⁵⁶ per fornire una struttura intellettuale in cui l'analisi spaziale, la geografia economica e l'economia potessero svilupparsi per servire da supporto all'economia regionale, alla pianificazione regionale e all'analisi geografica, come notato da John Pickles (1997). La discussione ha radici

⁵² Si vedano in proposito: Goodchild (1992); Pickles (1995a); Wright, Goodchild, Proctor (1997).

⁵³ Si veda al riguardo il pungente dibattito, apparso su *Environment and Planning A*, tra Openshaw (1991; 1992), Fotheringham (1991) e Clark (1992).

⁵⁴ Secondo alcune proposte essa si potrebbe sostanzialmente identificare con queste tre posizioni: “*Obtaining theoretical knowledge to form the design of a model*”; oppure “*developing theory on entities such as time and spatial phenomena*”; oppure ancora “*developing algorithms to test a theory*”. Si vedano in proposito Pickles (1997) e Wright, Goodchild, Proctor (1997).

⁵⁵ La visione positivista è stata finora privilegiata in GIS, pure se alcuni riconoscono nella crescente letteratura sull'epistemologia del GIS l'intero spettro di approcci, dal positivista al postmodernista [Wright, Goodchild, Proctor (1997)]; si veda anche Bertazzon e Waters (1996).

⁵⁶ Interessante è al riguardo notare come fino alla seconda metà degli anni '60 molti geografi (fra i quali T.R. Smith, Warntz, Stewart, Berry, Garrison, Ulmann, Dacey) erano intervenuti nei dibattiti sugli scopi della Regional Science: si vedano i primi numeri della rivista *Paper and Proceeding of Regional Science Association*.

profonde, e da lungo tempo si è ammessa la possibilità di paradigmi o matrici disciplinari diverse, tra cui si sono adottati il relativismo sociale, lo strutturalismo radicale, il neoumanesimo⁵⁷, in linea con la proposta di Wright, Goodchild, Proctor (1997) di una visione più generosa, in cui la scienza non appaia confinata ad alcuna particolare epistemologia.

Già in Bertazzon e Waters (1996) furono esaminate possibili alternative all'approccio funzionalista o logico-positivista, considerando le matrici disciplinari del relativismo sociale, dello strutturalismo radicale, del neoumanesimo e della conoscenza tradizionale della natura propria delle culture indigene non occidentali (*Traditional Environmental Knowledge*, o TEK). Da allora, molti progressi si sono compiuti nella realizzazione di sistemi denominati con diverse sigle, come PGIS (*participatory GIS*), PPGIS (*public participation GIS*), integrando nel GIS diverse matrici disciplinari alternative all'approccio funzionalista. Tra questi va ricordata non solo l'integrazione del TEK (*traditional environmental knowledge*) ma anche l'utilizzo dell'informazione geografica come strumento di riappropriazione dei propri territori e di decolonizzazione in varie parti del mondo⁵⁸.

L'utilizzo della GI (*Geographic Information*) come strumento di lotta e decolonizzazione rappresenta di per sé un suo utilizzo critico. In tale contesto spesso hanno più importanza le concettualizzazioni e le rappresentazioni mentre l'analisi è spesso limitata ad aspetti descrittivi ed elementari. Un aspetto che invece assume un'importanza preponderante è quello della rappresentazione cartografica, o visualizzazione, o, come si è visto, contro-cartografia⁵⁹. Il discorso è di particolare attualità in contesto cartografico, ma l'analisi spaziale potrebbe non essere lontano da venire con i *big data*, per cui non è impensabile che le "controcartografie" possano presto essere affiancate da una sorta di "contro-analisi-spaziali" che realizzino l'argomentazione di Brunson e Comber in favore di ripetibilità e trasparenza, arrivando potenzialmente a costruire narrazioni alternative di fenomeni come il cambiamento climatico o la diffusione della pandemia.

⁵⁷ Si veda in proposito Bertazzon e Waters (1996).

⁵⁸ Se vedano, fra l'altro: Akerman (2017); Hunt e Stevenson (2017); Mundy (2000); Sletto (2009); Wainwright e Bryan (2009).

⁵⁹ Come notano Akerman (2017) e Hunt e Stevenson (2017).

L'apertura di un dialogo inter- e intra-disciplinare che ha accompagnato il cammino dalla critica all'autocritica nella *GIScience* ha anche consentito un dialogo, non solo all'interno della geografia, ma anche con le scienze sociali e naturali, volto ad accettare e, ove possibile, superare i limiti di ciascuna prospettiva metodologica. Da qui sono stati proposti approcci nuovi, e in particolare le metodologie miste, o *mixed methods*, che affrontano un singolo problema di ricerca da una molteplicità di prospettive concettuali e metodologiche, nella convinzione che l'integrazione di metodi diversi possa fornire una risposta migliore di un singolo metodo ad una qualsiasi questione di ricerca⁶⁰. Seppure prospettive qualitative possano arricchire l'analisi, uno dei problemi classici è quello dell'inferenza statistica, che impone una sorta di trade-off tra la prospettiva statistico-quantitativa, che richiede campioni numerosi e rappresentativi, e quella qualitativa, che, richiedendo analisi approfondite, tende a considerare campioni di piccole dimensioni. La questione appare quasi paradossale ma di vitale importanza, poiché potrebbe costringere l'analisi spaziale a riconsiderare l'inferenza statistica in una prospettiva che spazia dal campione improprio perché troppo piccolo dell'analisi qualitativa, fino a quello improprio perché troppo grande ($n=all$) dei *big data*.

Interessante a questo proposito sono due recenti filoni di ricerca nell'ambito della GI (*Geographic Information*), entrambi oggetto di attenzione da parte della rivista *Transactions in GIS*. Il primo s'interessa agli aspetti geografici e culturali dell'informazione geografica. Affrontato in un workshop della conferenza AGILE 2019⁶¹, denominato GeoCultGIS, i contributi a questa discussione sono stati

⁶⁰ Si vedano in particolare gli articoli presenti nei due testi: Cope e Elwood (2009); Elwood (2010).

⁶¹ Si veda www.cs.nuim.ie/~pmooney/GeoCultGIS/, in cui i contenuti da sviluppare in questo contesto vengono così articolati nella call for papers: conceptual, methodological, or empirical short papers discussing the cultural and geographical aspects of geo-information, in particular, but not limited to, subjects such as: the effects of data production practices and of geographical peculiarities on data quality and representations, and approaches for treating related biases; novel and integrative (i.e. multi-perspective) representations of geographical realities and objects; developments and studies considering contextual dependencies in the perception and usage of geographical concepts and categories (e.g. land use nomenclatures, tagging of geo-objects, place boundaries); the effects of geographical and cultural diversities on the transferability and application of data processing methods, and context-sensitive methodological developments; GIScience research front and agendas in attending to geo-cultural dependencies).

pubblicati in una sezione speciale di *Transactions in GIS*. Questa piccola raccolta si occupa delle decisioni circa la codifica e modellizzazione dell'informazione geografica, che sono guidate dalla teoria e influenzate dai contesti sociali e culturali in cui sono prese. L'attenzione è rivolta in particolare a tre ruoli della geoinformazione: come rappresentazione di contesti geo-culturali, come prodotto di tali contesti e in relazione all'impatto che su tali contesti produce. Queste relazioni richiedono attenzione e perciò si può far leva sulle tecnologie geoinformatiche nel contesto del GIS partecipativo e del GIS a partecipazione pubblica⁶².

Anche nel secondo caso l'idea prende le mosse da un workshop tenutosi all'università di Heidelberg: PLATIAL'18⁶³ i cui risultati sono confluiti su un numero speciale di *Transactions in GIS* uscito nell'Agosto 2021. Il tema della "plazialità"⁶⁴ è ben rappresentato dall'editoriale dal titolo: "*A Place for Place: Modelling and Analysing Platial Representations*". Come affermano gli autori, il concetto di *place* è un argomento di interesse crescente tra gli studiosi di *GI-Science*. Si sono compiuti i primi tentativi di formalizzare l'informazione *platial* ed è sempre più generalmente accettato che i dati, particolarmente quelli generati dall'utente, sono per loro natura "*platial*". Allo stesso tempo, e specialmente in confronto con i concetti spaziali geometrici, il concetto di *place* è ambiguo, complesso e difficile da

⁶²Si veda Grinberger e Novack (2021, p. 577): «Decisions on how to encode and model information in a geographical database are theory-laden and contingent upon the social and cultural contexts in which they are made. Hence, the very meaning and structure of geoinformation and geodata intrinsically embed these contexts. Furthermore, representations of the world via geoinformation may emphasize certain perceptions of the world or promote new ones, hence facilitating social processes of digital transformation and marking a complete cycle of influences (i.e., from culture/society to geoinformation and back again). The first relation (geoinformation as a representation of geo-cultural contexts) suggests that within the field of geoinformation and GIS lie opportunities for geo-cultural inquiries that are (still) welcoming methodological advances. The second relation (geoinformation as a product of geo-cultural contexts) raises concerns regarding data interoperability and bias, emphasizing the need to understand how processes of production affect the product and how meaning is being transformed over space and culture. The third relation (geoinformation as an impact upon geo-cultural contexts) calls for leveraging geoinformation technologies and procedures for social change, as is being done within the fields of participatory GIS and public participation GIS».

⁶³PLATIAL'18: Workshop on Platial Analysis (<http://platial18.platialsience.net/>).

⁶⁴*Platial* è un aggettivo costruito sul sostantivo *place*. In lingua inglese i termini *space* e *place* presentano un'assonanza che non trova corrispondenza in italiano per cui è stato facile costruire il termine *platial*. Termine che non ha una traduzione propria in italiano: per questo utilizzeremo sempre *platial*.

catturare in termini formali e analitici, ciò che suggerisce la necessità di approcci interdisciplinari⁶⁵.

I curatori del workshop⁶⁶ riconoscono come la ricchezza del termine *place* ne comporti una vaghezza che finora ne ha reso difficile l'impiego in contesti quantitativi; tuttavia, essi riconoscono anche la necessità di definire il concetto di *place* in *GIScience*, assieme ad una teoria che consenta di caratterizzare, rappresentare ed utilizzare il *place* in *GIScience*. Anche in questo contesto sorgono seri dubbi sul fatto che i *geosocial media* e l'ondata di dati geospaziali possano fornire *da soli* l'impeto necessario a promuovere la ricerca e far muovere l'analisi "*platial*".

5.4. Le nuove frontiere dell'analisi spaziale

L'analisi spaziale, con la *GIScience*, ha dunque saputo fermarsi e considerare criticamente se stessa e il proprio oggetto. Si tratta di un passaggio importantissimo, in cui la disciplina ha saputo superare una posizione di arroccamento che aveva contraddistinto le frizioni disciplinari degli anni '90, e si tratta anche di una lunga fase che l'analisi spaziale sta vivendo non da sola, ma come parte di una più ampia *GIScience*. Questo contesto più vasto ne amplia la prospettiva, abbracciando dibattiti che spaziano fino alla cartografia, al *place*, al pensiero critico, avvicinandosi sempre più al nucleo dell'analisi spaziale.

La *GIScience* ha saputo superare il paradigma positivista e proprio in questo modo ha potuto aprirsi a prospettive diverse e fornire i propri strumenti ad istanze sociali, culturali, etniche, di genere impensabili in

⁶⁵ Westerholt, Mocnik, Comber (2020, p. 811): «The concept of place is a topic of increasing interest among GIScience scholars. First attempts to formalise *platial* information have been made and it is increasingly held that user-generated data sets in particular are often more *platial* than *spatial* in nature. At the same time, and especially when compared to geometric *spatial* concepts, the concept of *place* is ambiguous, complex and difficult to capture in formal and analytical terms, suggesting the need for interdisciplinary approaches».

⁶⁶ Westerholt, Mocnik, Comber (2020, pp. 811-812): «Consideration of *place* by GIScience scholars has increased over recent years due to the greater availability of *platial* data from diverse sources. [...] However, [...], the inherent vagueness in the definition of *place* has so far limited the progress on *platial* concepts within GIScience. [...] Yet, there is still a lack of a holistic consistent theory of how *places* can be characterised, represented, and used in a formal way. A *place*-based approach to GIS and analysis is nevertheless important, given the wealth of increasingly *place*-based information available to us in an increasingly digital world».

quel paradigma. Tutto questo non solo sancisce il superamento di quelle posizioni implicitamente conservatrici per cui quantitativismo e GIS erano stati da sempre criticati, ma le consente di contare su istanze e situazioni che si spingono ben oltre il dominio classico della geografia quantitativa. Chiave per il superamento delle vecchie posizioni è non solo il pensare spaziale, ma soprattutto il pensiero spaziale critico. Questo però non basta all'analista spaziale. Per poter *contare*, la geografia quantitativa non può più fermarsi ai numeri. La prospettiva critica ha squarciato i veli su come viene svolta l'analisi, e richiede oggi che le procedure siano trasparenti e replicabili. Richiede che l'analista spaziale sappia programmare le proprie analisi, scrivendone i codici. Ciò che non si poteva pretendere da chi fa GIS, si pretende da chi fa analisi spaziale: essere geografo, informatico, matematico e statistico, e critico. Si tratta di domande, di pretese enormi. Tutto questo è davvero necessario ma nel contempo è troppo per una sola figura, per una sola disciplina che rimane ancora di dimensioni ridotte. Eppure, la rivoluzione dei *big data*, la tecnologia pervasiva, richiedono proprio all'analisi spaziale di essere tutto questo e forse di più. Ed essa non può sottrarsi a queste domande.

Le questioni sono sempre più *grandi*. Benché non siano disgiunte dal metodo, esse si spingono ben oltre l'uso di metodologie quantitative, e perciò non possono essere affrontate nell'ambito della sola geografia quantitativa, ma richiedono una riflessione che possa avvalersi dell'esperienza e delle conoscenze maturate in altri rami della disciplina. Ciò che forse è davvero rivoluzionario in questa fase è proprio il fatto che i geografi quantitativi sembrano preparati ad affrontare le questioni non più da soli, ma in un dialogo aperto e costruttivo dentro e fuori la geografia⁶⁷.

Qualunque sia la natura della presunta rinascita del quantitativismo in geografia, si tratta di un fenomeno che potrebbe presentare delle opportunità, anche se non prive di difficoltà, non solo alla geografia quantitativa, ma alla geografia tutta. È dunque importante che la geografia sappia compiere una scelta matura e consapevole, così da poter, consapevolmente, cogliere oppure rifiutare le opportunità e le sfide poste dalle attuali circostanze. Le conseguenze di questa scelta saranno probabilmente importanti per l'intera disciplina, in Italia e altrove.

⁶⁷ Si vedano: Johnston *et Alii* (2014); Goodchild e Janelle (2010).

Parte III

La pratica

6. La pratica: dati, attrezzi, personale

[...] non ci sono “*dati*” non interpretati: non c’è niente, semplicemente, di “dato” a noi senza essere interpretato; niente da prendere come base. *Tutta* la nostra conoscenza è interpretazione alla luce delle nostre aspettative, delle nostre teorie, ed è perciò, in un modo o nell’altro, *ipotetica*.

Popper, *Poscritto alla logica della scoperta scientifica*, 2009, p.136.

6.1. Premessa

In questo volume abbiamo presentato l’analisi spaziale come una disciplina accademica e uno strumento per la gestione del territorio. Ne abbiamo ampiamente discusso i numerosi aspetti teorici, concettuali e critici, ma abbiamo anche voluto fornire una panoramica concreta di quale sia l’oggetto specifico dell’analisi spaziale. Un’applicazione di analisi spaziale è stata illustrata nell’Esempio 3.1, e un’ulteriore breve rassegna di applicazioni sarà presentata nel prossimo capitolo. Per realizzare analisi spaziali come quella dell’Esempio 3.1 occorrono alcuni “ingredienti”, che si possono sintetizzare in poche componenti fondamentali: dati, *software*, e personale qualificato. Questo capitolo si articola in una serie di considerazioni su ciascuna di queste componenti. Nessuno di questi aspetti può essere trattato in maniera esaustiva: ci limiteremo quindi ad alcune riflessioni che ci appaiono più rilevanti, e che riteniamo possano essere utili a chi intenda avvicinarsi alla pratica della disciplina.

Definiremo i dati partendo da una definizione minima e condivisibile sul termine “dati” e soffermandoci su ciò che distingue i dati spaziali. Presenteremo alcune considerazioni sulla raccolta e il campionamento dei dati spaziali, con un’attenzione critica all’errore insito in queste operazioni, che tende a propagarsi attraverso la manipolazione e l’analisi dei dati. Parleremo di incertezza dei dati spaziali, della sua influenza sui risultati analitici e dell’importanza delle tecniche statistiche per gestire l’errore spaziale insito sia nei dati sia nelle analisi spaziali. Ci soffermeremo sul cosiddetto diluvio di dati, sulle sue

implicazioni e sugli elementi critici che solleva, con qualche riflessione sulla transizione da un ambiente povero ad uno ricco di “dati”. È certo vero che senza dati non ci sarebbe analisi, come è vero che la qualità dei dati influisce sulla qualità dei risultati, come esprime efficacemente un noto adagio in lingua inglese: “*garbage in, garbage out*”¹. Ciò che ci interessa, dunque, non è tanto la quantità, ma la qualità dei dati.

Il secondo elemento è la cassetta degli attrezzi, e il ruolo dell’analista spaziale rispetto alla loro costruzione, utilizzo e perfezionamento. Ci occuperemo degli strumenti *software* dell’analisi spaziale, illustrando concretamente come si fa analisi spaziale e quali sono le domande e le potenzialità che il *software* presenta a chi lo avvicina. Vedremo che il *software* analitico è molto diverso dal tradizionale *software* GIS e comprenderemo come, per questa via, esso contribuisca a forgiare un professionista dell’analisi spaziale che è diverso da quello del GIS. Ritourneremo alla cosiddetta tensione tra elitismo e populismo, chiedendoci se davvero il populismo possa giovare alla diffusione dell’analisi spaziale.

L’analista spaziale (la componente umana) è l’ultimo – ma solo cronologicamente – elemento di cui ci occupiamo: il cosiddetto *live-ware*, l’elemento vivo, umano, senza il quale non avrebbero senso *software*, *hardware*, e neppure i dati. Questo elemento, così importante, è forse quello che riceve meno attenzione nel dibattito disciplinare, anche se alcuni dibattiti recenti si sono occupati dell’importanza dell’insegnamento della materia. Accenneremo quindi all’insegnamento, e alle difficoltà che presenta, e concluderemo con il ruolo potenziale di questa figura professionale nell’affrontare i problemi della società, e nello sviluppo futuro della disciplina.

Concluderemo infine con alcune considerazioni che ci traghettano proprio verso le applicazioni pratiche dell’analisi, che saranno l’oggetto del capitolo successivo.

¹ Che potremmo rendere come: “se entra spazzatura, esce spazzatura”, ovvero, se si analizza spazzatura, il risultato non può che essere spazzatura.

6.2. Dati, grandi dati, dati spaziali, analisi dei dati

Per chi, come noi, parla una lingua latina, il termine “dato” o il suo plurale “dati” è intuitivo. Il vocabolario Treccani² lo definisce come:

- a. “Ciò che è immediatamente presente alla conoscenza, prima di ogni forma di elaborazione”.
- b. “Con uso più generico, elemento, in quanto offerto o acquisito o risultante da indagini e utilizzato a determinati scopi [...]. Con uso più specifico, in informatica, *dati*, elementi di un’informazione costituiti da simboli (numeri, lettere: *d. numerici, alfabetici, alfanumerici*) che devono essere elaborati, per lo più elettronicamente, secondo un determinato programma”.

In inglese, la lingua in cui è scritta la gran parte della moderna analisi spaziale, il termine si usa direttamente in latino, soprattutto al plurale, *data*. Oltre a non evocare con la stessa immediatezza l’idea di “qualcosa che viene dato”, la forma insolita del sostantivo [la finale in *-a* per un plurale] crea confusione, tanto da dar luogo ad una curiosa pluralità grammaticale, per cui si legge tanto “*data are...*” quanto “*data is...*”. Il dizionario inglese Oxford lo definisce come un nome collettivo (*mass noun*), accettabile sia come singolare che come plurale, e ne fornisce, come prima definizione: “*Facts and statistics collected together for reference or analysis*”³.

Sia in italiano che in inglese, dunque, il termine “dati” è ormai quasi un nome collettivo, anche se in italiano rimane grammaticalmente un plurale, che si riferisce ad un’entità composita più che ad una pluralità di elementi singoli. Particolarmente condivisibile è la definizione che Treccani attribuisce al linguaggio informatico, secondo cui i dati sono costituiti da simboli; a questa dimensione simbolica aggiungiamo fin d’ora che, per noi, questi simboli (più spesso numerici, ma anche alfabetici e alfanumerici) rappresentano, in maniera codificata, un fenomeno geografico, e che questi simboli (numeri), direttamente o dopo manipolazioni, costituiscono l’input dell’analisi spaziale. A questo

² www.treccani.it/vocabolario/dato/ (visitato il 21/05/2021).

³ www.lexico.com/definition/data/ (visitato il 21/05/2021): «In Latin, data is the plural of datum and, historically and in specialized scientific fields, it is also treated as a plural in English, taking a plural verb, as in the data were collected and classified. In modern non-scientific use, however, it is generally not treated as a plural. Instead, it is treated as a mass noun, similar to a word like information, which takes a singular verb. Sentences such as data was collected over a number of years are now widely accepted in standard English».

ovviamene aggiungiamo che l'analisi spaziale poggia necessariamente su dati *spaziali*, sulle cui caratteristiche ci soffermeremo più avanti.

Se dunque questi simboli vengono *dati*, chi e perché li dà? Per molti di noi la prima vera esperienza con la ricerca e l'analisi di dati è la tesi di laurea⁴. Tradizionalmente⁵ la ricerca dei dati era una componente molto cospicua di quel progetto, spesso associata ad un viaggio fisico, per biblioteche e archivi o per svolgere ricerca sul campo, e per ottenere le informazioni di norma era necessario chiederle a qualcuno, che ce le avrebbe materialmente *date*.

Ai tempi della geografia teoretico-quantitativa, e anche dopo l'avvento del GIS e dell'analisi spaziale, l'acquisizione dei dati – spaziali e non – avveniva manualmente, con tecniche di tracciamento e copiatura manuale lunghe e soggette ad errore⁶. Oltre al costo dell'acquisizione dei dati, anche lo stoccaggio, l'organizzazione e l'elaborazione di “grandi” quantità di dati erano difficili e richiedevano molte ore di lavoro⁷. Come si è visto nel capitolo 1, lo stesso CGIS di Roger Tomlinson (1988) fu creato per automatizzare la produzione cartografica, riducendone i costi. Lo stoccaggio dei dati era problematico perché i dischi erano piccoli e i dati voluminosi. Per esempio, benché il modello *raster* presentasse dei vantaggi, era molto ingombrante nello spazio-disco, per cui sorsero metodi come il *quadtree*⁸ che consentiva di trascurare informazioni ridondanti. Sui limiti dello spazio-disco si accumulavano i limiti computazionali, non indipendenti dai precedenti; per cui, benché i dati fossero relativamente pochi (secondo gli standard odierni) si cercava di ridurli, con tecniche di riduzione (*data reduction techniques*) che consentivano di estrarre l'informazione rilevante eliminando ogni ridondanza⁹.

⁴ È ancora interessante rileggere, al riguardo, il testo che Eco (1977) ha scritto quasi mezzo secolo fa.

⁵ Almeno fino agli anni 1990, quando crebbe la disponibilità di dati anche in formato digitale (Openshaw e Albanides, 1999).

⁶ Errori che, come si vedrà poco più avanti, tendono a trascinarsi e anche amplificarsi attraverso i passaggi analitici successivi.

⁷ Si pensi solo alla necessità di impostare un modello ed inviarlo ad un sistema di elaborazione centralizzato, per poi recarsi fisicamente a ritirare le stampe dei risultati e, se necessario, ripetere l'operazione sino ad ottenere risultati soddisfacenti.

⁸ Si veda ad esempio Samet (1984).

⁹ Tra queste tecniche ricordiamo l'analisi delle componenti principali (*principal component analysis*), l'analisi fattoriale (*factor analysis*) e l'analisi dei clusters (*cluster analysis*),

Quando l'acquisizione e l'elaborazione dei dati erano dispendiose, i dati erano certo importanti per rispondere alla questione di ricerca, ma rimanevano un elemento ancillare rispetto alla questione. La grande disponibilità di dati degli ultimi decenni ha contribuito, paradossalmente, a sbilanciare questo rapporto, fino a capovolgerlo, cosicché spesso sono i dati a dettare, non solo le tecniche analitiche, ma persino le questioni di ricerca.

Oggi non serve più cercare o chiedere i dati, perché siamo sommersi dal cosiddetto *diluvio*¹⁰ per cui i dati ci piovono addosso da ogni parte: dai cosiddetti *social media* su cui comunichiamo incessantemente e da costellazioni di satelliti che ne inviano di continuo (*big data*), tramite la cessione gratuita di dati sul web (*open data*) e la cessione volontaria di informazioni corredate di posizione geografica (*volunteered geographical information - VGI*). Anche molti *big data* ed *open data*, oltre ovviamente ai *VGI*, sono riferiti ad una precisa posizione geografica, ossia georeferenziati. I computer, che tutti ormai possono permettersi, li possono stoccare in quantità enormi ed elaborare in tempi rapidissimi, e poi quei computer non sono neppure necessari, perché le nuove tecnologie (*web 2.0*) consentono di manipolare i dati anche lasciandoli nella grande nuvola (*cloud*), facendone, volendo, analisi spaziale e cartografia. I limiti tecnologici e computazionali del secolo scorso sono stati in gran parte superati; eppure, anche se non serve più ridurre i dati per stoccarli ed analizzarli, paradossalmente bisogna ancora sfoltirli per poterne estrarre informazioni sensate. Per questo ci sono tecniche nuove incardinate però nelle vecchie tecniche di riduzione dei dati¹¹.

Secondo la definizione data poco sopra, tutto ciò che ci piove addosso ogni giorno costituisce dei "dati". Ma quanti e quali di questi dati sono davvero *utili* ad affrontare questioni di ricerca *rilevanti*? C'è differenza tra la miriade di frammenti sconnessi palleggiati quotidianamente dagli adolescenti, rispetto a comunicazioni potenzialmente utili ad identificare focolai epidemici o cellule terroristiche. E in ogni caso i dati grezzi richiedono ancora trasformazioni laboriose, a monte di qualsiasi analisi spaziale¹². Di fronte all'inflazione di dati degli anni

tutte tecniche molto utilizzate dalla geografia teoretico-quantitativa. Si possono consultare, in proposito, i rispettivi volumi della serie CATMOG.

¹⁰ Si veda ad esempio Kitchin (2014, p. 3).

¹¹ Tra esse si ricordano il *data mining* e la *geocomputation*.

¹² Si vedano Burrough e McDonnell (1986) e Wan *et Alii* (2015).

'90 Stan Openshaw derideva come “pittoreschi” i collezionisti di dati¹³: coloro che, quasi come bambini, raccolgono dati, amano possederne *tanti*, li ordinano, li ammirano, ne fanno grafici e mappe, *ma non sentono neppure il bisogno di analizzarli*.

Quest'immagine può far sorridere, ma non ci si deve ingannare: il collezionismo di dati è facile, appagante e non richiede la fatica dell'analisi seria. Ma pretende grandi quantità: *big data*. E il fascino di quelle *grandi* collezioni, la lusinga delle statistiche superficiali e delle visualizzazioni d'effetto non solo può distogliere dall'impegno dell'analisi, ma anche trascinare a quella fede nei dati che giunge a capovolgere lo stesso metodo scientifico, privilegiando i dati, anziché le questioni da analizzare.

6.2.1. I dati spaziali e le loro caratteristiche

Spaziale è speciale, come affermava già oltre 30 anni fa Luc Anselin, e come è stato spesso ribadito ed enfatizzato, tra gli altri, da Michael Goodchild¹⁴. Che cosa, dunque, differenzia i dati spaziali da *altri* dati? Anche se è difficile, specie per un geografo, pensare a dati che non siano, almeno implicitamente, situati nello spazio, si pensi ai classici esperimenti statistici: il lancio di monete, di dadi, o l'estrazione di carte da un mazzo. Benché anch'essi si svolgano nello spazio [e nel tempo], il luogo [e il momento] dell'esperimento non sono semplicemente ininfluenti, ma anzi, generalmente, ai fini della loro validità, si richiede proprio che il risultato sia indipendente dal luogo [e dal momento] in cui sono condotti.

Al polo opposto ci sono invece i dati per i quali il luogo [e il momento] dell'osservazione riveste un'importanza fondamentale, tale da consentire la costruzione di serie spaziali e serie temporali di dati. Esempi classici di questo tipo sono le misurazioni di temperatura e

¹³ Openshaw (1994, p 6): «You can use any one of a dozen different statistical packages, some with linked map and graph displays. Software abounds that is able to help you visualise, plot, and analyse such data. Very picturesque!»

¹⁴ L'articolo di Anselin (1989) è intitolato proprio “What is Special About Spatial Data? Alternative Perspectives on Spatial Data Analysis”. La fortunata espressione è stata ripresa più volte anche informalmente in vari consessi; si vedano ad esempio Goodchild (2009) e Haining (2009).

umidità; la misurazione o il calcolo dell'altitudine, della densità di popolazione, del tipo di coltura o del patrimonio architettonico in un determinato luogo.

Per i dati temporali, l'attributo è accompagnato dal riferimento temporale; ad esempio, per la temperatura rilevata da una stazione meteorologica, il riferimento temporale è costituito dal momento del rilevamento e il dato è definito da una coppia di valori numerici: il riferimento temporale e la temperatura misurata. Prendiamo ad esempio la stazione di rilevamento della rete ARPA di Cortina d'Ampezzo: la temperatura minima (Temperatura aria a 2m) del 13 aprile 2021 è stata rilevata nell'istante solare 202104130315 e misurata in -2.4 gradi centigradi¹⁵. Essa è quindi definita dalla coppia numerica [202104130315, -2.4].

I dati spaziali sono analogamente definiti dal riferimento spaziale e dall'attributo, ma con l'importante differenza che il riferimento spaziale è costituito da – almeno – due valori numerici, che sono spesso le coordinate geografiche espresse in latitudine e longitudine. Il dato spaziale, quindi, è espresso da una coppia di coordinate geografiche, oltre l'attributo, per un totale di – almeno – tre valori numerici. Nell'esempio, la stazione ARPA di Cortina è definita dalle coordinate 46.53668964 (latitudine) e 12.12739699 (longitudine)¹⁶. La temperatura minima di Cortina è quindi espressa dalla tripletta numerica [46.53668964, 12.12739699, -2.4]. Non di rado il riferimento geografico necessita anche dell'altitudine, come nel caso di Cortina. In tal caso il riferimento spaziale è formato da tre elementi (le due coordinate, più l'altitudine in metri), oltre ovviamente, all'attributo; la definizione è così costituita dall'insieme di quattro numeri [46.53668964, 12.12739699, 1271, -2.4]. Nei *database* GIS è possibile associare alle coordinate geografiche numerose caratteristiche, come per esempio il momento della rilevazione, l'umidità relativa o la direzione del vento, per cui il dato spaziale, oltre ai riferimenti spazio-temporali può contenere una molteplicità di attributi.

Rispetto ad altri dati, il dato spaziale è dunque caratterizzato da una maggiore dimensionalità numerica. Inoltre, se i dati temporali possono essere ordinati cronologicamente senza ambiguità, nello spazio non è

¹⁵ www.arpa.veneto.it/bollettini/meteo/h24/minmaxoggi/minmaxodierne.xml

¹⁶ Coordinate angolari x, y, espresse in decimali.

data una direzione privilegiata, e il problema dell'ordinamento può essere ulteriormente complicato dall'eventuale irregolarità delle unità spaziali. Nel caso di maglie regolari, i cosiddetti *raster* (ad esempio i *pixels* di un'immagine satellitare) l'ordinamento è arbitrario ma privo di ambiguità¹⁷, ma nel caso di unità spaziali di forma e dimensione irregolare, come regioni, province, comuni, anche un ordinamento da ovest a est e poi da nord a sud (come per i *pixels*), risulta ambiguo.

In tutto questo s'innesta la variazione dell'attributo (es. densità di popolazione) sulle unità spaziali. L'analisi della distribuzione, o variazione spaziale dell'attributo, quindi, si integra con quella del riferimento spaziale, dando luogo ad una complessità superiore a quella di altri tipi di dati, e alle proprietà note come dipendenza ed eterogeneità spaziale ampiamente discusse nel capitolo 2.

Come è noto, nella ricerca applicata non è sempre possibile conoscere un'intera popolazione d'interesse, per cui sono necessarie forme di campionamento, ad esempio sistematico, casuale, o a cascata. I dati spaziali richiedono considerazioni specifiche, poiché le proprietà dei fenomeni spaziali richiedono schemi di campionamento più elaborati rispetto ai dati privi di riferimenti spazio-temporali¹⁸.

Campionamento spaziale. Per analizzare la fertilità del suolo o l'inquinamento dell'aria si può ipotizzare una base relativamente omogenea, su cui impiegare schemi di campionamento regolare; se invece si studia l'incidenza di una malattia, il campionamento deve considerare il sostrato formato dalla maglia irregolare dei comuni, su cui è misurata la popolazione. Diverso è l'esempio di un faunista che studia la presenza di volpi in una certa area; tipicamente lo studioso avanza su un percorso dal quale osserva l'eventuale presenza dell'animale entro una fascia di visibilità (*buffer*). Nei casi più fortunati egli può camminare su un prato aperto e senza barriere, dove definirà uno schema di linee perpendicolari (detto transetto o *transect*) che gli consenta di coprire, dato il *buffer*, l'intera area (ossia ogni linea disterà dalla successiva il doppio dell'ampiezza del *buffer*). La regolarità decresce se

¹⁷ Le caselle, o *pixels*, vengono di norma numerate da ovest a est e poi da nord a sud, ovvero a partire da quella più in alto a sinistra, procedendo poi come si fa con le righe di testo. Si veda ad esempio Jensen (2007).

¹⁸ Si veda ad in proposito qualsiasi manuale di analisi geografica quantitativa, ad esempio Burt, Barber, Rigby (2009, pp. 282-289).

il percorso deve adattarsi a versanti ripidi, corsi d'acqua, ecc., ove lo schema di campionamento dipende anche dalle caratteristiche della zona e del sentiero.

Incertezza posizionale o localizzativa. Oltre l'incertezza nella misurazione dell'attributo, comune a tutti i dati, ai dati spaziali si aggiunge l'incertezza sulla misurazione della posizione spaziale¹⁹. Si consideri l'Esempio 3.1, in cui la posizione delle chiazze oleose sull'oceano è dedotta da foto inviate da un aeromobile dotato di apparecchiature fotografiche e GPS. In parte, l'incertezza è legata alla modalità di campionamento (la rotta), in parte alla scala (l'altitudine del volo e le caratteristiche della macchina fotografica), in parte dalla variabilità della superficie oceanica, condizionata da venti e correnti. La risoluzione di un'immagine satellitare condiziona la precisione e l'accuratezza dei dati da essa desunti: minore è la risoluzione, maggiore è l'incertezza localizzativa del dato.

Dipendenza spaziale. L'esempio del campionamento regolare su transetti è forse quello che meglio illustra la dipendenza spaziale introdotta dal campionamento. In quel caso, infatti, ogni osservazione è legata, ossia "dipende" dalla precedente, e la probabilità di avvistare una volpe è condizionata anche dal fatto che la persona ha già camminato sul percorso immediatamente precedente. Le osservazioni lungo il percorso sono quindi legate da dipendenza spaziale²⁰. Va da sé che all'aumentare della distanza tra le osservazioni aumenta il tempo intercorso tra i rispettivi passaggi in quei punti, e quindi decresce la dipendenza tra le osservazioni, secondo la logica espressa informalmente dalla legge di Tobler e rappresentata da una funzione di attrito della distanza. Se ciò è evidente nel campionamento a transetti, forme di dipendenza spaziale sono comunque indotte dal campionamento, si può dire "per costruzione". In ogni caso, quindi, le osservazioni su un processo spaziale sono affette non solo dalla dipendenza ed eterogeneità spaziale intrinseche al processo, ma anche da quelle indotte dal campionamento. All'atto dell'analisi sarà quindi difficile (pur se possibile) districare la porzione di tali proprietà intrinseca al processo da quella indotta dal campionamento, che comunque influisce sulla rappresentatività dei dati spaziali e sulle proprietà delle analisi successive.

¹⁹ Come già notavano Burrough e McDonnel (1986).

²⁰ Si veda ad esempio Dale e Fortin (2014).

Scala e MAUP. Qualsiasi schema di campionamento spaziale è legato a varie “scale”: quella su cui si estende il fenomeno, quella d’indagine e quella scelta per il campionamento dei dati, ciò che rimanda al problema dell’unità areale modificabile²¹. Inoltre, è importante ammettere che la conoscenza del fenomeno d’interesse (es. presenza volpi, incidenza di malattia, fertilità del suolo) che si ottiene dai dati è non solo necessariamente limitata, ma dipende anche dal campione. Se infatti, l’indagine fosse ripetuta, scegliendo uno schema di campionamento diverso, il fenomeno potrebbe manifestare caratteristiche diverse, legate alla differenza nei campioni e alla variabilità spaziale del fenomeno considerato, come studiato fin dagli anni ’70 dalla geostatistica²². È evidente come l’incertezza indotta dalla scelta del campione si somma all’incertezza determinata dalla scelta di una scala di analisi e dall’aggregazione spaziale²³. La scala del campionamento e della misurazione influenzano quindi i dati utilizzati nell’analisi, cosicché dati campionati a scale diverse o con schemi diversi, o acquisiti da unità spaziali diverse non sono necessariamente confrontabili. Analisi condotte su dati a scale diverse o a modalità di aggregazione diversa conducono potenzialmente a risultati diversi, come discusso nel capitolo 2 (*MAUP*).

Incongruenze tra unità spaziali. Il livello di aggregazione, o *risoluzione* dei dati influisce su qualità e incertezza dei dati anche in altri modi: si pensi ad un’indagine sull’incidenza di una malattia, in cui si usino dati clinici ottenuti alla scala (o risoluzione) dei distretti sanitari, dati sulla popolazione ottenuti alla scala delle unità censuarie, e dati d’inquinamento dell’aria ottenuti da stazioni di monitoraggio sparse sul territorio. I confini delle unità di censimento non coincidono con quelli dei distretti sanitari, e *buffers* tracciati attorno alle stazioni di monitoraggio dell’aria disegnano una maglia di unità spaziali ancora diversa. In tali – frequenti – situazioni, in cui i dati non collimano nello spazio, il loro utilizzo in analisi multivariate soffre di un errore indotto da questa incertezza localizzativa (*spatial mismatch* e *misalignment*).

²¹ Si veda in proposito il capitolo 2, e in particolare il *MAUP* (*modifiable areal unit problem*).

²² Si veda Cressie (1989). Obiettivo della geostatistica, definita nel capitolo 4, era la stima dell’entità di giacimenti minerari sotterranei in base a campioni necessariamente limitati (in quanto costosi) ottenuti da trivellazioni eseguite in superficie.

²³ Si veda l’analisi che ne fa Openshaw (1983).

Come si è già osservato si tratta di errore spaziale, che si somma e s'intreccia alle caratteristiche intrinseche dei dati spaziali, per cui si rende necessario l'impiego di metodi esplicitamente spaziali²⁴.

Incertezza temporale. A tutto ciò si aggiungono errore ed incertezza temporali. Si pensi ad indagini statistiche effettuate a cadenze diverse. Per indagini spazio-temporali si può porre il problema della validità del raffronto tra indagini successive, se, durante il loro intervallo temporale, vengono mutate le definizioni delle unità spaziali, che periodicamente sono rivedute²⁵.

6.2.2. Dati spaziali, processi stocastici, errore e incertezza

I dati spaziali possono essere ottenuti in vari modi, tra cui (ma l'elenco non è esaustivo) la collezione manuale sul campo, secondo vari schemi di campionamento, la ricezione di dati d'archivio (per esempio l'ISTAT), l'acquisizione di immagini telerilevate, fino alla lettura dei diversi tipi di dati *open source* disponibili sul web.

Nella gran parte di queste situazioni, il ricercatore non ottiene una serie di dati pronta per l'analisi spaziale, ma dati grezzi, che richiedono manipolazioni atte a trasformarli in forme utilizzabili in analisi spaziale. Le procedure necessarie variano secondo il tipo di collezione, l'origine del dato e il grado di elaborazione richiesto. In ogni caso si tratta di operazioni frequentemente condotte in ambito GIS, utilizzando varie funzioni di tali sistemi. Si osservi che in tutti i casi in cui il ricercatore (che si appresta a condurre un'analisi spaziale) riceve dati (ovvero non è egli stesso a curarne la collezione diretta sul campo) è probabile che questi siano già stati, almeno parzialmente, elaborati da altri, e quindi le operazioni preliminari di tipo GIS possono già essere state condotte. La trasparenza con cui tali operazioni vengono condotte e riportate varia a seconda del tipo di dati e di chi li ha curati: le informazioni salienti dovrebbero trovarsi nei cosiddetti "metadati", che dovrebbero accompagnare i dati.

Chiunque svolga le operazioni preliminari di GIS, gli errori contenuti nei dati vengono propagati tramite tali *operazioni* secondo

²⁴ Una discussione in proposito si può trovare in Liu e Bertazzon (2016).

²⁵ Si veda ad esempio www.istat.it/it/archivio/222527.

modalità che dipendono dalle funzioni, dall'errore, e dal dato, come è da lungo tempo riconosciuto nella letteratura GIS²⁶. Frequentemente l'errore umano che viene commesso nella fase di immissione dei dati nel sistema, quando l'operatore può commettere errori di posizionamento, di calcolo, di unità di misura, di banale distrazione. In generale, quanto più i processi vengono automatizzati, tanto minore è la probabilità di commettere errori. Tuttavia, l'inserimento automatico non elimina i problemi di incertezza localizzativa, scala e differenze tra le unità spaziali. Come si è sostenuto altrove²⁷ l'automazione del sistema non può sostituire la perizia di chi lo usa, e in tutti i casi è importante che i compiti automatizzati vengano verificati coscienziosamente da un operatore preparato.

Esistono poi una serie di errori derivati dall'analisi e da un uso incorretto della logica²⁸. Molte operazioni analitiche (elementari o complesse) di GIS possono produrre errore come, ad esempio, la sovrapposizione di mappe e l'utilizzo delle successive intersezioni. L'interpolazione è un'altra comune operazione che consiste nello stimare il valore di un attributo in un punto ove non sia stato rilevato, sulla base del valore di tale attributo nei punti vicini, ove esso sia stato rilevato. Per definizione l'interpolazione contiene un elemento di approssimazione, e anche se si utilizzano algoritmi sofisticati, l'interpolazione induce una regolarità artificiale. Nell'applicazione di tale tecnica sono inoltre possibili contraddizioni logiche: per esempio, data la densità di popolazione di un quartiere fiorentino a sinistra dell'Arno e quella di un quartiere a destra dell'Arno, è facile ottenere la densità di popolazione nel mezzo del fiume. Il problema non è tanto l'invalidità di quella particolare stima, quanto un suo eventuale utilizzo in analisi successive, ad esempio in un'analisi *location/allocation*²⁹ che se ne serve per definire la superficie di domanda per il posizionamento di un nuovo ospedale: in tal caso l'analisi finale definirebbe la posizione del

²⁶ Si vedano in proposito Heuvelink *et Alii* (1999); Wan *et Alii* (2015).

²⁷ Bertazzon (2005).

²⁸ Gong e Mu (2000, p. 189): «Logical error refers to the inconsistency of relationship among different features presented in a database. It is usually manifested through other types of errors. Thus, logical relationships of mapped features can be checked for error detection. Positional error has been».

²⁹ Per una definizione operativa si può consultare <https://pro.arcgis.com/en/pro-app/latest/help/analysis/networks/location-allocation-tutorial.htm>

nuovo ospedale in modo tale da servire anche gli abitanti incorrettamente stimati nel mezzo del fiume.

Vi è quindi l'aspetto importante della propagazione dell'errore tramite operazioni analitiche³⁰. Si pensi ad una funzione, espressa in termini matematici molto semplici, dalla formula: $y = f(x)$. Funzioni analitiche spaziali complesse possono contenere diverse operazioni matematiche successive, per esempio somme, prodotti, logaritmi e derivate. In questi casi ogni operazione matematica successiva agisce sul dato, ma se questo contiene un errore, la funzione agisce su entrambi. Se, nell'esempio, la variabile x è stata misurata con errore, essa contiene una componente ε , e la funzione $y = f(x)$ in realtà sarà una funzione di entrambe le componenti $y = f((x \pm \varepsilon))$, con effetti su y difficili da prevedere e diversi a seconda dell'operazione³¹.

È quindi evidente che i vari tipi di errore e di incertezza che affliggono i dati interagiscono nel corso del processo che, da un'osservazione sul campo, porta ad un'analisi statistica spaziale. Le varie forme di errore ed incertezza vengono propagate e s'intrecciano con modalità difficili da seguire e rintracciare lungo le diverse fasi. È certamente possibile che circostanze fortunate portino alla cancellazione o perlomeno alla riduzione di componenti di errore o incertezza, come è anche possibile che circostanze meno fortunate portino all'accumulo o all'acuirsi delle loro conseguenze. In taluni casi, ponendo molta cura nelle operazioni di inserimento e manipolazione dei dati, è possibile ridurre l'errore nelle primissime fasi di un progetto analitico, ma raramente è possibile correggerlo nelle fasi successive. In ogni caso, quindi, è importante documentare in maniera accurata e trasparente tutte le fasi del processo, senza cercare di mascherare la presenza di errori di cui si sia a conoscenza. Per quanto esistano tecniche in grado di assistere il ricercatore nell'identificazione dell'errore, è sempre buona norma: comprendere a fondo il fenomeno in esame e i dati che lo rappresentano; eseguire ogni passaggio analitico, anche il più banale, con competenza e cura; esaminare i risultati – anche parziali – con attenzione e senso critico.

³⁰ Si veda Zhang *et Alii* (2016); Cressie e Kornak (2003).

³¹ Si veda anche il capitolo 2.

I dati spaziali: rappresentazione di processi stocastici. Infine, è evidente come, anche per tutte le ragioni legate all'incertezza e all'errore dei dati, sia importante che l'analisi quantitativa sia condotta con metodi statistici, che rimangono i più adatti a gestire, quantificare e, se possibile, minimizzare l'errore nell'analisi. L'incertezza, oltre che il risultato di errore, è un elemento intrinseco dei dati spaziali. Lo è perché di essi non si conoscono perfettamente localizzazione e caratteristiche; perché i fenomeni geografici non si risolvono in un'unica localizzazione, ma possono essere mobili o estendersi su scale diverse; perché la realtà geografica è rappresentata da dati, e i dati sono, come si è visto, incerti; ma, soprattutto, perché i fenomeni geografici sono per natura incerti. Si pensi ad un fenomeno come il tempo meteorologico: forse il fenomeno incerto per antonomasia. Si estende sulla superficie della terra; è mobile, perché i sistemi meteorologici si spostano, spinti dai venti e dai moti planetari; si estende su più scale, perché fenomeni locali solo legati a fenomeni di più vasta scala. Per tutte queste ragioni è difficile prevedere dove poverà domani o quando poverà su Milano. Soprattutto, è difficile prevedere dove o quando poverà perché non abbiamo ancora compreso perfettamente i meccanismi che regolano i fenomeni meteorologici. Chiamiamo quindi incertezza la parte del fenomeno che sfugge alla nostra comprensione.

Se, dunque, una parte dell'incertezza proviene dai dati, parte proviene anche dai limiti della nostra comprensione dei fenomeni. Per questo, sin dall'inizio abbiamo definito i fenomeni geografici come processi *stocastici* spaziali, definizione su cui ci siamo soffermati nel capitolo 2 e che enfatizza come l'incertezza non sia solo un accidente della misurazione, ma anche una parte intrinseca dei fenomeni, associata alla loro variabilità e ai limiti della nostra capacità di comprenderli. Le proprietà intrinseche dei processi spaziali, ossia dipendenza ed eterogeneità spaziale, sono alla base dei cosiddetti effetti spaziali, che conducono ad un incremento di varianza, e quindi di incertezza, nelle stime dei modelli analitici spaziali. Errore ed incertezza associate ai dati e alla nostra comprensione dei fenomeni, e comunque propagate tramite le procedure analitiche in GIS, vanno quindi a confluire nell'errore dei modelli, aumentandone l'incertezza. Tutto questo rafforza la necessità di una cornice stocastica e di un apparato statistico spaziale finalizzato alla conoscenza, misurazione e minimizzazione dell'errore spaziale.

6.3. La cassetta degli attrezzi dell'analisi spaziale

Ancora una volta è la storia delle discipline sorelle a fornire un utile punto di partenza per una discussione critica di alcuni aspetti relativi al *software* analitico spaziale e alle sue implicazioni. Dal dibattito che vide emergere la *GIScience* emerse una pluralità di posizioni tra la scienza, la fabbricazione e l'uso degli attrezzi GIS, e in quel contesto si discusse la generale incapacità o impreparazione dei geografi a costruire quegli attrezzi³². La situazione dell'analisi spaziale è un po' diversa, e nel seguito si cercherà di comprendere il coinvolgimento di chi fa analisi spaziale nella fabbricazione dei propri attrezzi. Come si è visto nel capitolo 4, il termine analisi spaziale o geospaziale, continua a designare sia quella che abbiamo definito analisi spaziale elementare sia l'analisi statistica spaziale, benché anche dal punto di vista dell'esecuzione (via *software*) vi siano importanti differenze tra le due. Nel seguito del paragrafo, col termine "analisi spaziale" ci riferiamo all'analisi spaziale statistica o complessa.

Le analisi spaziali elementari possono essere eseguite attraverso il menu a cascata dei pacchetti standard di GIS: un singolo "click" apre una "finestra" che conduce l'utente direttamente all'esecuzione della funzione oppure lo guida attraverso alcune ulteriori scelte e quindi lo conduce indirettamente al risultato; in entrambi i casi, l'operazione si esaurisce in pochi passaggi, a conclusione dei quali il *software* restituisce il risultato dell'operazione. L'utente può scegliere entro una gamma data di parametri, che definiscono i dettagli dell'operazione, quali la larghezza del buffer o le unità di misura. Di norma il risultato di queste operazioni è la definizione di nuove entità geometriche, la modifica di entità esistenti oppure la selezione di un sottoinsieme di entità; in tutti questi casi, le entità risultanti dall'operazione ereditano le proprietà di quelle originali, per cui la validità dell'operazione analitica è garantita dalle proprietà degli oggetti utilizzati come *input*³³.

Viceversa, le operazioni di analisi statistica spaziale richiedono una serie di compiti più complessi, sia dal punto di vista computazionale che da quello concettuale: la formulazione di ipotesi, la definizione di relazioni funzionali, la rappresentazione dei fenomeni e dello spazio

³² Si vedano, in proposito: Wright, Goodchild, Proctor (1997); Pickles (1997).

³³ Queste operazioni sono state introdotte nel capitolo 4 e illustrate in Fig. 4.1.

in cui si svolgono, oltre a richiedere, frequentemente, l'uso di strumenti matematici e statistici relativamente complessi. Si consideri ad esempio una regressione spaziale multivariata³⁴. Assumendo che tutti i dati siano già stati raccolti, ripuliti ed omogeneizzati, una tipica sequenza di operazioni potrebbe essere sintetizzata secondo la Fig. 6.1.



Fig. 6.1 - Sequenza dei passaggi di una regressione spaziale

Benché più complesso delle operazioni analitiche spaziali elementari, anche un procedimento analitico spaziale di questo tipo può essere ridotto ad una sequenza meccanica analoga allo schema della figura³⁵. Tuttavia, alcuni dei passaggi più delicati, come la definizione del modello o l'implementazione di correttivi nel caso di violazione delle ipotesi [rappresentato nella figura dalla freccia tratteggiata], richiedono un giudizio, e la procedura capace di condurre a risultati *validi* non sempre può essere determinata *a priori*. Diversamente dalle operazioni elementari, alcune delle scelte dell'utente avranno un impatto pesante

³⁴ Per una definizione dell'analisi si veda il capitolo 2, e per le applicazioni si vedano l'Esempio 3.1 e il capitolo 7.

³⁵ Si veda la trattazione proposta da ArcGIS Pro: <https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-statistics/regression-analysis-basics.htm> (visitato il 31/01/2022)

sui risultati analitici, influenzandone il significato e la validità. In particolare, date le scelte compiute, saranno le fasi concettuali relative alla definizione e valutazione del modello a determinare il significato e la validità dei risultati attendibili. Il prodotto finale di un modello analitico spaziale non è un'entità geometrica, ma un insieme di parametri con precise proprietà statistiche, che non si limitano ad ereditare le proprietà delle variabili che sono state immesse (*input*) nel modello. Si potrebbe affermare che lo svolgimento di analisi statistica spaziale richieda maggiore competenza statistica e matematica di quanto richiesto per svolgere operazioni elementari.

Le analitiche spaziali che utilizzano tecniche di *machine learning* si situano in una posizione ancora diversa: esse poggiano invero su modelli analitici di complessità anche maggiore dell'analisi statistica spaziale. Tuttavia, per ragioni che comprendono proprio la loro maggiore complessità ma anche per la filosofia che ne regge l'utilizzo e la frequente applicazione ai *big data*, la pratica di queste analisi tende a prescindere dalla verifica di ipotesi sui dati e sulle loro relazioni, dal condurre analisi esplorative preliminari, e dal verificare le proprietà statistiche dei dati, optando invece per la semplice immissione dei dati, lasciando che sia poi la macchina a trovare la soluzione. Si tratta dell'utilizzo del modello come una *black box*, una scatola magica in cui entrano dati ed escono soluzioni, ove è frequente che l'utente non sia a conoscenza dell'apparato modellistico che utilizza³⁶. In qualche modo questa posizione costituisce un paradosso: operazioni per natura complesse vengono svolte per definizione in maniera meccanica. Paradossalmente, è quindi più *facile* svolgere analitiche spaziali che analisi statistica spaziale.

Va infine osservato che chi si appresta ad eseguire un'operazione analitica, sia essa elementare o complessa, si trova di fronte ad un compito tecnico o matematico, che per natura lo porta a focalizzare le questioni tecniche o matematico-statistiche, e non a soffermarsi sugli aspetti concettuali che stanno a monte, come la rappresentazione dei fenomeni, l'equità nella scelta del campione o le questioni etiche. Da questo sorge anche la necessità di un apparato critico, che sopperisca a questa tendenza applicativa, come discusso nel capitolo 5.

³⁶ Un esempio di modelli di questo tipo si può trovare in Mirzaei *et Alii* (2021).

6.3.1. *Fabbricare gli attrezzi dell'analisi spaziale*

Secondo una tradizione decennale, le operazioni elementari si eseguono in ambiente GIS tramite menù a cascata che guidano l'utente attraverso le fasi varie dell'operazione³⁷, mentre le analisi statistiche spaziali sono condotte tramite *software* di tipo statistico-analitico. Tra i comuni pacchetti commerciali di statistica³⁸, quello che per lungo tempo si è imposto, fornendo un'ampia gamma di funzioni analitiche spaziali è stato S-plus³⁹ scritto nel linguaggio di programmazione S, un linguaggio sviluppato a partire dagli anni '70 che si evolse come linguaggio *object-oriented* con spiccate capacità grafiche⁴⁰. Le versioni commerciali di S-plus che si susseguirono dagli anni '90 fornivano funzioni statistiche standard tramite i classici menu a cascata (*GUI - graphical user interface*), mentre l'estensione *S-plus spatial* consentiva di svolgere una serie di funzioni statistiche spaziali, che però non erano proposte tramite *GUI* e pertanto richiedevano l'esecuzione di comandi, spesso raccolti nelle sequenze di comandi dette *scripts*. Un esempio di *script* è presentato nell'Esempio 6.1. Era possibile integrare le operazioni di GIS con quelle di analisi spaziale con operazioni di import-export⁴¹.

L'ambiente R. Il linguaggio di programmazione R⁴² è molto simile ad S, tuttavia, mentre S fu impiegato in *software* commerciale, R viene pubblicato gratuitamente (*free*) e *open source* a partire dai primi anni '90. La modalità *open source* consente la massima trasparenza di esecuzione, poiché chiunque può accedere gratuitamente al programma

³⁷ Di questo ci siamo occupati in un articolo dal titolo eloquente: "L'arte dell'analisi spaziale: perché saper scegliere i parametri del modello se li puoi semplicemente selezionare dal menu a cascata" (Bertazzon, 2005).

³⁸ Tra i pacchetti commerciali notiamo innanzitutto, SPSS (www.ibm.com/products/spss-statistics), poi SAS (www.sas.com/), STATA (www.stata.com/) e Minitab (www.minitab.com/), tra gli altri.

³⁹ Si trovano anche le grafie Splus e S plus. Ora acquistato da TIBCO, se ne trova traccia su: <https://s-plus.software.informer.com/8.0/> e www.tibco.com/ (visitato il 31/01/2022)

⁴⁰ Si vedano Chambers (1998); Becker (1994).

⁴¹ Si veda Bao *et Alii* (2000).

⁴² «R, like S, is designed around a true computer language, and it allows users to add additional functionality by defining new functions. Much of the system is itself written in the R dialect of S, which makes it easy for users to follow the algorithmic choices made. For computationally-intensive tasks, C, C++ and Fortran code can be linked and called at run time. Advanced users can write C code to manipulate R objects directly». (www.r-project.org/).

(code), e questo a sua volta permette a chiunque di comprendere e verificare come vengono eseguite le singole operazioni, e anche di modificarle e di mettere a disposizione di altri le proprie modifiche o estensioni. Secondo il «Progetto R», R viene definito non come pacchetto, ma come *ambiente software*, le cui funzionalità vengono estese da pacchetti (*packages*)⁴³. Esistono vari pacchetti utilizzati per l'analisi spaziale, tra cui “*spdep*”⁴⁴ che contiene numerose funzioni per l'analisi della dipendenza spaziale e il calcolo dei relativi test e modelli, come illustrato dall'Esempio 6.1. Come osservato nel capitolo 1, tra gli odierni *software* per l'analisi statistica spaziale, l'ambiente R è probabilmente il più popolare, ma non l'unico⁴⁵.

L'indice di autocorrelazione spaziale di Moran è stato definito in termini matematici e statistici nel capitolo 2. Da chi lo utilizza, quindi, si presume una certa familiarità con quegli strumenti. Se, per calcolare l'indice, è sufficiente “cliccare” un'opzione da un menù a cascata, l'utente potrebbe ottenere il risultato anche senza una conoscenza minima di quegli strumenti matematici e statistici, mentre si potrebbe invece pensare che tale conoscenza sia richiesta per comporre e lanciare uno *script*. Tuttavia, data la natura *freeware* e *open source* di R e lo spirito della “comunità R”, si possono trovare online molti *scripts* per l'esecuzione di varie analisi, per cui a chi conduce l'analisi basta saper riconoscere lo *script* e modificarne i dettagli per adattarlo al proprio problema. Può sembrare poco, ma saper leggere, comporre e utilizzare stringhe di comando richiede capacità che vanno oltre la cassetta degli attrezzi di chi utilizza il pacchetto Office, pacchetti tradizionali di GIS,

⁴³ «Many users think of R as a statistics system. We prefer to think of it as an environment within which statistical techniques are implemented. R can be extended (easily) via *packages*. There are about eight packages supplied with the R distribution and many more are available through the CRAN family of Internet sites covering a very wide range of modern statistics». (www.r-project.org/).

⁴⁴ Si veda in proposito: <https://cran.r-project.org/web/packages/spdep/spdep.pdf>.

⁴⁵ Oltre ad esso, vogliamo brevemente ricordare *Geoda* e *GWR*, sviluppati rispettivamente da Luc Anselin e dal gruppo di Stuart Fotheringham. Anche *Geoda* (<https://geodacenter.github.io/>) è distribuito gratuitamente e contiene una serie di funzioni specializzate, comunicate attraverso un'interfaccia grafica attraente, che ne rende l'uso intuitivo e accessibile, interfacciandosi anche facilmente con dati sia in formato GIS (*shapefiles*) sia in formato tabulare (*csv*). È anch'esso open source, benché tramite l'interfaccia grafica il codice non sia visibile all'utente. Le ultime versioni di *GWR* (<https://sgsup.asu.edu/sparc/software>) sono distribuite gratuitamente e offrono numerose varianti della regressione *GWR*.

e altro *software* comunemente usato dai geografi. Questo ha alcune implicazioni importanti.

Da un lato, questa soglia, o bagaglio minimo di conoscenza, costituisce una barriera che tende a confinare al di fuori dell'analisi spaziale coloro che non ne sono (o non ritengono di esserne) in possesso, e questo punto sarà ripreso nel prossimo paragrafo.

D'altro canto, la capacità di capire, anche solo superficialmente, lo *script*, consente una comprensione dell'attrezzo utilizzato che è di gran lunga superiore a quella consentita dall'uso del menu. L'interfaccia attraverso cui l'operazione viene presentata, quindi, può favorire (nel caso dello *script*) oppure inibire (nel caso del menu cliccabile) la comprensione dell'attrezzo. Infine, se lo *script* richiede una soglia minima di conoscenza, il suo utilizzo prolungato e la sua applicazione in circostanze e per operazioni e dati diversi promuovono l'incremento e il consolidamento di quella conoscenza da parte dell'utente. Questa conoscenza, più profonda e matura rispetto a quella "soglia", consente all'utente anche di divenire sempre più padrone dell'attrezzo, capace di modificarlo e adattarlo e, in casi non rari, anche di migliorarlo. Potremmo così definire una gamma di utenti, che, secondo la loro perizia, varia da principiante a intermedio a esperto. Con una semplice metafora potremmo definire l'utente intermedio come un sarto, capace di tagliare e cucire gli *script* confezionando abiti su misura anziché acquistarli da un grande magazzino, ovvero svolgere le operazioni tramite click. L'utente esperto sarebbe quindi paragonabile quasi ad uno stilista capace di creare modelli e apportare migliorie al processo produttivo. Diversamente dal redditizio mercato della moda, questi utenti rimettono in rete i prodotti della loro creatività, a uso e consumo di altri utenti, favorendo così un circolo virtuoso.

L'intera gamma di utenti, compresi i principianti, non si limita ad utilizzare gli attrezzi di lavoro, ma li modifica, adatta e migliora in proporzione alla propria perizia. Benché generalmente incapaci di scrivere *software* "a tenuta industriale", rimangono pochi dubbi che questi utenti incarnino non solo l'utilizzatore, ma anche il fabbricante di attrezzi.

ESEMPIO 6.1. Autocorrelazione spaziale con stringhe di comandi

Uno script per il computo dell'indice di Moran (*spdep* in R)

```
#installazione
install.packages("spdep")
library("spdep")

#importazione dati
dati <- read.csv("../dati.csv")

#definizione della matrice dei pesi spaziali
xy<-SpatialPoints(dati$var1[,c("Long","Lat")])
var1_k2<-knn2nb(knearneigh(xy, k=2, longlat=T))
var1_k2_w<-nb2listw(var1_k2, style="w")
var1_k2_w

#I di Moran
moran.test(data1$var1, listw=var1_k2_w, alterna-
tive="two.sided")
```

- Il simbolo “#” designa commenti, ignorati dal programma;
- Le prime due stringhe installano e avviano *spdep*; la terza importa i dati. Di queste operazioni preliminari non compare *output*;
- Il terzo gruppo (1) legge le coordinate dei punti; (2) calcola i vicini più prossimi secondo $k=2$; (3) definisce la matrice dei pesi spaziali sull'attributo (Var1), e (4) elenca i valori della matrice ottenuta;
- L'ultima stringa calcola l'indice per (Var1), secondo la matrice definita sopra e precisa l'ipotesi statistica (“two-tailed”).

Risultato o *output* relativo al precedente script

```
Moran I test under randomization

data: (data1$var1)
weights: listw=var1_k2_w

Moran I statistic standard deviate = 4.317041, p-
value = 1.58135e-05
alternative hypothesis: two.sided
sample estimates:
Moran I statistic      Expectation      Variance
      0.4866721          -0.0217391      0.0138694
```

- L'output richiama dati, variabile e matrice dei pesi spaziali;
- Segue il valore del test Z corredato dal valore del p statistico;
- Infine, compaiono valore calcolato, atteso e varianza dell'indice.

È quindi importante ribadire che questo utente competente è pur sempre un geografo, che conosce gli strumenti computazionali discussi nel capitolo 2 e, sempre più, l'apparato critico esposto nel capitolo 4, e a queste conoscenze e capacità affianca dunque quella di fabbricare o almeno modificare i propri attrezzi di lavoro. Siamo così in una situazione molto diversa da quella discussa per il GIS sul finire del secolo scorso⁴⁶, ed è importante ribadire come questa importante competenza cresca anche in risposta allo stimolo dato dalla forma in cui gli attrezzi gli sono presentati e in conseguenza si amplifichi grazie alla condivisione gratuita sul web di attrezzi e risultati. Si potrebbe concludere che in questo modo si realizzino quel controllo sulle operazioni svolte e quella trasparenza dell'analisi evocati da Chris Brunsdon e Alexis Comber (2021).

6.3.2. Dall'elitismo al populismo: gli attrezzi come strumento per diffondere l'analisi spaziale

Ciascuno dei *software* GIS e analitici utilizza un'interfaccia diversa e un modo proprio di presentare analisi e risultati. Il leader del *software* GIS, Esri, ha ampliato la gamma delle funzioni analitiche spaziali specializzate disponibili attraverso l'interfaccia grafica dei suoi prodotti tradizionali, e contemporaneamente offre anche soluzioni che consentono ai professionisti – non semplici utenti del GIS – di compilare *scripts*, nel linguaggio di programmazione Python, nel pacchetto pensato proprio per loro, ArcGIS Pro⁴⁷, e anche pacchetti analitici specializzati come *Geoda* e *GWR* offrono interfacce relativamente *user-friendly*. Questi sviluppi sembrano suggerire che gli ambienti *software* di analisi spaziale e GIS tendano a convergere in un ambiente capace di attrarre anche l'utente inesperto, consentendogli tuttavia di crescere e sviluppare le proprie competenze. Persiste però una differenza tra il *software* di matrice analitica, più orientato alla statistica, e quello di

⁴⁶ Si veda il dibattito presentato nel capitolo 4, e in particolare i lavori di Wright, Goodchild, Proctor (1998) e Pickles (1997).

⁴⁷ Si veda www.esri.com/en-us/arcgis/products/arcgis-pro/overview e www.python.org/ (visitati il 1/2/2022).

matrice GIS, che tende a privilegiare gli aspetti visuali⁴⁸. Si confronti, in proposito, il procedimento per il calcolo dell'indice di autocorrelazione spaziale di Moran in R, illustrato nell'Esempio 6.1, con le immagini riprodotte nelle Fig. 6.2 e 6.3, che sintetizzano, rispettivamente, l'*output* relativo al computo al medesimo indice in ArcGIS⁴⁹, e la procedura per il suo calcolo in *Geoda*⁵⁰.

La Fig. 6.2 non illustra il metodo per ottenere l'indice in ArcGIS, ponendo invece attenzione al risultato, che è sintetizzato nella rappresentazione grafica del relativo test Z, corredato di colori e schemi per valutare l'associato *pattern* spaziale. Analogamente la Fig. 6.3 mostra, nel riquadro di sinistra, l'interfaccia per il calcolo della matrice dei pesi spaziali e, in quello di destra, il risultato, che qui è sintetizzato da un grafico cartesiano che confronta le unità spaziali con quelle definite "vicine" (*neighbouring*) dalla medesima matrice.

Presentare all'utente i risultati delle analisi in forma visuale, grafica, o cartografica le rende certo più attraenti e probabilmente più intuitive, e può così attirare più persone verso la pratica dell'analisi spaziale, anche se, privilegiare ciò che è grafico su ciò che è statistico, tende a diminuire l'enfasi sugli aspetti analitici e matematico-statistici dell'operazione, con il rischio di attirare utenti che non sono preparati a comprendere tali questioni.

⁴⁸ Nella primavera del 2020 gli studenti del corso di 'Advanced spatial analysis and modelling' tentarono di persuadermi che è possibile svolgere analisi spaziale in ArcGIS pro con comandi e risultati equivalenti a quelli di R. Ne parlai con alcuni colleghi, ricevendo l'impressione che le possibilità di ArcGIS pro fossero effettivamente molto vicine a quelle dei pacchetti analitici. Con riluttanza e molte precisazioni, consentii agli studenti che lo volessero di utilizzare ArcGIS pro per condurre il progetto richiesto dal corso. Rimasi delusa e amareggiata nello scoprire che il risultato prodotto da ArcGIS pro per una regressione *GWR* era *solo* un insieme di mappe, privo di qualsiasi tabella o testo che illustrasse coefficienti, indici e diagnostiche della regressione.

⁴⁹Si veda: <https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/spatial-autocorrelation.htm> (visitato il 31/01/2022). Si osserva che l'autocorrelazione spaziale è illustrata in termini grafici e calcolata tramite menù interattivi (*dialog box*). Infine, il risultato è presentato in termini grafici, benché corredato da parametri statistici.

⁵⁰ https://geodacenter.github.io/workbook/5a_global_auto/lab5a.html

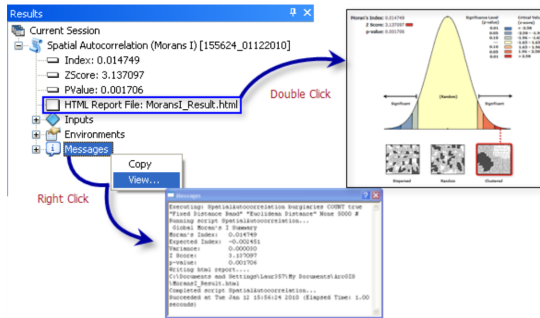


Fig. 6.2 - Output del calcolo dell'indice di autocorrelazione spaziale in ArcGIS

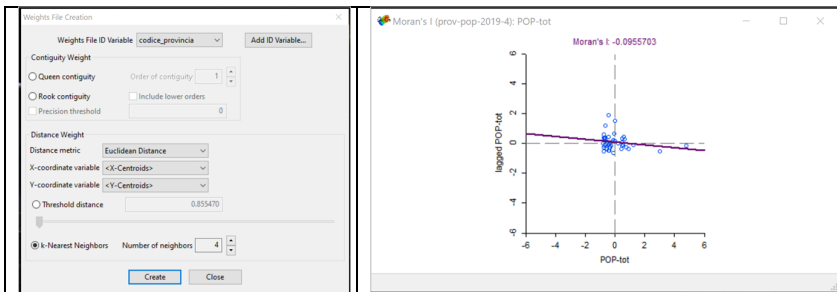


Fig. 6.3 - Procedura per il calcolo dell'indice di autocorrelazione spaziale in Geoda

L'odierna tecnologia dell'informazione geografica presenta enormi possibilità di avvicinare un largo pubblico di studenti, giovani, utenti casuali, quelle "masse" individuate da Michael Goodchild nella tensione tra una visione populista ed una elitaria della scienza dell'informazione spaziale⁵¹. Da sempre il *software* GIS è proattivo e propone operazioni di varia complessità tramite GUI (*Graphical User Interface*) e menù a cascata semplici e intuitivi. A questo si sono aggiunte più recentemente le diverse funzionalità reperibili online, a cominciare da Google Maps e Google Earth⁵², che potrebbero costituire un mezzo, o almeno un punto di partenza, per promuovere il pensiero spaziale, soprattutto nelle generazioni più giovani.

⁵¹ Goodchild (2000).

⁵² Si vedano www.google.com/, <https://earth.google.com/web/>

A questo proposito rimane interessante l'argomentazione di Michael Goodchild (2000) che identificava una tensione tra una visione populista del GIS, in cui la tecnologia è facile da usare ed accessibile a tutti, ed una visione elitaria, in cui solo coloro che conoscono bene i principi della teoria spaziale della scienza dell'informazione geografica sono in grado di usarla efficacemente⁵³. L'esperienza degli ultimi decenni insegna che la popolarizzazione del GIS, benché fondata su un approccio superficiale e tecnicistico, ha condotto, se non ad un'esplosione, almeno ad una significativa espansione delle applicazioni di GIS. Ciò può far pensare che analisi spaziale e GIS si trovino di fronte ad un'opportunità che non possono permettersi di perdere. C'è una grande domanda di applicazioni di analisi spaziale di alta qualità, in settori sempre più numerosi e rilevanti, dal cambiamento climatico alle migrazioni, alla gestione delle emergenze. Per rispondere adeguatamente a questa domanda non bastano i tecnici né gli entusiasti, sono necessari dei professionisti che sappiano pensare lo spazio in modo critico, integrando concetti, teorie, conoscenze statistiche e un approccio critico. Tutto questo non si può ottenere agitando lo specchio per le allodole della tecnologia, ma questa può essere uno strumento per avvicinare le masse, da cui potranno emergere dei professionisti, preparati attraverso un'adeguata formazione.

Non bastano certo i *big data* ad innescare una nuova rivoluzione quantitativa e nemmeno il *software*, gratuito o accattivante, basta a plasmare un futuro radioso per l'analisi spaziale e la geografia quantitativa, però entrambi servono a dare grande visibilità e opportunità che possono essere colte pienamente solo dalla geografia. Il *software* [analitico] per il GIS è un prodotto di considerevole valore commerciale, come lo sono molte delle sue applicazioni, dai database spaziali, alle analisi, alla cartografia. Il valore commerciale del *software* per la manipolazione dei dati spaziali è probabilmente un fattore importante per lo sviluppo di interfacce user-friendly, che possono servire ad espandere il mercato di quel *software* a più ampi segmenti di utenza. Paradossalmente, però, il *software* leader nell'analisi spaziale è un prodotto distribuito gratuitamente entro una comunità interessata al suo utilizzo

⁵³ Goodchild (2000, p. 8): «basic tension [in GIS] between the populist view, in which technology is easy to use and accessible to all, and the elitist view in which only those well versed in the principles of spatial theory and geographic information science are able to use it effectively».

e sviluppo. In questo paradosso sta una grande differenza tra l'analisi spaziale e il *big business* del GIS. Senza l'incentivo del profitto, forse, non sarà facile ottenere una significativa espansione dell'analisi spaziale e della moderna geografia quantitativa.

Rimarrebbe un elemento da discutere, nella cassetta degli attrezzi dell'analisi spaziale: l'*hardware*, ossia i computer e le varie macchine che consentono di far girare il *software*, di cliccare i menu o di far girare gli *scripts*. Di questo elemento, pur essenziale, non abbiamo trovato traccia di discussione nella letteratura. Ci limitiamo quindi a questa breve menzione e osserviamo che, quando affermiamo che l'analista spaziale è capace di forgiare i propri attrezzi, è solo a quelli *software* che ci riferiamo, ma questa capacità non sopperisce alla necessità di utilizzare macchine, che pure si acquistano sul mercato, anche se a prezzi generalmente accessibili, e che pure hanno un impatto sulle nostre scelte, sull'economia e sull'ambiente.

6.4. Chi è speciale nell'analisi spaziale

Spaziale è speciale, si è più volte ribadito, perché speciali sono i fenomeni e i dati che li rappresentano. Ma c'è ancora un elemento, troppo spesso trascurato, di cui vogliamo occuparci in questo capitolo, dedicato alla pratica dell'analisi spaziale. Ci riferiamo a *chi* fa analisi spaziale, ovvero l'elemento umano che pure ne è, forse, l'elemento più qualificante. Secondo una delle più classiche definizioni, un GIS sarebbe «una raccolta integrata di *hardware*, *software*, dati e personale (*liveware*) che opera in un contesto istituzionale»⁵⁴. Certo, nei 30 anni che ci separano da quella definizione molto è cambiato, ma gli elementi fondamentali del GIS, inteso anche come scienza, rimangono fondamentalmente i medesimi. E si tratta dei medesimi elementi che compongono l'analisi spaziale, definita come «il processo che genera conoscenza utile per rispondere a domande sull'ordine, la disposizione

⁵⁴ Maguire (1991, p. 17): «A Geographical Information System is best described as an integrated collection of hardware, software, data and liveware which operates in an institutional context».

e la struttura inerenti ai problemi spaziali»⁵⁵; e «benché l'analisi spaziale possa comprendere alcune funzionalità GIS, essa va ben oltre le possibilità offerte dall'impiego del GIS»⁵⁶. Chi compie questo processo, chi implementa queste funzionalità è la componente umana, il personale, che Maguire definisce con un termine insolito, *liveware*, termine che lo accomuna agli elementi *hardware* e *software*, ma da essi lo distingue su basi sostanzialmente biologiche se, diversamente dagli altri elementi, questo è *vivo*.

Molta della tecnologia odierna mette a disposizione potenzialmente a chiunque gli attrezzi per costruire la propria analisi spaziale (elementare), e per costruire la propria cartografia utilizzando gli *open data*, proprio come consente a tutti di scrivere un testo utilizzando la tastiera di un PC. Paradossalmente, lo stesso ambiente R è gratuito e aperto a tutti: chiunque, quindi, può procurarsi *open data* e condurre analisi spaziale statistica avanzata in un ambiente *software* che è *freeware* e *open source*. Questa possibilità teorica però è limitata dalla difficoltà stessa di condurre l'analisi programmandone autonomamente le funzioni. Ci troviamo forse di fronte a un paradosso: da un lato l'importanza del *liveware* nel GIS e nell'analisi spaziale, dall'altro la facilità con cui chiunque, in possesso di *hardware*, può procurarsi gli altri elementi necessari (*software* e dati) anche al di fuori di un contesto strutturato (scuola, azienda). È importante allora chiedersi quale sia la differenza tra chi costituisce quel *liveware* e chi si accosta indipendentemente all'analisi, e quali siano le differenze tra il prodotto analitico degli uni e degli altri.

La tecnologia odierna, come già il tradizionale *software* GIS con *GUI* e menù a cascata costituiscono un ambiente protetto e confortevole in cui l'utente sa che gli attrezzi a sua disposizione gli consentiranno di ottenere i risultati cercati: analisi e cartografia. Parafrasando una fortunata metafora di Umberto Eco (2004, pp. 165-166) sulle differenze tra Mac e DOS, quella tecnologia alla portata di tutti,

⁵⁵ Wolf e Murray (2016, p. 1): «Primarily, spatial analysis is the process of generating useful knowledge to answer questions about order, pattern, and structure inherent in spatial problems».

⁵⁶ Wolf e Murray (2016, p. 2): «Because spatial analysis is often strongly coupled with explicit GIS functionality, discussion of spatial analysis is sometimes equated to GIS operations. This is an incomplete view. While spatial analysis may include GIS functionality, it goes beyond capabilities offered through the use of GIS».

l'illusione dei dati aperti a tutti, sono pervasi da quel cattolicesimo secondo cui l'utente è sollevato dal fardello dell'interpretazione e dalla difficoltà della scelta, gli viene spiegato cosa fare e come farlo per raggiungere con certezza, se non l'eternità, almeno risultati validi, persino corredati da grafici animati e cartografia sontuosa. All'estremo opposto ci sono gli attrezzi analitici spaziali, soprattutto quelli *duri e puri* che, come l'ambiente R, non imbelliscono la statistica vestendola da immagine o cartografia: essi non fanno quella promessa, né, ovviamente, la mantengono.

Ciò accade perché l'analisi spaziale richiede ipotesi e scelte preliminari, verifiche di ipotesi ed esecuzione di calcoli anche complessi: non è facile promettere che tutto ciò approdi ad una soluzione, certa e valida. Chi svolge l'analisi non è un semplice utente, perché nelle fasi della definizione e dello svolgimento dell'analisi statistica spaziale gli si impongono scelte difficili e soggettive, si pretende che sappia giungere da sé a definire il modello ottimale. In altri termini, gli si richiede di praticare il pensiero spaziale, e anche di praticarlo in maniera critica, pena l'insuccesso, ovvero il mancato raggiungimento del risultato. Nello svolgimento dell'analisi spaziale, diversamente dallo svolgimento di operazioni analitiche elementari sotto la guida di un *software*, il *pensare spaziale* non è solo un utile accessorio, ma un elemento fondamentale.

Da queste considerazioni emerge la differenza tra un semplice utente e qualcuno che svolga consapevolmente l'analisi, proprio grazie a quel suo *pensare criticamente lo spazio*. Come è ormai evidente, non abbiamo un termine per definire questa figura: utente è certo inadeguato, e neppure *liveware* ci soddisfa. Dobbiamo accontentarci di generici analista o ricercatore, forse professionista. Questa figura va plasmata nelle scuole forgiando persone che conoscano profondamente l'analisi spaziale e siano capaci non solo di pensiero spaziale, ma anche di pensiero spaziale critico, e va poi nutrita tramite una cultura che per certi versi va contro la filosofia del tutto e subito.

6.4.1. Insegnare l'analisi spaziale

Attraverso le questioni delineate lungo questo volume si è implicitamente tratteggiata una figura complessa e sfaccettata del

professionista dedito allo studio e alla pratica dell'analisi spaziale. Il bagaglio dell'analista spaziale deve abbracciare i concetti matematico-statistici discussi nel capitolo 2, fino a quelli critici illustrati nel capitolo 4, e quelli tecnici discussi nei paragrafi precedenti di questo capitolo. Si tratta di una figura complessa, con una preparazione articolata, difficile da costruire in un corso universitario, se non, forse, in un corso di laurea strutturato. Le capacità richieste all'analista spaziale vanno costruite con una preparazione di base di tipo quantitativo e tecnico (*scripting*), una preparazione al pensiero critico, una preparazione a quel pensare spaziale che con esso va integrato nel *critical spatial thinking* definito da Michael Goodchild e Donald Janelle (2010). Pensare in maniera spaziale significa porsi la domanda “*dove*”, perché la risposta alla domanda, a questa domanda, potrà fornire risposte anche su *come, quando e perché*⁵⁷.

Dell'insegnamento della *GIScience* si è parlato, e si è visto come molte siano le circostanze che hanno contribuito a dar forma ai programmi universitari di GIS. Come quello della *GIScience*, il contenuto disciplinare dell'analisi spaziale è vasto e complesso. Se un suo nucleo importante è la moderna geografia quantitativa, col passare del tempo questo contenuto si va ampliando, incorporando sempre più contenuti relativi ai fondamenti, paradigmi, etica, critica, che pure hanno da sempre sostenuto le discipline dell'informazione geografica, ma solo in epoche relativamente recenti si sono imposti all'attenzione degli studiosi, e che solo faticosamente si stanno cercando un posto nei curricula dei corsi universitari. Mentre la materia cresce, il tempo per apprenderla rimane lo stesso, o forse diminuisce. Infatti, la disciplina si arricchisce di contenuti computazionali e tecnologici, che richiedono impegno didattico per l'insegnamento e l'apprendimento, mentre cresce l'esigenza di insegnare ed apprendere concetti e teorie. Così anche i corsi universitari sembrano rincorrere il tempo nel Paese della Regina Rossa, e chi li coordina si trova di fronte a scelte difficili tra l'impartire capacità operative (*skills*) oppure formazione teorica e concettuale. Della questione si è occupato recentemente *The Canadian Geographer*, in un numero speciale dedicato proprio a “*Teaching critical GIS/Teaching GIS critically*”. Come afferma Warren (2020) «ormai è

⁵⁷ Si visiti il Center for Spatial Studies dell'Università della California a Santa Barbara (<http://spatial.ucsb.edu/>).

quasi un cliché osservare che lo stesso termine “GIS critico” sfugge al tentativo di definirlo. Ancora più discutibile è la questione di cosa comporti insegnare il GIS critico, quando è stato osservato che “insegnare il GIS critico” e “insegnare criticamente il GIS” non sono necessariamente la stessa cosa. Come ricercatori che praticano il GIS, faticiamo a incorporare nei programmi prospettive teoriche che potrebbero alienare l’interesse degli studenti; come geografi critici, faticiamo ad accettare tecnologie intricate che minacciano di dominare l’attenzione degli studenti. Come docenti in sistemi accademici nordamericani, osserviamo queste questioni pedagogiche danzare sul palco scenico, davanti allo sfondo delle oscillazioni nel numero degli iscritti, le ridefinizioni amministrative e le crescenti strette di bilancio»⁵⁸.

Argomentazioni analoghe sono espresse da Matthew Wilson⁵⁹, che muove da considerazioni sui curricula delle università (seppure l’autore prenda in considerazione soltanto quelle americane), ove la questione verte sulla necessità, o quanto meno l’opportunità, che tali curricula contengano teoria e critica non meno che tecnica e pratica del GIS. E se l’autore, non da solo, argomenta in favore di questa duplicità di contenuti, egli stesso cita nel testo alcune valutazioni degli studenti, di segno opposto. Secondo tali valutazioni, sarebbero proprio gli studenti ad esprimere forte insoddisfazione nei confronti di corsi universitari di GIS in cui vengono insegnate critica e teoria che, oltre ad essere (a detta loro) di scarsa utilità pratica, sottraggono tempo prezioso al ben più rilevante insegnamento dei metodi e delle tecniche (*sic*)⁶⁰.

Nei guai sono quegli studenti (delle università americane) che vogliono imparare a fare le mappe senza perdere tempo a imparare il perché, senza chiedersi le ragioni, le implicazioni e i limiti del loro fare, il *cui prodest*. Il guaio di costoro è che non comprendono che

⁵⁸ Warren (2020, p. 647): « It is almost a cliché by now to observe that the term “critical GIS” skirts easy definition [...]. Even more contentious is the question of what teaching critical GIS entails, echoing questions raised by Elwood and Wilson (2017) who observe that teaching critical GIS and teaching GIS critically are not necessarily the same thing. As GIS practitioners, we struggle with the incorporation of critical theoretical perspectives into our curricula that do not alienate our students; as critical geographers, we struggle with reliance on intricate technologies that threaten to dominate our students’ attention. As university faculty within North American systems, we watch these pedagogical issues play out against a backdrop of changing enrolment patterns, administrative realignments, and increasing budget constraints».

⁵⁹ Wilson (2015; 2017).

⁶⁰ Wilson (2017).

queste distinzioni non hanno perduto il loro senso, che la linea del fare ormai non si distingue più dalla linea del criticare. Ai guai della mappa si aggiunge il fatto che ormai troppi, che non solo non hanno imparato teoria e critica, ma non hanno neppure imparato le tecniche, si improvvisano mappatori – non cartografi, ma dilettanti ingenui con l’illusione che la tecnologia da sola li renda capaci di fare⁶¹.

Abbiamo citato il *Dialog* del 2016 sulla geografia quantitativa nei curricula britannici, e il numero speciale del *Canadian Geographer* del 2020 sull’insegnamento critico del GIS. Non è poco. Sulla situazione in Italia si tornerà nell’ultimo capitolo, ma non si può tacere che da noi, proprio in questo periodo, la geografia è scomparsa dai curricula delle scuole superiori. Frattanto, il master in GIS dell’Università di Calgary ha ormai più di 20 anni, e sono in corso trasformazioni per renderlo più adatto ai tempi. Un adattamento modulare, che certamente segue diverse motivazioni, ma che forse contiene anche possibilità per introdurre moduli snelli anche nelle nostre università.

6.4.2. Potenzialità e limiti di chi fa analisi spaziale

Vi è una grande domanda di analisi spaziale di alta qualità, applicata a questioni che sorgono in settori importanti, dall’ambiente al clima, dalla geografia medica e sanitaria alla pianificazione urbana e regionale. Come si è visto, l’abbondanza di dati, la disponibilità di tecnologie accessibili, e la grande domanda di soluzioni ai problemi del nostro tempo potrebbero, assieme, costituire un’opportunità per lo sviluppo e l’insegnamento dell’analisi spaziale, anche nel nostro Paese, per la formazione di una figura professionale che sappia rispondere con competenza a questa domanda. Al momento, a quanto ci è dato sapere, non esistono purtroppo corsi universitari che preparino questa figura, intesa come qualcosa di distinto dal GIS e dalla geografia.

L’analisi spaziale, assieme al GIS e alla geografia quantitativa, possiedono una prerogativa forse unica, che è fornita loro dalla capacità di comunicare tramite i linguaggi dello spazio, della matematica e della geometria. Con questi strumenti linguistici la disciplina sa comunicare con le scienze sociali, ambientali, ingegneristiche, medico-

⁶¹ Wilson (2017); Bertazzon (2019).

sanitarie. Forse la differenza principale tra chi fa analisi spaziale e chi fa GIS sta nella fabbricazione degli attrezzi, che è una capacità, o forse una necessità, un'abitudine, che rimane specifica di questo gruppo, ac-comunato in questo a chi fa geografia quantitativa. La capacità di forgiare e affinare i propri attrezzi è importante, perché rende l'analista meno dipendente da altri, i costruttori di *software*, fornendogli la potenzialità di costruire analisi sempre più trasparenti e verificabili.

Essenziale per l'evoluzione della disciplina è il diffondersi di una cultura dell'analisi spaziale, di una consapevolezza delle tecniche disponibili e dei benefici che l'applicazione di tali tecniche può apportare: da un lato la superiorità delle tecniche specificamente spaziali rispetto ai metodi tradizionali, dall'altro la complementarità di tali analisi rispetto ad altri metodi per la comprensione dei processi spaziali⁶². Al fine di raggiungere tale cultura dell'analisi spaziale è necessario che un numero crescente di utenti e di professionisti sia disposto ad utilizzare tale insieme di tecniche, ma è pure necessario che tali tecniche siano applicate in maniera rigorosa e corretta.

L'analisi spaziale è un campo disciplinare relativamente giovane, in cui molte delle soluzioni attualmente praticate sono mutate da altre discipline⁶³: sono quindi soluzioni necessariamente limitate nello scopo e nel significato, talvolta così acerbe che mentre risolvono un aspetto del problema, ne inducono dei nuovi. Strategie di questo tipo potevano reggere durante l'infanzia della disciplina, ma da più parti oggi risuona l'appello di Michael Goodchild e Donald Janelle (2010) assieme a Daniel Sui (2001) affinché l'analisi spaziale sviluppi un corpus di teorie e tecniche sue proprie, capaci di approfondire l'analisi dei problemi e fornire, nel lungo termine, soluzioni a tutto campo, tali da aprire nuove frontiere di ricerca e da configurare una disciplina matura e indipendente. Si vede quindi che se da un lato l'analisi spaziale possiede una forza in più, la competenza e l'indipendenza di saper fabbricare i propri attrezzi, oltre la forza di saper comunicare oltre i propri confini, la sua debolezza rimane nello sviluppo di teorie proprie, nelle difficoltà di praticare la disciplina in maniera critica.

⁶² Guarrasi (2002).

⁶³ Berry (2004).

6.5. Verso la pratica

I dati non sono i fenomeni; come abbiamo detto in questo capitolo, sono solo cifre e simboli che utilizziamo per analizzare e comprendere i fenomeni geospaziali. Benché i dati, soprattutto spaziali e spaziotemporali, siano sempre di più e sempre più alla portata di molti, non vi è ragione di ritenere che per questo rappresentino meglio i fenomeni geospaziali. La quantità e la risoluzione spaziale dei dati, lungi dal risolvere i problemi analitici dei dati spaziali, rischiano di acuirli. Neppure gli strumenti e le tecnologie per la loro analisi e la rappresentazione, benché sempre più sofisticati, potenti ed accessibili, bastano a migliorare l'analisi dei dati geospaziali. Si può allora sottoscrivere l'opinione⁶⁴ secondo cui la crescente disponibilità di dati spaziali in grandi quantità e ad elevata risoluzione potrebbe addirittura costituire il presupposto per una nuova rivoluzione quantitativa in geografia?

Ci sono pochi dubbi che la grande e crescente abbondanza di dati possa costituire una grande opportunità per l'analisi spaziale e per la geografia quantitativa, ma se ci sarà una rivoluzione, non saranno certo i dati a farla. Il diluvio di dati geospaziali ad alta risoluzione pone delle sfide alla geografia quantitativa, che essa dovrà affrontare debitamente se vorrà cogliere quell'opportunità. La prima sfida, che forse contiene tutte le altre, è proprio quella di non lasciarsi sopraffare dai dati.

I dati sono importanti per comprendere i fenomeni geospaziali, soprattutto in termini quantitativi, ma è necessario comprendere non solo ciò che i dati sono, ma soprattutto ciò che essi non sono, quali ne sono i limiti, ed acquisire la consapevolezza che essi non possono sostituire né i fenomeni geospaziali né un'adeguata analisi spaziale. Colpita dalla crescente disponibilità di dati, che ancora non era un diluvio, la geografia quantitativa reagì inventando tecniche che potessero gestire grandi quantità di dati. Si trasformò verso una ricerca *data driven*, lasciando che fossero i dati a guidarla. E quando la pioggia si intensificò, lasciò che fossero i dati a decidere non solo la direzione della ricerca, ma anche le tecniche da applicare. Quando iniziò a piovere (dati), negli anni '90 del secolo scorso, la geografia quantitativa (e non solo) si lasciò incantare dalle sirene dei dati, e solo recentemente essa stessa

⁶⁴ Wyly (2014).

comincia a comprendere quanto sia rischioso cedere ai dati il posto di comando, consegnando loro le chiavi della ricerca.

Oggi viviamo una terribile realtà, la pandemia che ci spinge anche a pensare in maniera nuova a fenomeni che sono anche geospaziali. Quasi due secoli fa il Dr. Snow ebbe un'idea rivoluzionaria: pensare in maniera spaziale all'epidemia di colera, disegnare la mappa dei decessi. Il Dr. Snow, medico, non cartografo, s'ingegnò a rappresentare i dati epidemiologici su uno stradario. Senza i dati non avrebbe potuto segnare i decessi sulla mappa, ma non furono i dati a consentirgli di individuare l'origine del contagio: fu la sua capacità di pensare l'epidemia come un processo spaziale, fu la sua analisi dei dati epidemiologici. Rivoluzionario fu il suo pensiero spaziale.

Durante la pandemia da COVID-19 *dashboards* e cartografie hanno conquistato le prime pagine dei giornali e i vertici dei social media, diventando essi stessi *virali*: *l'analisi geospaziale è diventata virale*. Essa può insegnarci molto su questa pandemia, e questa pandemia può insegnarci molto sulle potenzialità dell'analisi geospaziale. La geografia quantitativa non può accontentarsi di lasciare ai dati il comando della ricerca e il controllo dei metodi, ma deve tornare al posto di guida, trovando modi non già per aggirare i metodi statistici, ma per incorporarli nelle analisi, non già per tralasciare le questioni etiche, ma per affrontarle compiutamente. Tra le difficoltà che l'analisi spaziale deve affrontare, c'è che tutto ciò richiede tempo, impegno e investimenti di lungo termine. Forse le facili e redditizie *analytics* sono solo chimere che l'allontanano dal suo mestiere e dalla sua vocazione.

Questo prepotente, tragico fenomeno geospaziale sta svelando utilità, efficacia, fascino e limiti dell'analisi geospaziale, spingendoci ad affrontare discorsi che il GIS ha a lungo trascurato, se già nel 1997 Wright, Goodchild, Proctor affermavano che «i dibattiti della geografia sul GIS sono insolitamente esposti all'attenzione generale»⁶⁵. In questo senso, la situazione odierna non è molto diversa, e se non fosse bastato il diluvio dei dati, è la pandemia che ci costringe ad un impegno e a un dibattito che vanno ben oltre i dati. Un impegno cui non possiamo sfuggire, un dibattito che non può più rimanere *inter nos*.

⁶⁵ Wright, Goodchild, Proctor (1997, p. 346): « Of course geography is no stranger to methodological debates, and some of the arguments over GIS echo the arguments over quantification in the 1950s. [...]. Geography's debates over GIS are thus unusually exposed to general view».

7. L'analisi spaziale in azione: qualche esempio

Ὁ λόγος δηλοῖ ὅτι ...

Esopo, VI secolo a.C.¹

7.1. Premessa

Dopo aver discusso numerose questioni astratte relative all'analisi spaziale, in questo capitolo si propone una breve selezione di studi che illustrano come l'analisi spaziale possa essere impiegata, concretamente, per la comprensione e gestione del territorio. Ciascuna di queste ricerche racconta la storia di un'applicazione dell'analisi spaziale, la cui morale ci parla del territorio analizzato.

La letteratura offre numerosi studi applicativi²; qui si è scelto di presentare solo studi guidati dall'autrice, che rimane la sola responsabile di errori ed omissioni. La scelta e la presentazione degli studi è limitata da molti fattori, tra cui i vincoli alla riproduzione di articoli già pubblicati. Si sono scelti cinque studi, che si snodano in un percorso quasi circolare partendo dalla sanità pubblica per passare all'ambiente urbano, fino all'ambiente costiero per poi tornare alla sanità e portare esempi in ambito italiano. Il primo studio di sanità pubblica, già discusso nell'Esempio 2.1, illustra l'intreccio di processi spaziali sanitari e demografici; l'analisi individua quartieri a maggiore rischio sanitario, su cui si propongono interventi sociosanitari mirati. Segue uno studio a cavaliere tra sanitario e ambientale, che stima l'inquinamento atmosferico in ambiente urbano ad elevato dettaglio spaziale, per analizzare la relazione tra inquinamento e salute umana; questa

¹ «Questa storia insegna che...» (Esopo, 1996, p. 6).

² Basta scorrere gli indici di riviste come *Geographic Analysis*, *International Journal of GIS*, *JOSIS* tra le altre.

ricerca ha prodotto numerosi articoli e le stime prodotte da questi modelli sono state impiegate sia in studi epidemiologici sia per promuovere la scelta di zone meno inquinate per svolgere attività fisica all'aperto. Il terzo è uno studio ambientale, discusso anche nell'Esempio 3.1, che analizza la relazione tra inquinamento oceanico da idrocarburi e attività marittime; ne emerge una complessità che va oltre la narrativa del trasposto (merci e petroli) come unico responsabile dell'inquinamento dei mari, facendo emergere il ruolo delle attività di diporto: per questo si suggeriscono programmi di educazione ambientale che ne promuovano la consapevolezza per ridurre, nel tempo, l'impatto. Segue uno studio dell'inquinamento nella laguna di Venezia che, pur se datato, presenta un interesse metodologico e geografico: in ambito italiano presenta un intreccio di geostatistica e regressione spaziale per la stima della presenza di inquinanti nelle vongole che traggono nutrimento dal sedimento lagunare; oltre l'aspetto ambientale, la questione investe la salute umana e l'economia locale. L'ultimo studio, ancora in fase di elaborazione mentre scriviamo, presenta un'analisi descrittiva della mortalità durante la prima ondata dell'epidemia da COVID-19 in Lombardia e Veneto, le prime regioni in occidente ad esserne colpite: adattando strumenti analitici spaziali all'analisi spazio-temporale a grande scala consente di evidenziare dinamiche locali poco studiate; studiando le prime regioni colpite dall'epidemia si può comprendere la relazione tra mortalità e caratteristiche locali, traendone elementi utili a gestire futuri eventi epidemici.

Gli esempi sono presentati in maniera schematica, secondo una struttura uniforme che ne evidenzia ipotesi, dati, metodi e risultati, concludendosi con osservazioni sul ruolo dell'analisi spaziale e l'applicazione pratica dello studio.

La presentazione di ogni studio integra tre linguaggi: quello narrativo testuale, quello grafico e cartografico, e quello matematico e statistico: questa forma di comunicazione intende stimolare una riflessione sulle persistenti tensioni tra visuale e numerico, tra elitismo e populismo nell'analisi spaziale.

Nello spirito che pervade questo volume, questo capitolo intende anche promuovere la comprensione e l'impiego dell'analisi spaziale, attraverso studi in cui l'analisi non è solo rigorosa, ma anche critica e, ove possibile, trasparente e replicabile. In tutti i casi è precisato il *software* impiegato che, per l'analisi spaziale, è quasi esclusivamente

l'ambiente R. Purtroppo, in quasi tutti i casi, i dati utilizzati sono soggetti a vincoli etici o clausole contrattuali, che ne impediscono la pubblicazione. In larga parte è per questa ragione che si è inserito l'ultimo studio, che è interamente condotto su *open data* e *software freeware*, realizzando un'analisi che, come auspicato in letteratura, è completamente trasparente e replicabile.

7.2. Malattie cardiovascolari in contesto urbano

Questo studio illustra una linea di ricerca sull'incidenza di malattie cardiovascolari nella città di Calgary (Canada); avviato diverso tempo fa, continua a produrre studi, i più recenti legati anche all'ambiente³. L'incidenza della malattia è concettualizzata come un processo spaziale [Fig. 7.1(a)] ed analizzata in relazione a fattori di rischio non modificabili (variabili demografiche) e modificabili (variabili socioeconomiche) [Fig. 7.1(b)].

Ipotesi

- a) L'incidenza della malattia nella città presenta una distribuzione spaziale non casuale.
- b) La distribuzione spaziale dell'incidenza della malattia è associata a fattori di rischio non modificabili (variabili demografiche) e modificabili (variabili socioeconomiche).
- c) L'associazione spaziale tra all'incidenza della malattia e fattori di rischio modificabili può coadiuvare interventi e politiche di prevenzione sociosanitaria.

Dati

- 1) L'incidenza di malattia è misurata dalle cateterizzazioni cardiache effettuate nelle strutture sanitarie competenti per territorio. Si tratta di dati clinici soggetti ai vincoli del protocollo etico approvato. I casi clinici sono registrati al livello di codice postale di residenza

³ Bertazzon (2009); Bertazzon, Olson, Knudtson (2010); Liu *et Alii* (2020). Si veda anche l'Esempio 2.1.

dei pazienti⁴, aggregati al livello delle unità censuarie rappresentate in Fig. 7.1(a) e rapportati alla popolazione residente.

2) I fattori di rischio sono misurati da variabili censuarie⁵ alle unità spaziali rappresentate nelle Fig. 7.1(a) e 7.2.

Metodi

Sono elencati solo i metodi più rilevanti in relazione alle ipotesi e all'impiego di analisi spaziale, secondo una numerazione corrispondente a quella delle ipotesi. Le analisi statistiche spaziali sono condotte in S-plus⁶ e la cartografia in ArcGIS.

- a) Analisi autocorrelazione spaziale.
- b) Analisi di autoregressione spaziale (*SAR*) multivariata.
- c) Visualizzazione cartografica.

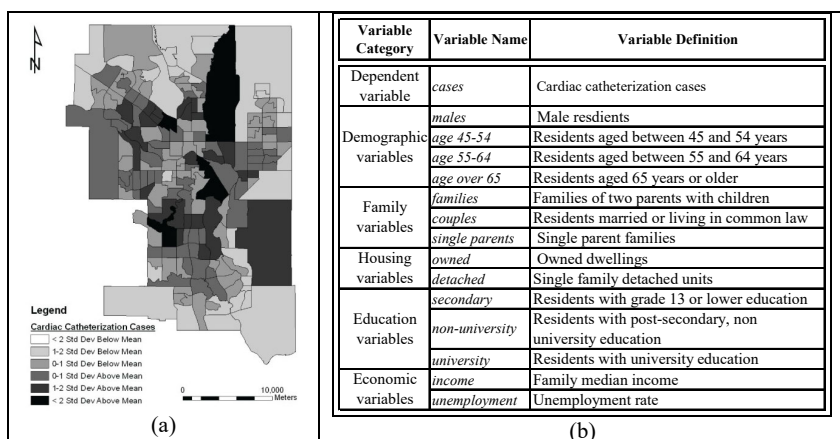


Fig. 7.1 - Distribuzione dell'incidenza di malattie cardiovascolari a Calgary (Canada) (a) e variabili demografiche e socioeconomiche considerate (b).

Riprodotta da Springer, Applied Spatial Analysis and Policy, 3:1-23.

⁴ Come notato nel capitolo 5, i *postal codes* corrispondono ad isolati nelle città del Canada. Si veda: www.canadapost-postescanada.ca/cpc/en/support/articles/addressing-guidelines/postal-codes.page?. Visitato il 3/02/2022.

⁵ www12.statcan.gc.ca/census-recensement/2006/dp-pd/index-eng.cfm. Visitato il 3/02/2022.

⁶ Questo perché il primo studio, illustrato in maggior dettaglio, è datato. Come discusso nel capitolo 6, le differenze tra S-plus e R sono minime, e la transizione molto semplice.

Risultati

- a) L'incidenza di malattia e le variabili demografiche e socioeconomiche manifestano autocorrelazione spaziale positiva significativa. ***L'autocorrelazione spaziale dell'incidenza di malattia è associata alla distribuzione residenziale della popolazione anziana (età >65) secondo un processo spaziale interattivo [Fig. 7.1(a) e 7.2(a)].***
- b) L'incidenza di malattia è significativamente associata alle variabili demografiche e socioeconomiche riassunte in Fig. 7.1(b).

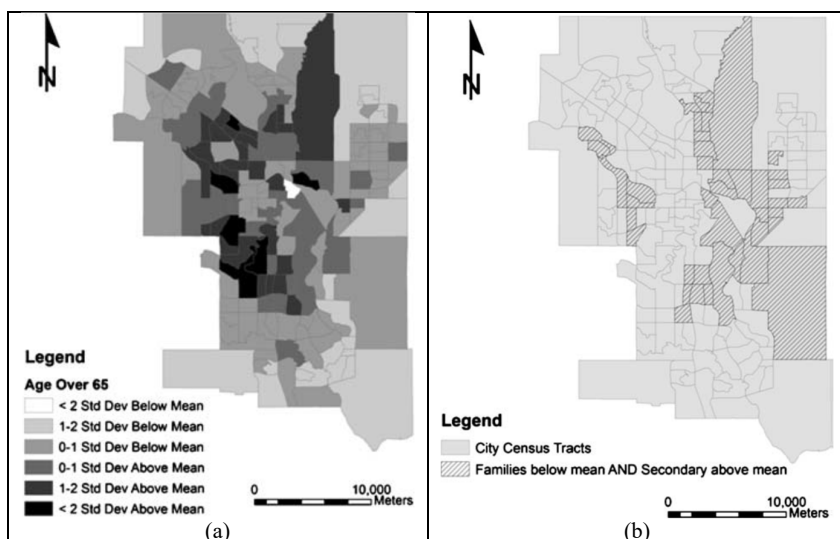


Fig. 7.2 - Distribuzione spaziale della popolazione over 65 (a) e aree con maggiore presenza di fattori di rischio modificabili, identificate tramite l'analisi spaziale. Riprodotta da Springer, Applied Spatial Analysis and Policy, 3:1-23.

Fattori di rischio modificabili e non modificabili. Un primo modello di regressione, M_1 , esprime l'incidenza di malattia (variabile dipendente) in funzione delle variabili (esplicative) sia demografiche che socioeconomiche. ***Risultano statisticamente significative quasi esclusivamente variabili demografiche⁷, tra cui domina***

⁷ Le variabili esplicative sono il sesso (maschio), le tre classi di età, e, unica variabile non-demografica, l'istruzione universitaria, unica variabile associata negativamente alla dipendente.

- età>65. M_1 presenta un'elevata *goodness-of-fit* ($R^2=0.79$) e l'autocorrelazione spaziale nel residuo non è significativa ($I=3.9e-03$). *Fattori di rischio modificabili*. Un secondo modello di regressione, M_2 , esprime l'incidenza di malattia in funzione delle sole variabili socioeconomiche⁸. **La significativa autocorrelazione spaziale nel residuo ($I=0.63$) richiede una specificazione autoregressiva spaziale.** M_2 presenta una *goodness-of-fit* inferiore ad M_1 ($pR^2=0.32$).
- c) **La localizzazione dei fattori di rischio modificabili, identificati da M_2 , indica zone a maggior rischio sociosanitario [Fig. 7.2(b)].**

Osservazioni e ruolo dell'analisi spaziale

- i. L'impiego di analisi esplorativa spaziale evidenzia autocorrelazione spaziale positiva e significativa sia nell'incidenza di malattia che nella distribuzione residenziale della popolazione anziana: ciò **suggerisce un processo spaziale interattivo**.
- ii. L'autocorrelazione spaziale della variabile dipendente (incidenza di malattia) si riversa nel residuo del modello di regressione, aumentandone la varianza e diminuendone l'affidabilità. **Questo si osserva in M_2 ma non in M_1 , dove età>65 'cattura' l'autocorrelazione spaziale della dipendente.**
- iii. Benché affidabile e con elevata *goodness-of-fit* (quasi 80%), M_1 non costituisce un utile strumento di supporto alle decisioni, poiché esprime l'incidenza della malattia quasi esclusivamente in termini di fattori di rischio non modificabili (età e sesso). **M_1 mostra che l'autocorrelazione osservata nell'incidenza di malattia è riconducibile alla distribuzione residenziale della popolazione anziana, tramite un processo spaziale interattivo.**
- iv. M_2 costituisce un utile strumento di supporto alle decisioni, poiché esprime l'incidenza della malattia solo in termini di **fattori di rischio modificabili che, da soli, spiegano oltre il 30% della variabilità dell'incidenza di malattia**. Trai fattori significativi, alcuni sono *problematici* (istruzione bassa e isolamento sociale), diversamente da altri (reddito alto e istruzione non-tecnica). **Senza**

⁸ I fattori di rischio modificabili contenuti nel modello sono: reddito (con segno positivo), istruzione tecnica (con segno negativo), istruzione media o inferiore (con segno positivo) e famiglie di due genitori con figli in casa (con segno negativo).

L'impiego del metodo autoregressivo spaziale l'analisi sarebbe inaffidabile, a causa dell'autocorrelazione spaziale nel residuo.

- v. Ulteriore *analisi spaziale individua le aree a maggiore presenza dei fattori di rischio problematici, segnalate all'autorità per eventuali interventi di politiche e prevenzione sociosanitaria.*

L'analisi spaziale si rivela utile per la conoscenza e la prevenzione del rischio sociosanitario legato all'incidenza di malattie cardiovascolari. L'analisi ne quantifica la distribuzione spaziale, specificando modelli affidabili per l'identificazione di fattori di rischio modificabili e l'individuazione di aree a maggior rischio sociosanitario.

7.3. Analisi spaziale dell'inquinamento atmosferico⁹

Questo esempio illustra una feconda linea di ricerca avviata con una collaborazione con *Health Canada*¹⁰. Obiettivo della ricerca è la stima dell'inquinamento atmosferico in ambito urbano a grande scala, o dettaglio spaziale. Obiettivo delle stime è informare modelli epidemiologici e di rischio, per comprendere gli effetti dell'inquinamento atmosferico sulla salute umana.

Ipotesi

- a) Lo studio degli effetti dell'inquinamento sulla popolazione umana e la stima di modelli di esposizione e rischio richiedono informazioni sull'inquinamento a grande dettaglio spaziale. La rete di monitoraggio dell'inquinamento gestita dai servizi ambientali¹¹ è troppo rada per servire a questo scopo (11 stazioni).

⁹ Bertazzon *et Alii* (2015); Bertazzon e Shahid (2017); Bertazzon e Couloigner (2018); Bertazzon, Couloigner, Underwood (2019); Bertazzon *et Alii* (2020); Bertazzon, Couloigner, Mirzaei (2021).

¹⁰ Si veda: www.canada.ca/en/health-canada.html. Visitato il 4/02/2022.

¹¹ Si veda <https://craz.ca/monitoring/calgary-southeast/>, sostanzialmente corrispondente agli ARPA italiani. Visitato il 7/02/2022.

- b) Una rete fitta e regolare¹² (50 stazioni nel 2010/11 e 100 nel 2015/16)¹³ e una doppia campagna di monitoraggio (2 settimane in estate e 2 in inverno) forniscono dati sufficienti per modelli LUR.
- c) L'inquinamento dell'aria e le variabili esplicative sono processi spaziali e i modelli sono soggetti agli effetti spaziali.

Dati

- 1) Campionamenti dell'inquinamento atmosferico elaborati in laboratorio da *Health Canada*. Il campionamento poggia su una rete di volontari disposti ad ospitare un sensore negli spazi esterni della propria abitazione (*VGI - volunteered geographic information*).
- 2) Variabili d'uso del suolo ottenute da fonti statistiche, dal comune e calcolate con metodi GIS.

Metodi

- a) I modelli LUR (*land use regression*) consentono stime molto dettagliate dell'inquinamento dell'aria: concettualizzano la quantità di inquinante nei siti campionati come funzione dell'uso del suolo in un intorno, tramite variabili quali la presenza di strade e traffico, la presenza di stabilimenti industriali, o la densità di popolazione.
- b) Modelli autoregressivi spaziali.
- c) Modelli *GWR*.

Risultati

- a) I modelli LUR producono stime affidabili dell'inquinamento a grande scala spaziale. Le mappe in Fig. 7.3 rappresentano le stime di biossido di azoto (NO₂) in estate e inverno¹⁴.

¹² Il numero di sensori era dettato dalla disponibilità di *Health Canada*; il loro posizionamento fu determinato da un modello *location/allocation* che fornisce la posizione ottimale di ciascun sensore in base alle informazioni disponibili sulle fonti di inquinamento e la distribuzione residenziale della popolazione. I siti indicati dal modello furono poi aggiustati per assicurare l'accessibilità del sito, l'accesso alla rete elettrica, il permesso di installare il sensore.

¹³ Una terza campagna di monitoraggio nel 2020-21 non ottenne il supporto di *Health Canada*, che preferì l'uso di strumenti analitici diversi. Il sopraggiungere della pandemia da COVID-19 ne avrebbe comunque impedito la realizzazione. Frattanto, il gruppo di studio acquistò una serie di sensori a basso costo per il monitoraggio del particolato fine (Bertazzon, Couloigner, Mirzaei, 2018). Nonostante risultati promettenti, anche questo studio fu rallentato dalla pandemia.

¹⁴ Queste mappe illustrano la campagna 2010-11. Per le mappe di NO₂ nel 2015-2016 si veda Bertazzon, Couloigner, Underwood (2019), e per quelle del particolato fine Bertazzon,

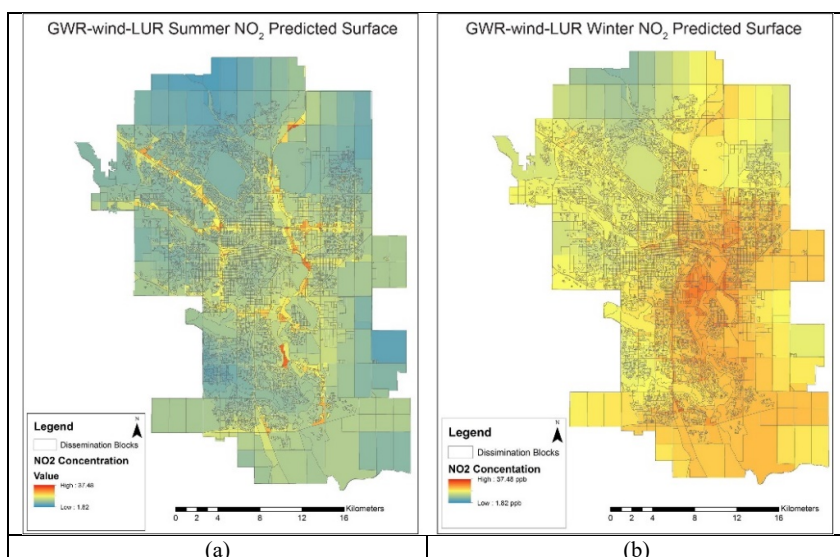


Fig. 7.3 - Stime LUR dell'inquinamento atmosferico da biossido di azoto a Calgary in estate (a) e inverno (b).

Riprodotta da Elsevier, Spatial and Spatio-temporal Epidemiology, 14-15: :9-21.

b) I modelli LUR presentano autocorrelazione ed eterogeneità spaziale nei residui. Come nello studio illustrato in 7.2, è possibile identificare una variabile [il vento dal nord] capace di spiegare l'autocorrelazione spaziale osservata nell'inquinamento. Il modello contenente questa variabile viene specificato in forma *GWR*, riducendone così anche la varianza indotta dall'eterogeneità spaziale.

Osservazioni e ruolo dell'analisi spaziale

i. Nelle mappe illustrate in Fig. 7.2 spiccano le principali arterie stradali, soprattutto in estate (a) quando l'inquinamento è minore; l'inverno (b) mostra l'inquinamento diffuso legato a riscaldamenti domestici ed emissioni industriali. Le stime di inquinamento poggiato sulle variabili LUR risultate significative nel modello.

Couloigner, Mirzaei (2021) e per quelle di numerosi inquinanti sull'hinterland di Calgary, Bertazzon e Couloigner (2018).

- ii. ***L'implementazione di modelli LUR da parte di un gruppo di geografi porta alla considerazione esplicita degli effetti spaziali***, con conseguente aumento dell'affidabilità delle stime.
- iii. La ricerca produsse numerosi modelli per diversi inquinanti. Le campagne di monitoraggio, ripetute dopo un intervallo di 5 anni, consentirono la ***definizione di modelli multi-temporali e spazio-temporali***¹⁵. Questi studi confermano la ***stabilità nel tempo della distribuzione (pattern) spaziale degli inquinanti e delle variabili significative*** nei modelli LUR.
- iv. Lo studio è incardinato sulla VGI, che garantisce la qualità dei dati, dimostrando l'interesse del pubblico per l'analisi¹⁶.
- v. ***Le stime di inquinamento furono successivamente utilizzate in studi sull'incidenza di diverse malattie*** (asma, malattie autoimmuni, malattie cardiache e intestinali).
- vi. ***Le mappe dell'inquinamento furono pubblicizzate, per comunicare al pubblico la distribuzione dell'inquinamento, favorendo la scelta di zone meno inquinate per l'attività fisica all'aperto.***

7.4. Inquinamento da idrocarburi e attività marittime lungo la costa Pacifica canadese¹⁷

Numerosi aspetti di questo progetto sono stati discussi nell'Esempio 3.1. Tuttavia, esso fornisce alcuni ulteriori spunti che meritano un accenno in questo contesto: l'impiego di regressione Poisson, la definizione di modelli regionali a partire da quelli locali e la valutazione dei test statistici spaziali su distribuzioni diverse dalla normale.

¹⁵ Per le stime di inquinanti diversi si veda Bertazzon e Couloigner (2018), per i modelli multi-temporali Bertazzon, Couloigner, Underwood (2019) e Bertazzon, Couloigner, Mirzaei (2021), e per quelli spazio-temporali Liu e Bertazzon (2016).

¹⁶ Per la campagna 2010-11, gli studenti volontari bussarono alle porte delle zone identificate del modello location/allocation chiedendo di poter installare un sensore. Viceversa, in previsione della seconda campagna (2015-16), lo studio fu pubblicizzato dagli autori inviando un volantino ai propri contatti e-mail (collegli, amici e parenti), in cui si chiedeva chi fosse disposto ad ospitare un sensore nel proprio giardino o terrazzo. Le offerte furono così numerose che si dovette/poté procedere ad una selezione in base alla localizzazione. Gli ospiti seguirono attentamente la campagna, segnalando subito eventuali malfunzionamenti. Per queste ragioni, la qualità dei dati fu superiore alla campagna precedente.

¹⁷ Bertazzon et Alii (2014).

Ipotesi

- a) La distribuzione delle chiazze oleose non è casuale.
- b) Esiste una relazione tra le chiazze oleose osservate e numerose attività marittime – oltre il trasporto di petrolio e merci – in prossimità della costa Pacifica canadese.
- c) I velivoli di sorveglianza documentano la presenza di chiazze oleose ma fungono anche da deterrente per scarichi illegali.
- d) La gamma di attività marittime varia secondo gli ambiti locali.

Dati

- 1) Per la variabile dipendente: dati forniti dal programma di sorveglianza marina NASP.
- 2) Per le esplicative: database di attività marittime curato da partner del progetto e variabili create in GIS (es. distanza dalla costa).

Metodi

- a) Analisi di autocorrelazione spaziale e *clustering* su *raster*.
- b) Regressione Poisson, con variabile *offset* (la “porzione sorvolata” viene rapportata alla dipendente).
- c) Regressione *GWR* Poisson.
- d) Impiego di *expert opinion* per l’interpretazione e valutazione delle analisi quantitative (tutte condotte in *R*).

Risultati

- a) La distribuzione spaziale delle chiazze presenta eterogeneità e autocorrelazione spaziali, anche in relazione al numero elevato di osservazioni nulle (0 chiazze).
- b - c) Numerose variabili sono significativamente associate alla presenza di chiazze; i residui presentano elevata eterogeneità spaziale.
- d) ***In base ai coefficienti calcolati da GWR si delineano 3 regioni e si ricalcolano modelli regionali***, illustrati in Fig. 7.4 e sintetizzati di seguito. 1) Nord: imbarcazioni da diporto, distanza dalla costa, navi passeggeri; 2) Stretto di Georgia: porti turistici, imbarcazioni da diporto, rimorchiatori, pescherecci, distanza dalla costa; 3) costa sudovest dell’isola di Vancouver: navi da carico, pescherecci.

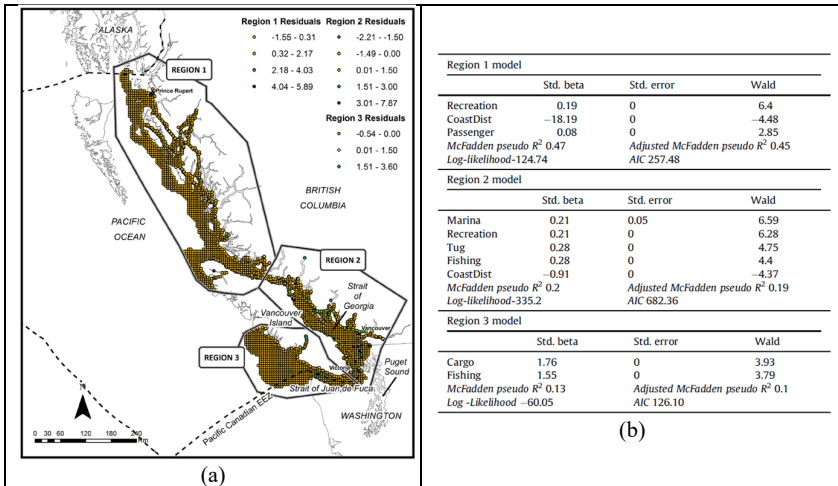


Fig. 7.4 - Regioni della costa pacifica canadese delineate in base all'analisi locale (a); modelli regionali di regressione Poisson (b).
 Riprodotta da Elsevier, Applied Geography, 52: 78-89 con licenza CC-BY-NC-ND.

Osservazioni e ruolo dell'analisi spaziale

- i. La significatività delle proprietà spaziali (eterogeneità e autocorrelazione) è valutata da test statistici che assumono normalità, rendendone problematica la valutazione in una distribuzione Poisson.
- ii. La dimensione dell'area di studio e delle celle del *raster* influisce sui risultati statistici (*MAUP*).
- iii. L'eterogeneità spaziale gonfia l'errore dei risultati globali. L'analisi locale, *GWR*, produce un numero tanto elevato di coefficienti locali da complicarne l'interpretazione. **L'analisi quantitativa e qualitativa dei coefficienti locali consente l'individuazione di regioni omogenee su cui definire modelli regionali**, trascendendo l'approccio quantitativo col coinvolgimento di *expert opinion*.
- iv. Le variabili significative nelle diverse regioni suggeriscono, **contrariamente all'opinione comune**, che l'inquinamento diffuso lungo la costa non sia ascrivibile al solo trasporto merci (navi da carico), ma anche ad attività turistiche e di diporto. **Col supporto degli esperti si ritiene che elementi come l'imperizia nel maneggiare i carburanti, la presenza prolungata di imbarcazioni vecchie (impregnate di carburante) nei porti turistici e lente ma continue perdite dai serbatoi in acqua e lungo la costa**

contribuiscono all'inquinamento marino in prossimità della costa. Queste forme di inquinamento, nel tempo, possono essere mitigate da adeguati programmi di educazione ambientale.

- v. L'impiego della variabile *offset* consente una *considerazione esplicita del ruolo di deterrenza dei velivoli di sorveglianza, chiarendo che il modello non spiega azioni illegali* (risciacquo dei serbatoi in mare) che vengono evitate in presenza del velivolo.
- vi. *I modelli elaborati in questo studio furono successivamente impiegati per stimare l'inquinamento in zone non sorvolate dai velivoli, e le stime servirono a valutare l'inquinamento da idrocarburi come fattore di rischio per gli uccelli marini*¹⁸.

7.5. Analisi spaziale del rischio ecologico nella laguna di Venezia

Lo studio, benché datato, è rilevante per i metodi e l'ambito geografico italiano¹⁹. Geostatistica e modelli *SAR* analizzano il bioaccumulo di inquinanti (metalli pesanti e PCB) in una vongola (*tapes philipinarum*) che si nutre dal sedimento che, nella laguna di Venezia, è stato a lungo soggetto a scarichi industriali di inquinanti. Dato l'impiego alimentare e il valore commerciale della specie, l'inquinante nel sedimento può danneggiare la salute umana e l'economia locale.

Ipotesi

- a) La non-coincidenza spaziale tra i siti campionati di organismo e sedimento richiede un'interpolazione spaziale dei dati di sedimento.
- b) L'inquinante nei tessuti della vongola è funzione di: inquinante nel sedimento; caratteristiche dell'organismo (peso e frazione lipidica); caratteristiche locali del sedimento (frazione fine, pH e carbonio organico totale), che favoriscono l'assorbimento dell'inquinante.

Dati

- 1-2) Dati da analisi di laboratorio su campioni di vongole e di sedimento prelevati in laguna (Magistrato alle Acque²⁰).

¹⁸ Fox *et Alii* (2016).

¹⁹ Bertazzon *et Alii* (2000); Bertazzon *et Alii* (2006); Bertazzon e Micheletti (2006).

²⁰ Per una discussione di questi dati si veda Bertazzon *et Alii* (2006).

- 3) Lo studio utilizza dati da analisi preesistenti, per cui i campioni di vongole e di sedimento provengono da siti diversi.

Metodi

- a) Interpolazione spaziale, effettuata con metodi geostatistici (*kriging*). Il metodo poggia sulla stima del semivariogramma [Fig. 7.5(b)] che è un modello di autocorrelazione spaziale.
- b) Analisi spaziale *SAR*.

Risultati

- a) L'interpolazione spaziale *kriging* produce una mappa dell'inquinamento nel sedimento (per ciascun inquinante considerato), da cui sono desunti i valori nei siti di campionamento delle vongole. La Fig. 7.5 mostra, come esempio, l'interpolazione per il mercurio.

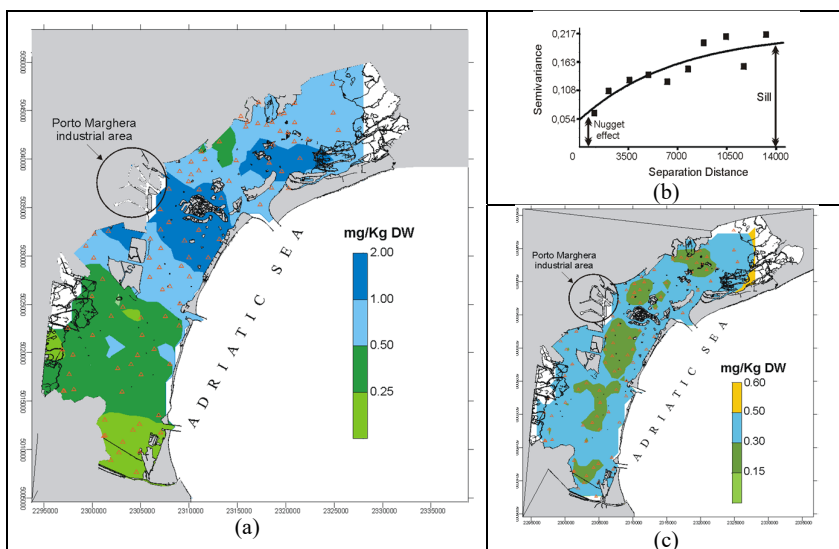


Fig. 7.5 - Interpolazione *kriging* del mercurio (Hg) nel sedimento superficiale della laguna di Venezia (a); semivariogramma omnidirezionale per il mercurio (b); incertezza (deviazione standard) dell'interpolazione *kriging* (c).

Riprodotta da Proceedings of GIS and Environmental Modeling 2000. CD-ROM, No. 122.

- b) L'interpolazione spaziale *kriging* produce una mappa dell'inquinamento nel sedimento (per ciascun inquinante considerato), da cui

sono desunti i valori nei siti di campionamento delle vongole. La Fig. 7.5 mostra, come esempio, l'interpolazione per il mercurio.

c) I modelli di regressione spaziale per ciascun inquinante [Fig. 7.6(a)] forniscono i coefficienti per la previsione spaziale del bioaccumulo [Fig. 7.6(b)]. La significatività delle variabili relative alle caratteristiche ambientali varia secondo gli inquinanti.

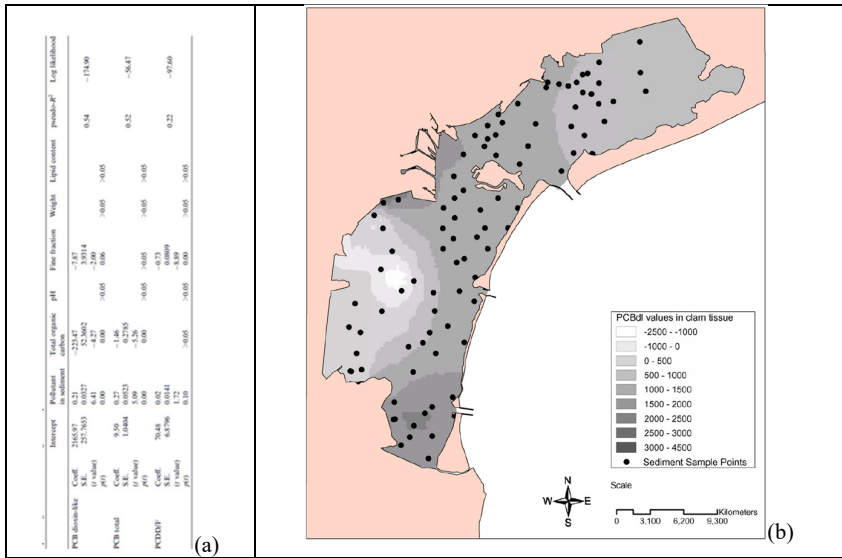


Fig. 7.6 - Regressioni spaziali per gli inquinanti organici (a) e previsione dell'inquinamento da PCBll nei tessuti delle vongole secondo la regressione spaziale (b). Riprodotta da Elsevier: Computers, Environment and Urban Systems, 30(6): 880-904.

Osservazioni e ruolo dell'analisi spaziale

- i. La considerazione esplicita delle caratteristiche ambientali locali del sedimento lagunare evidenzia l'importanza dell'impiego di metodi analitici spaziali.
- ii. **La complessità dell'ambiente lagunare sconsiglia l'impiego di interpolazione, tuttavia il metodo statistico utilizzato produce non solo stime affidabili, ma anche una stima dell'incertezza ad esse associata, come illustrato in Fig. 7.5.**
- iii. Sia la definizione del semivariogramma che quella della matrice dei pesi spaziali richiedono uno studio dell'autocorrelazione spaziale. **I parametri del semivariogramma informano la matrice**

dei pesi spaziali, producendo uno studio articolato dell'autocorrelazione spaziale per ciascun inquinante.

- iv. Il confronto dei parametri dell'autocorrelazione spaziale per ciascun inquinante illustra le caratteristiche spaziali dell'inquinamento nell'ambiente lagunare: *l'auto-similitudine esibisce un raggio di dimensione molto variabile secondo gli inquinanti.*
- v. Le previsioni dei *modelli di regressione spaziale per i vari inquinanti forniscono informazioni utili a delineare le zone di pesca in relazione all'inquinamento nel sedimento.*

7.6. Analisi spaziotemporale della mortalità nella prima ondata di COVID-19

Lo studio, scelto per l'area di studio italiana e l'impiego di soli *open data*, illustra un'analisi preliminare, presentata all'ECQTG 2021²¹ e in fase di elaborazione mentre si scrive. Tra i numerosi studi sulla pandemia²², questo contribuisce all'analisi della variazione spaziale a diverse scale, analizzando la mortalità nella prima ondata epidemica in Lombardia e Veneto, le prime regioni colpite in Occidente. L'analisi della diffusione in questi territori, allora relativamente impreparati, può chiarire il ruolo delle caratteristiche locali nella diffusione del virus, e supportare gestione e prevenzione di eventi epidemici futuri.

Ipotesi

- a) La diffusione della mortalità durante la prima ondata di COVID-19 ha seguito modalità diverse in Lombardia e Veneto.
- b) L'analisi spazio-temporale a grande scala rivela caratteristiche locali dei territori e della diffusione.
- c) L'analisi spazio-temporale globale sintetizza le caratteristiche regionali delle fasi dell'ondata epidemica.

Dati

- 1) ISTAT *open data*: decessi giornalieri per comune dal 2011²³.

²¹ Bertazzon *et Alii* (2021).

²² Gibertoni *et Alii* (2021); Rosenkrantz *et Alii* (2021).

²³ www.istat.it/it/archivio/240401

2) Definizione della variabile d'interesse:

$$EMS_{i(2020)} = \frac{\left(MS_{i2020} - \frac{\sum_{j=2015}^{2019} MS_{ij}}{5} \right)}{Pop_{Gen2020}} \times 100.000$$

Equazione 7.1

dove $EMS_{i(2020)}$ = eccesso di mortalità nella settimana i del 2020;
 MS_{ij} = somma dei decessi giornalieri per la settimana i dell'anno j ;
 $Pop_{Gen2020}$ = popolazione a gennaio 2020.
 EMS è definita per comune nelle prime 14 settimane di pandemia.

Metodi

- a) Visualizzazione cartografica dei dati di eccesso di mortalità (Fig. 7.7). La cartografia è realizzata in ArcGIS Pro.
- b) Analisi della correlazione spazio-temporale tramite indice di autocorrelazione spaziale bi-variato globale e locale²⁴. L'indice confronta la variabile "eccesso di mortalità nella settimana i " con la variabile "eccesso di mortalità nella settimana $i+I$ ". L'indice globale è calcolato in *R* (*spdep*) e quello locale in *Geoda*²⁵.

Risultati

- a) La distribuzione dell'eccesso di mortalità in entrambe le regioni presenta tratti locali difformi, caratterizzata anche dall'alternanza di valori positivi ed elevati e di valori negativi o nulli in comuni contigui o prossimi. (Fig. 7.7).
- b) L'autocorrelazione spazio-temporale alterna valori positivi e negativi [Fig. 7.8(a) - (c)]. L'indice globale di autocorrelazione spaziale bi-variato, calcolato per ciascuna regione, evidenzia un andamento analogo per le due regioni, ma caratterizzato da differenze temporali e di intensità. [Fig. 7.8(d)].

²⁴ Analogamente all'indice di autocorrelazione spaziale, discusso nel capitolo 2, l'indice bi-variato confronta le realizzazioni di una variabile con quelle di un'altra variabile ad una data distanza: https://geodacenter.github.io/workbook/5b_global_adv/lab5b.html. Visitato il 3/02/2022.

²⁵ Una volta calcolai gli indici in *Geoda*, le mappe sono state realizzate in ArcGIS Pro.

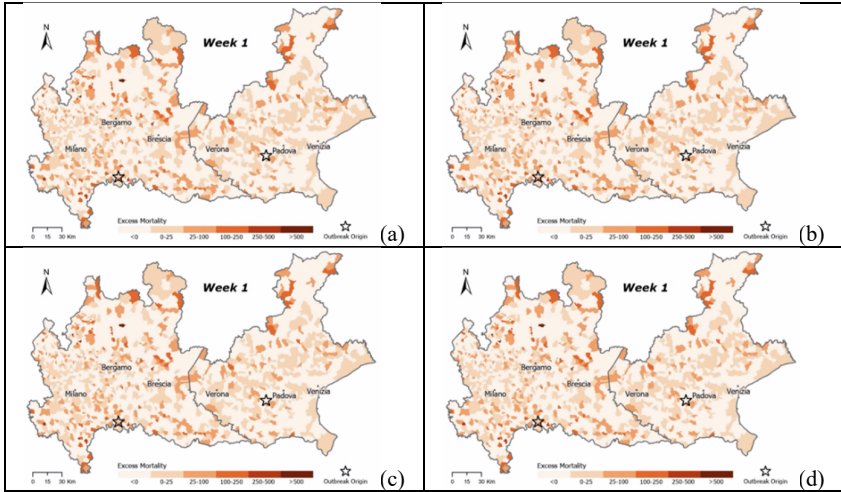


Fig. 7.7 - Distribuzione dell'eccesso di mortalità in Lombardia e Veneto durante la prima ondata pandemica nelle settimane II (a), V (b), VII (c) e X (d).

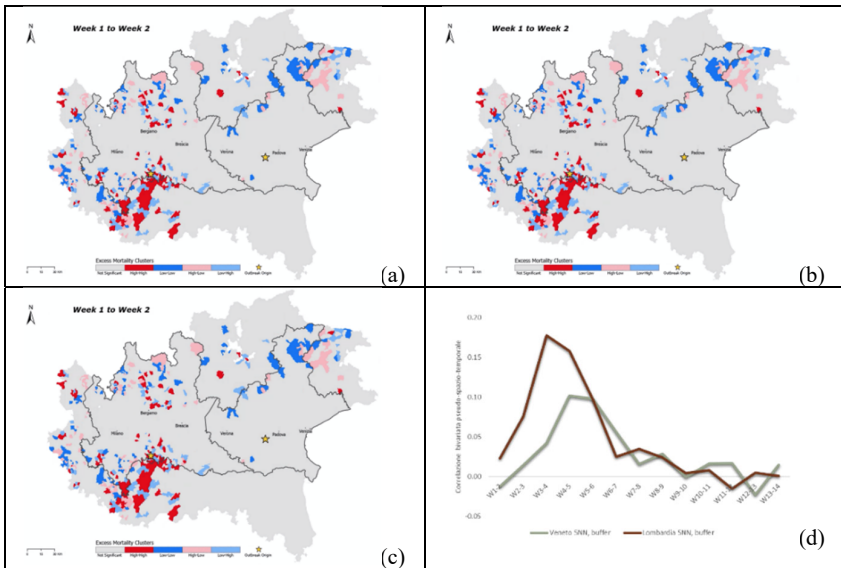


Fig. 7.8 - Correlazione spazio-temporale: indice bivariato di Moran, su vicinato di 5 vicini ($k=5$) e buffer corrispondente. Indice locale (a) - (c) e globale (d).

Osservazioni e ruolo dell'analisi spaziale

- i. Nelle fasi iniziali e finali dell'ondata epidemica l'eccesso di mortalità è prevalentemente minore o uguale a zero in entrambe le regioni. Gibertoni *et Alii* (2021) notano che il valore atteso dell'eccesso di mortalità nel 2020 sarebbe infatti inferiore alla media 2015-2019, in virtù del *trend* decrescente durante il periodo.
- ii. ***L'analisi spazio-temporale a grande scala evidenzia l'eterogeneità spaziale*** dell'eccesso di mortalità.
- iii. ***L'autocorrelazione spazio-temporale globale per le due regioni ne delinea la diffusione secondo 4 fasi***: 1) correlazione spazio-temporale inizialmente bassa e in crescita rapida; 2) valori elevati e stabili per un periodo breve; 3) discesa rapida; 4) discesa lenta e incostante fino al ritorno allo stato iniziale. ***Le differenze tra regioni sono probabilmente legate alle caratteristiche delle popolazioni e dei territori, ma anche all'implementazione di interventi volti ad arginare il contagio.***
- iv. L'indice di autocorrelazione spazio-temporale produce risultati interpretabili con matrici dei pesi spaziali estese (5 vicini più prossimi). Come illustrato in Fig. 7.8, l'area di studio è stata estesa con l'aggiunta di un corrispondente *buffer* di 5 vicini, ad esclusione delle parti confinanti con altri Stati e della costa.
- v. Si nota che ***i comuni della Lombardia sono 1.506, contro i 563 del Veneto.*** Oltre le considerazioni sulla dimensione demografica, ***questa disparità numerica può avere influenzato, oltre che i valori osservati, anche l'analisi, in relazione al problema dell'aggregazione delle unità spaziali, o MAUP.***

7.7. Morale delle storie

A conclusione di ciascuno studio si sono espresse alcune considerazioni sul ruolo dell'analisi spaziale nella comprensione del fenomeno e sull'impiego dello studio nella gestione dei territori esaminati.

In tutti gli studi l'analisi spaziale ha incrementato l'affidabilità delle stime dei modelli. L'importanza di tale affidabilità si comprende osservando che tali stime sono state impiegate per lo sviluppo di politiche socio-sanitarie o ambientali, direttamente, o tramite ulteriori studi che le hanno impiegate in modelli successivi. È utile ricordare che,

senza l'impiego dell'analisi spaziale, atta a correggere gli effetti spaziali, le stime sarebbero affette da una maggiore variabilità, che ne renderebbe le stime meno affidabili. Così, l'affidabilità delle stime si rifletterà sull'efficacia delle politiche territoriali che su esse poggiano.

In diversi casi, inoltre, l'analisi spaziale aiuta ad evidenziare aspetti meno noti, come nel caso degli studi sulla mortalità da COVID-19 e sull'inquinamento atmosferico, oppure a supportarli con un'argomentazione rigorosa, come nel caso dello studio sull'inquinamento marino. Questi ultimi due studi sollevarono anche l'interesse della stampa locale, che stimolò ulteriormente il dibattito. In particolare, la pubblicazione – non solo in ambito accademico – dello studio sull'inquinamento atmosferico, evidenziò quanto scarsa fosse la consapevolezza della distribuzione spaziale di tale inquinamento da parte del pubblico. La città di Calgary gode di una rete di percorsi pedonali e ciclabili tra le più lunghe del Nordamerica, di cui alcuni segmenti sono molto popolari soprattutto nella stagione estiva. Tuttavia, molti segmenti di quei percorsi, tra cui proprio quelli più frequentati, sono contigui ad alcune tra le principali arterie stradali, il cui traffico veicolare abbassa la qualità dell'aria sui percorsi ciclo-pedonali. Il nostro studio contribuì a promuovere questa *conoscenza spaziale* tra gli utenti di quelle piste, che così poterono operare una scelta più consapevole delle zone e dei percorsi dove svolgere attività motoria all'aperto, anche in funzione dell'inquinamento atmosferico locale. Lo studio fu molto apprezzato anche dal personale medico e soprattutto infermieristico della clinica per l'asma del locale ospedale pediatrico, che volle acquisire le nostre mappe per utilizzarle come strumento per coadiuvare le famiglie dei piccoli pazienti nella scelta di parchi e percorsi che consentissero loro di trascorrere tempo all'aperto in luoghi più sicuri in termini di qualità dell'aria. Forse non è superfluo ribadire che non sarebbe stato possibile ottenere questi risultati senza l'impiego di strumenti analitici spaziali statistici rigorosi, che furono poi verificati in ambito interdisciplinare, con l'apporto di conoscenze specifiche in materia di ambiente e sanità pubblica.

Quest'ultima considerazione ci porta a concludere ricordando l'importanza degli aspetti meno rigorosi dell'analisi, dall'impiego di *expert opinion* nello studio dell'inquinamento marino al successo di quello sull'inquinamento atmosferico grazie all'uso della VGI. Questo sottolinea la necessità di un approccio multidisciplinare e aperto che

completi il rigore dell'analisi spaziale. Aprendo così la ricerca a prospettive diverse, questo approccio ne espone la metodologia e i risultati a critiche e contributi che non possono che migliorarne la qualità. È ancora in questa prospettiva che ribadiamo la necessità di condurre l'analisi in maniera replicabile e trasparente, ciò che si è fatto in tutti gli studi presentati per quanto riguarda il *software*, e ove possibile per quanto riguarda i dati.

Dopo questa rassegna di studi applicativi, che si è chiusa volutamente in ambito geografico italiano, concludiamo questa parte con una panoramica, purtroppo breve, della vicenda dell'analisi spaziale in Italia.

8. *L'analisi spaziale in Italia*

- ...Di là l'uomo si parte e cavalca tre giornate tra greco e levante...
- riprendeva a dire Marco, e a enumerare nomi e costumi e commerci d'un gran numero di terre. Il suo repertorio poteva dirsi inesauribile, ma ora toccò a lui d'arrendersi. Era l'alba quando disse:
 - Sire, ormai ti ho parlato di tutte le città che conosco.
 - Ne resta una di cui non parli mai.
- Marco Polo chinò il capo.

Calvino, *Le città invisibili*, 1978, p. 94

8.1. Premessa

L'analisi spaziale nasce e cresce prevalentemente in ambiente anglosassone e in lingua inglese. Molti contributi importanti provengono da autori europei per origine o affiliazione e, col passare del tempo, sempre più frequenti sono quelli esterni a questi ambiti, primo fra tutti quello asiatico. In questo volume non ci siamo occupati esplicitamente di letterature diverse da quella in lingua inglese, consapevoli, da un lato, che questa nostra scelta comprende molta della letteratura internazionale, ma dall'altro anche che essa può escludere involontariamente contributi importanti, scritti in altre lingue o pubblicati in contesti nazionali o locali. Premesso che questo volume non ha alcuna pretesa di esaustività, desideriamo considerare un'unica letteratura *non* in lingua inglese: quella italiana. Ce ne occupiamo in quest'ultimo capitolo, necessariamente più breve dei precedenti. Le ragioni di questa scelta sono sin troppo evidenti. Pur sapendo in partenza che la letteratura sull'analisi spaziale nella nostra lingua non è vasta, è tuttavia importante posizionare questo lavoro nel contesto nazionale, collocandolo in un dibattito di cui auspichiamo sviluppi futuri, forse anche stimolati da questo nostro contributo.

8.2. L'analisi spaziale tra il GIS e la geografia italiana

Lungo tutto questo volume non ci siamo ancora soffermati sulla storia e la vicenda dell'analisi spaziale in Italia, soprattutto perché non siamo al corrente di un dibattito, una letteratura, una comunità, specificamente centrati sull'analisi spaziale nel nostro Paese; tuttavia, non riteniamo che questo basti a chiudere questo nostro discorso senza un accenno alla vicenda dell'analisi spaziale in Italia.

Come altrove, la vicenda dell'analisi spaziale potrebbe essere intrecciata a quella del GIS, la cui vicenda, nel nostro Paese, è diversa da quella dei paesi anglofoni, a partire dal nome che lo designa. Fin dai suoi primordi, il mondo anglosassone affermava la centralità della geografia con la *G* di GIS, e altrettanto faceva quello francofono con il termine *Geomatique*. In Italia invece non furono tradotti i termini che formano l'acronimo GIS, ciò che avrebbe portato a rovesciarne l'ordine in SIG¹, ma si scelse di mantenere l'acronimo anglofono GIS. Frattanto, si assistette anche ad un tentativo di tradurlo in italiano, tuttavia sottraendovi la *G* di geografia, coniando invece l'acronimo SIT, Sistemi Informativi Territoriali, da cui l'elemento *Geografico* scompariva del tutto. Perché l'aggettivo Geografico venne sostituito con Territoriale, togliendo la Geografia dal GIS? Le radici di questa vicenda vanno forse ricercate nell'evoluzione del GIS nel nostro Paese, legata più alla pratica svolta da urbanisti, ingegneri e altri professionisti che da geografi: il forte interesse, manifestato fin dagli inizi, da quelle professioni, e il minore, o forse solo più tardivo interesse dei geografi sono probabilmente fra le ragioni della vicenda del GIS in Italia. Questa vicenda riconduce l'appetibilità del GIS, forse manifestatasi più nel contesto professionale che in quello accademico, a quella caratteristica di *big business* che anche in Italia, forse, ne condizionò le sorti dentro e fuori l'accademia.

Diversamente dal mondo anglofono e da altri Paesi europei, il GIS tardò ad entrare nelle università italiane, e ove lo fece fu come mero strumento di cartografia, o come insegnamento prettamente tecnico,

¹ Con la diffusione del GIS oltre il mondo anglofono, anche il suo nome si trasforma adattandosi ai diversi contesti linguistici. Così, volendo considerare solo i principali ambiti europei, l'acronimo rimane inalterato nel mondo germanico, mentre in quello francofono e ispanico si capovolge in SIG, secondo la sintassi neolatina. Oltre le differenze di grafia, esistono quelle di pronuncia: "Gi-Ai-Es" in inglese, "Ghis" in tedesco, "Gis" in italiano.

anzi, non di rado addirittura insegnato da tecnici. Per ragioni che trascendono lo scopo di questo volume, il GIS non si conquistò mai un posto centrale nella geografia italiana, né nei corsi universitari, né nella ricerca accademica². Relegato nei corsi professionalizzanti e nella ricerca applicata ne rimase caratterizzato dalla debolezza dell'apparato accademico e di pensiero che avrebbe dovuto formare i professionisti del GIS. Questa situazione pesò sulle sorti del GIS, condannandolo ad essere soprattutto strumento. E come strumento il GIS fu impugnato da altri: ingegneri, urbanisti, ambientalisti e una gamma di non-geografi che, nella professione come nella letteratura, ne affermano la titolarità.

La geografia italiana, così, non visse l'ondata di piena che aveva travolto quella nordamericana³ e solo tardivamente e timidamente cominciò ad occuparsi di GIS, inteso non solo come strumento⁴. Le riviste di geografia ospitarono contributi, sia applicativi che di riflessione, e furono organizzati convegni, tra i quali spicca ASITA⁵, forse il più longevo, che raggruppa un insieme variegato di accademici e professionisti, aziende private ed enti pubblici, interessati, più che alla ricerca, alle applicazioni. In questa comunità i geografi sono solo una delle tante voci, né occupano una posizione privilegiata accanto a quelle delle discipline ingegneristiche, le scienze della terra e quelle urbanistiche.

² Occorre considerare però che ci fu un preciso e pesante anatema nei riguardi della validità delle “nuove geografie”, venute da olttralpe e oltre oceano, ed in particolare della “geografia quantitativa”: per questo ci si può riferire al testo di Ferro (1983) in cui tra l'altro affermava: «oggi che la “moda” di una tale geografia [la geografia quantitativa] sta passando [...] ma gli anni perduti sono ugualmente parecchi, le energie sciupate sono molte [...] pensando che l'utilità della maggior parte di tali lavori era soprattutto (e soltanto) quella di tenere in esercizio le facoltà intellettuali».

³ Di tutte le questioni che hanno lacerato l'ambiente del GIS in Nord America per oltre 30 anni, in Italia non è giunta che qualche debole eco. Alla creazione del GIS la geografia italiana non ha partecipato, e quando è arrivato era già pronto per l'uso. Tutto questo non ha però fermato la marcia del GIS, che anche in Italia oggi è “big business”, ma molto meno per la geografia.

⁴ Si vedano al riguardo le raccolte curate da Guarrasi (1996) e Muscarà (1998).

⁵ Federazione delle Associazioni Scientifiche per l'Informazione Ambientale e Territoriale.

8.3. La letteratura in lingua italiana

Se la vicenda dell'analisi spaziale è intrecciata con quella del GIS, non c'è certo da aspettarsi una sua marcata presenza nella geografia italiana. Abbiamo provato a saggiare il terreno effettuando una banale ricerca⁶ utilizzando *Google Scholar* per esplorare non solo la letteratura accademica tradizionale, ma anche la cosiddetta letteratura grigia, alla ricerca di contributi contenenti le parole chiave “analisi spaziale” (*sic*). La ricerca, con tutti i limiti del metodo⁷, ha prodotto risultati scarni e datati: un articolo sulla centralità urbana (Porta, Crucitti, Latorra, 2006), una monografia sulla scienza dell'informazione geografica (Boffi, 2004) e la traduzione dall'inglese di un lavoro ancor più datato (Unwin, 1986). A questo si aggiunge una serie di contributi che contengono i termini chiave ma applicativi che spaziano dall'archeologia alla politica e all'economia⁸. Abbiamo poi effettuato alcune delle ricerche correlate suggerite dallo stesso *Google Scholar* modificando i termini chiave, ad esempio con “Analisi spaziale gis” (*sic*). I risultati non sono molto diversi: lavori applicativi, dall'archeologia all'architettura alla pianificazione territoriale, in pubblicazioni eterogenee⁹ e un manualetto di GIS¹⁰, anch'esso datato.

Si è optato quindi per una ricerca sistematica negli indici delle riviste e delle case editrici che si occupano di geografia, direttamente oppure tramite Google Scholar¹¹. Non si sono effettuate ricerche

⁶ Ricerca effettuata il 05/07/2021.

⁷ Questo tipo di ricerca poggia sul presupposto dell'esistenza di indici digitali completi e aggiornati, e non abbiamo verificato questo presupposto. Sui risultati potrebbero quindi pesare carenze nell'indicizzazione, soprattutto delle riviste.

⁸ Riportiamo solo i primi tre risultati: Bagolini e Dalmeri (1987); Zucchini (2013); Russo (2008).

⁹ I primi risultati sono: Sanna (2012); Chavarría Arnau e Valente (2009); Monti (2007).

¹⁰ Favretto (2000).

¹¹ Ricerca effettuata il 05/07/2021. *Rivista Geografica Italiana*: contiene diversi contributi che si snodano lungo un periodo abbastanza lungo, tuttavia gli autori più presenti sono affiliati ad istituzioni straniere. *Bollettino della Società Geografica Italiana*: la ricerca produce un solo articolo sull'ontologia della geografia e rappresentazioni cartografiche, scritto da un autore affiliato all'università di Bucarest (Tambassi, 2018). *Geotema*: anche qui si trova un lavoro importante (Goodchild, 1996) che probabilmente è una traduzione, anche se l'articolo non riporta né l'autore della traduzione né il titolo dell'articolo in lingua originale. In ogni caso, al pari di (Unwin, 1986), si tratta di una presenza importante nella letteratura italiana. *Ambiente Società Territorio*. Troviamo la nota di lavoro di Coccia (2000), che utilizza un modello di interazione di tipo gravitazionale, che appare nel filone della geografia quantitativa

specifiche in atti di convegni, sia per i limiti della loro indicizzazione, sia per la probabile sovrapposizione con articoli già emersi dall'esame dei periodici. Si sono infine esaminati i titoli delle principali Case Editrici¹².

Nemmeno nei risultati dell'importante Convegno di Varese (Corna Pellegrini e Brusa, 1980) né nei due volumi curati da Giacomo Corna Pellegrini (1987), che rappresentano le più importanti e basilari rassegne relative agli studi e tendenze in atto nella geografia italiana dagli anni '60 agli anni '80, vi è un accenno a GIS, né, tanto meno, ad "analisi spaziale", ma solo a brevi riferimenti a matematica e statistica, il più pertinente dei quali è il richiamo all'analisi fattoriale.

Tra i gruppi di lavoro dell'AGEI non compare l'analisi spaziale (e neppure il GIS); vi compare tuttavia un gruppo denominato "Nuove tecnologie per la conoscenza e la gestione del territorio", tra i cui obiettivi e riferimenti bibliografici si incontrano parole chiave e autori sovrapposti all'analisi spaziale (e.g., spatial turn, Goodchild, Sui)¹³, benché l'interesse del gruppo sia chiaramente centrato su *GIScience* e non su analisi spaziale.

tradizionale. *J-Reading Journal of Research and Didactics in Geography*: produce una serie di risultati pertinenti: Pesaresi (2020), che si configura come un lavoro di analisi spaziale; Santini, Miracco, Santucci (2018), che implementa una *multi-criteria analysis* in ambiente GIS; Pesaresi e Marta (2014), anche questa un'analisi spaziale; Primi e Marchioro (2020), che conducono un'analisi di diffusione rifacendosi ad Hägerstrand. *Bollettino dell'Associazione Italiana di Cartografia*: troviamo un solo lavoro che propone una riflessione su GIS e nuova geografia, comprensibilmente orientata verso la cartografia e più al GIS che all'analisi spaziale (Borruso, 2013). *Semestrale di studi e ricerche di geografia*: spicca il volume N° 1 (2017), *Geografia e digital technologies*, ove tuttavia nessun contributo si occupa di analisi spaziale. *Prolegòmena gheographikà*: anche qui il volume N° 2 (2015), *Crossing "spatial turn"*, non si concentra sugli aspetti dello spatial turn relativi all'analisi spaziale. Anche nel N° 2 (2008) "Flussi, Reti e Sistemi Informativi Geografici", al di là del già citato articolo di Borruso, non spiccano contributi di analisi spaziale.

¹² *FrancoAngeli*: Unwin (1986). *Utet Università*, nella sezione geografia produce risultati relativi non all'analisi spaziale, ma al GIS, e quindi relativamente rilevanti: Pesaresi (2017) e Migliacco e Carrion (2016). *Pàtron*: ricerca non disponibile; tuttavia, un esame manuale dei titoli della materia 'Geografia, scienze ambientali e territoriali' produce, oltre a una manciata di titoli relativi a cartografia e GIS, anche un manuale sull'applicazione di metodi statistici (perlopiù descrittivi) a dati spaziali (Conti Puorger, 2018). La ricerca del termine "analisi spaziale" nel sito di *Mimesis* produce "Geolocalizzare per decidere" di D. Dapiaggi, un'applicazione del GIS in ambito bibliotecario. Ulteriori ricerche con il termine 'geografia' e l'esame dei titoli di 'cartografie sociali' non producono risultati pertinenti all'analisi spaziale.

¹³ www.ageiweb.it/gruppi-di-lavoro/nuove-tecnologie-per-la-conoscenza-e-la-gestione-del-territorio/. Visitato il 19/01/2021.

Questa breve ricerca bibliografica, sicuramente incompleta e preliminare¹⁴, evidenzia alcune caratteristiche dell'analisi spaziale nella letteratura geografica in Italia. Innanzitutto, si tratta di una presenza sporadica, tantopiù perché questa sparuta letteratura comprende sia lavori in traduzione che contributi di autori affiliati ad istituzioni straniere. Si osserva anche che i lavori in traduzione provengono dalla sola letteratura anglofona. Tra le riviste spicca, per numero di presenze, *J-Reading* – una rivista di didattica, che totalizza un numero di contributi analogo alla *Rivista Geografica*, ma molto più recenti. Tra le case editrici spicca UTET, e anche in questo caso si tratta di una collana dedicata alla didattica. Notiamo infine che la nostra ricerca è stata condotta volutamente nell'ambito della geografia, soprattutto per quanto riguarda le riviste, mentre il panorama italiano potrebbe comprendere presenze in settori disciplinari diversi.

Difficile trarre una conclusione sullo stato dell'analisi spaziale nella geografia italiana. La stessa sporadicità dei risultati bibliografici potrebbe suggerire che le parole chiave utilizzate nella ricerca non fossero le più appropriate. A sua volta questo potrebbe rimandare all'esistenza di altri termini chiave (di cui non siamo al corrente) o, più probabilmente, ad una presenza ancor meno strutturata, ossia l'assenza di termini chiave coerenti e, in definitiva, l'assenza di una comunità che si riconosca nell'ambito di ricerca che abbiamo chiamato analisi spaziale. Eventuali carenze nell'indicizzazione delle riviste, oltre a condizionare i risultati di questa ricerca, contribuirebbero ad oscurare la presenza di un eventuale corpus di ricerca, rendendolo di fatto invisibile alla comunità scientifica nazionale ed internazionale.

La presenza relativamente maggiore di contributi relativi all'analisi spaziale e/o al GIS in ambito didattico fa supporre una presenza della materia nella letteratura italiana più come bagaglio di strumentazione che come produzione scientifica. È evidente, inoltre, come questa letteratura sia strettamente legata al filone GIS, ciò che confermerebbe l'intreccio tra le due discipline, ma anche la percezione di entrambe come strumenti, ancillari ad altri studi, dentro e fuori la geografia.

¹⁴ Una ricerca ancor meno strutturata, basata su conoscenze personali e ricerca manuale tra gli indici, rivela qualche ulteriore presenza riconducibile prevalentemente all'ambito GIS / GIScience. Preferiamo comunque non riportare questi risultati, data la loro natura estremamente soggettiva.

Si osservi poi che la ricerca bibliografica discussa sopra è stata condotta inizialmente solo con i termini chiave “analisi spaziale”, solo successivamente abbinati ai titoli di riviste e case editrici. La ricerca iniziale, dunque, avrebbe dovuto evidenziare eventuali presenze anche in contesti disciplinari diversi dalla geografia. Ciò per certi versi è accaduto, restituendo comunque una serie disparata di contributi, e questo confermerebbe l’ipotesi che l’analisi spaziale sia utilizzata “all’occorrenza” in contesti diversi, senza che ne emerga un filone coerente in alcun contesto. A queste considerazioni vorremmo aggiungere il posizionamento disciplinare della geografia nell’accademia italiana, che comporta una più cospicua presenza di ricercatori di formazione umanistica; non è quindi scontato che studiosi di questa formazione dimostrino entusiasmo nei confronti delle metodologie matematiche e statistiche proprie dell’analisi spaziale. Analoghe considerazioni valgono per gli studenti iscritti ai corsi di laurea in geografia, ai vari livelli.

Storicamente, furono invece i geografi legati al raggruppamento economico ad abbracciare il filone quantitativo¹⁵. I risultati bibliografici mostrano infatti come la penetrazione dell’analisi spaziale in Italia sia avvenuta molto tempo addietro (Unwin, 1986), facendo supporre addirittura che questa entrata sia legata al dibattito inerente alla ‘vecchia’ geografia teoretico-quantitativa: una letteratura relativamente cospicua attorno al volgere del secolo che, come discusso nel capitolo 3, fu assai vivace in quel periodo.

Infine, osserviamo la presenza, nella letteratura anglofona internazionale, di contributi che potremmo dire italiani per l’argomento, il cognome degli autori, e in certi casi la loro affiliazione¹⁶. Anche in questo contesto va ribadito che la nostra analisi si è limitata alla letteratura anglofona, e quindi potrebbero esserci sfuggiti contributi significativi scritti in lingue altre, rispetto all’italiano e all’inglese. Si tratta di un insieme di contributi potenzialmente vasto, ma anche sfuggente: in ogni caso pare difficile riconoscerne una comunità o un’identità disciplinare. Rimane comunque importante riaffermare che il lavoro di autori italiani si estende oltre le riviste italiane e i lavori in lingua italiana.

¹⁵ Fra gli altri: Vagaggini e Dematteis (1976); Lando (1985); Zanetto (1979, 1986).

¹⁶ Si tratta chiaramente di un’argomentazione aneddotica e soggettiva. Citiamo, solo a titolo di esempio: Giordano e Cole (2018); Gilbertoni *et Alii* (2021); Bressan (2021); Bertazon, Couloigner, Underwood (2019).

8.4. L'analisi spaziale nella geografia italiana

Si può forse tentare una conclusione affermando che l'analisi spaziale sembra non avere *ancora* trovato una propria collocazione o identità nella geografia italiana. Tuttavia, è importante osservare che ne esistono applicazioni sparse, dentro e fuori la geografia, e che esiste un insieme di ricercatori che sono italofoeni e nel contempo anglofoeni. La presenza di questo insieme di persone nella letteratura internazionale suggerisce che potenzialmente ci sia *già* un gruppo di individui dotato di conoscenze e capacità tali da poter portare l'analisi spaziale nella letteratura geografica italiana. Per questo potrebbe essere auspicabile la costituzione di una sorta di gruppo d'interesse attorno a questi temi, al fine di portare esplicitamente la riflessione sull'analisi spaziale all'interno della geografia italiana, magari finalizzata al sostegno di applicazioni più consapevoli nella pratica professionale, anche nei settori pubblico e privato. Anche se forse è prematuro pensare ad una penetrazione più strutturata dell'analisi spaziale nella geografia italiana, se non in tempi più lunghi, questa sorta di presenza virtuale potrebbe costituirne un preludio.

Benché una maggiore presenza dell'analisi spaziale nella geografia italiana sia auspicabile, non è facile presagire una penetrazione in tempi brevi. Tuttavia, come si è più volte ribadito, l'analisi spaziale, al tempo degli *spatial big data*, della *spatial data science* e delle *spatial analytics*, potrebbe costituire un importante molla per lo sviluppo della geografia, anche in Italia. La geografia italiana fu forse troppo lenta nel comprendere e accogliere il GIS: oggi l'analisi spaziale potrebbe fornirle una nuova occasione da non perdere.

Verso una conclusione: la geografia che conta

Le quai cose tutte sono d'altissima et bellissima speculatione , facendoci elle per vie ò arti Matematiche comprendere et conoscere il cielo come sia di natura sua percioche egli ne circonda girando attorno ; et dimostrarci la terra per sembianza ò imagine . Percioche la vera terra , per esser' in se stessa grandissima , et non si aggirando , come il cielo , d'intorno a noi , non è possibile, che da gli huomini sia cercata tutta insieme, nè à parte à parte.

Tolomeo, *Geografia*. Libro I, p. 2 (tradotto da G. Ruscelli, 1574)¹

Lo spazio del nostro mondo è teatro di fenomeni complessi: dagli esiti incerti di una pandemia, al clima impazzito, all'incertezza alimentare, alle crisi economiche e geopolitiche. Nel 1854 la rudimentale analisi spaziale condotta del medico inglese John Snow sulla mortalità associata ad un focolaio epidemico di colera fornì un apporto straordinario alla conoscenza della malattia, contribuendo ad arginare il contagio e a scalzare la teoria medica dominante. Oggi disponiamo di teorie, strumenti e dati per condurre analisi spaziali capaci di aggiungere tasselli importanti alla nostra comprensione dei fenomeni: questa analisi spaziale ne fornisce una valida chiave di lettura attraverso metodi quantitativi rigorosi che si spingono ben oltre paradigmi riduttivamente quantitativi, valutati in una prospettiva critica, in un dialogo aperto con le scienze sociali, mediche, ambientali.

Attraverso una serie articolata di capitoli abbiamo intrapreso un dialogo con chi conosce la disciplina e con chi verrebbe conoscerla, cercando di dire che cos'è l'analisi spaziale, ma anche che cosa non è, cosa fa bene e cosa potrebbe fare meglio. Abbiamo tracciato il ritratto di una disciplina che affonda le radici principali nella geografia teorico-quantitativa, nel GIS e nella statistica, e che, reciprocamente, stimola lo sviluppo di quelle discipline. Ne è emersa l'immagine di una disciplina che fu e rimane la ragion d'essere del GIS, che rimane parte qualificante della moderna geografia quantitativa, e che continua a

¹ Il testo digitalizzato si trova in: <http://websrv.archeo.unict.it:8080/items/show/203#?c=0&m=0&s=0&cv=88&xywh=109%2C20%2C654%2C568> (visitato il 9/02/2022).

stimolare nuovi sviluppi della statistica, capaci di affrontare gli effetti generati dai processi e dai dati spaziali.

Radicata nel dibattito disciplinare delle discipline madri, l'analisi spaziale fatica ancora a conquistare una propria autonomia. Figlia di geografia quantitativa e *GIScience*, da queste si distingue per la profondità del legame con la statistica, che ne fa una scienza rigorosa e attenta all'incertezza e all'errore. L'analisi spaziale ha maturato una grande consapevolezza dei problemi posti dai processi e dai dati spaziali, riassunta dallo slogan *spaziale è speciale*. Ha quindi approntato metodi per affrontarli e gestirli, incrementando l'affidabilità dei modelli quantitativi. Una volta individuati i problemi, la ricerca di una soluzione è questione urgente, e come Alice nel Paese della Regina Rossa, la disciplina ha cercato al più presto dei metodi per arginare i problemi, ma senza perdere di vista l'urgenza di rispondere alle questioni concrete. Così ha cercato (e trovato) soluzioni metodologiche in altre discipline, ma proprio perché nate altrove, quelle soluzioni rimangono parziali, non del tutto capaci di affrontare la complessità dei processi spaziali. All'analisi spaziale rimane molto lavoro da compiere: la ricerca di teorie sviluppate al proprio interno, da cui possa discendere una metodologia coesa, capace di affrontare compiutamente la gamma di effetti ascrivibili a proprietà quantitative diverse. Le proprietà speciali dei processi spaziali sono misurate da indici quantitativi, e su queste misure quantitative poggiano i metodi statistici. Tuttavia, come si è visto, nessuna proprietà è del tutto separabile dalle altre, e ciascuna proprietà è solo la manifestazione quantificabile di caratteristiche qualitative. Per questo è necessaria una conoscenza profonda e anche qualitativa dei processi spaziali, da cui possano discendere rappresentazioni dello spazio e dei suoi processi che siano complementari al rigore dei metodi statistici.

All'alba di questo secolo è finalmente iniziato l'importante cammino verso un dialogo proficuo con altre discipline, paradigmi e metodi. Intraprendere questo sentiero non è stato facile, per il fraporsi di tensioni e antagonismi che a lungo hanno punteggiato la vicenda storica della geografia e delle geografie. Superati gli ostacoli, l'analisi spaziale non è sola su questo sentiero, ma ha per compagna la più vasta e consolidata *GIScience*, che ha saputo anche avviare un discorso critico, che si occupi non solo di trovare metodi quantitativi e statistici efficienti, ma che abbia anche il coraggio di riflettere sui propri

presupposti, sul come e sul perché, sulle influenze del potere, della politica, dell'economia, che sappia riconoscere non solo i meriti, ma anche i limiti della ricerca. Molto del discorso critico contemporaneo s'innesta nel dibattito sui dati, in un momento storico in cui la letteratura ha coniato il termine *diluvio*, con riferimento ai dati che quotidianamente ci piovono addosso. Torneremo sulle prospettive che questo diluvio offre alla geografia e all'analisi spaziale, ma intanto qui ribadiamo l'imperativo dell'uso critico dei dati. I dati rimangono un elemento imprescindibile dell'analisi quantitativa, e proprio per questo richiedono una riflessione profonda e articolata che non trascuri l'epistemologia, l'etica, la politica dei dati. Proprio mentre scriviamo queste ultime righe si avvia alla conclusione la serie di seminari organizzata dall'AAG sulla *geoetica*: un segnale importante dell'impegno della comunità (non solo accademica) che si riconosce nella scienza dell'informazione geografica.

La rivoluzione dei cosiddetti *big data* e la tecnologia alla portata di tutti ci inducono nella tentazione di lasciare che siano i dati a parlare, che sia la tecnologia a guidare l'analisi dei fenomeni. Da questa tentazione è emersa anche una geografia quantitativa che non è analisi statistica spaziale. Proprio in risposta all'abbondanza di dati e all'esplosione della tecnologia, l'analisi spaziale riafferma la necessità di conoscere i dati, di formulare e verificare ipotesi, di poggiare su teorie solide e di rammentare sempre che i dati non sono i fenomeni.

Questa ostinazione a perseguire rigore quantitativo, statistico e teorico ha paradossalmente ostacolato lo sviluppo e la popolarità dell'analisi spaziale: non è con questa ostinazione che si sfornano *analitiche* redditizie su miriadi di dati corredate da mappe fantasmagoriche. Già molto tempo addietro Michael Goodchild identificava, nell'allora nascente scienza dell'informazione geografica, una tensione tra elitismo e populismo, tra numerico e visuale. In questa duplice tensione, sempre attuale, tra analisi spaziale e GIS, riconosciamo anche uno dei principali limiti della nostra disciplina. Torneremo anche sul nome della disciplina, ma occorre qui ricordare che essa è anche nota col nome di statistica spaziale: in essa quindi centrale è il rigore matematico e statistico, solo corredato da visualizzazioni e mappe. Questa posizione la contrappone anche filosoficamente al GIS, il cui fulcro è la rappresentazione visuale, capace, anche da sola, di produrre i propri risultati. Negli anni questa differenza e questa

tensione hanno segnato la distanza tra le due discipline, che solo in epoche recenti hanno tentato un avvicinamento e forse un intreccio.

Per ragioni storiche e metodologiche, la cassetta degli attrezzi dell'analisi spaziale contiene stringhe di codice analitico e statistico che bisogna saper leggere, comporre e far girare, innestate in concetti matematici e statistici da comprendere profondamente e da applicare con perizia e flessibilità. Anche questa cassetta, in contrapposizione a quella del GIS, è percepita dai più come pesante, poco attraente, e soprattutto difficile da usare. Su queste basi è cresciuta intorno all'analisi spaziale la fama di disciplina ostica, per soli iniziati. Questa fama non ha giovato alla sua immagine, relegandola a dominio di pochi: eletti, iniziati, magari anche bravi, ma pur sempre pochi. E così rimane disciplina di nicchia, nota a pochi, praticata da pochi, ambita da pochi. Questa debolezza può essere forse superata per vie diverse. La prima è la divulgazione della materia oltre il circolo dell'élite, illustrandone le potenzialità e mostrando che proprio quel suo rigore può incrementare la conoscenza dei processi spaziali, e a sua volta contribuire a definire politiche e coadiuvare la gestione del territorio. La seconda è proprio l'avvicinamento al GIS, che con la sua visione populista può aumentare l'attrattiva dell'analisi, favorendo la comprensione del suo potenziale attraverso un'interfaccia più visuale, intuitiva e rassicurante. Non si sta certo suggerendo che interfacce intuitive e semplici *click* possano sostituirsi a stringhe di comando e analisi rigorose, ma solo che anche queste interfacce possono contribuire a traghettare più geografi verso l'analisi spaziale.

La cassetta degli attrezzi è stata una metafora importante nella transizione del GIS da sistema a scienza dell'informazione geografica. In questa metafora si è definito *chi* fa analisi spaziale, qualcuno che è speciale perché oltre a utilizzare i propri attrezzi li deve anche saper fabbricare, perché deve saper costruire e maneggiare quelle stringhe di codice che porta nella cassetta. Richiamando il vecchio dibattito sul GIS, questa capacità non avvicina necessariamente la figura dell'analista spaziale a quella dello scienziato, ma certo la allontana da quella del semplice utente. Questa figura professionale, intimamente coinvolta sia nella produzione che nell'impiego degli attrezzi, per questa stessa ragione non è avulsa dallo sviluppo stesso della disciplina. Le stringhe di codice dell'analisi spaziale girano in un ambiente *software* che non è un pacchetto commerciale, ma è prodotto dalla comunità

stessa degli analisti, gratuitamente e in modalità *open source*, cosicché chiunque può accedere al codice, modificarlo, migliorarlo, e ripetere indipendentemente l'analisi. È proprio l'approccio critico alla *spatial science* che la spinge ad impiegare metodi trasparenti e replicabili, capaci di soluzioni efficaci e credibili, e anche questo può contribuire a sfatare il mito di una disciplina arroccata dietro codici arcani per soli eletti.

Si è detto di un'identità che la disciplina fatica a conquistare. L'elitismo che la caratterizza, almeno nella percezione, certo non l'aiuta, perché la si conosce poco e, senza il beneficio della conoscenza, la si qualifica male, la si confonde con ciò che non è, la si definisce con una terminologia molteplice. Anche questa è una debolezza della disciplina, che deve ancora maturare una chiara percezione di sé e un'immagine da proiettare al suo esterno. Vi è poi un altro elemento che pesa sul suo nome rendendola sfuggente: il richiamo del profitto. Anche in questo non si può scordare la vicenda del GIS, il cui sviluppo fu trainato, ma anche distorto, dal *big business*, dai profitti che con esso si potevano ottenere, dentro e fuori l'accademia. L'analisi spaziale, piccola ed elitaria, non fu mai *big business*, ma è proprio l'*analisi* che oggi è *big business*, sempre più appetibile in virtù della crescente disponibilità di *big data*. Sui *big data*, si è visto, non si fa necessariamente *analisi [statistica]*, si fanno *analytics*, appetibili all'industria e foriere di guadagni, che si sfornano in rapida successione senza lasciarsi distogliere da teorie, statistiche rigorose, verifica di ipotesi. E poiché i dati che piovono sono spaziali, la disciplina che chiamiamo analisi spaziale vorrebbe anche essere *spatial analytics*, forse per rincorrere quei guadagni, o forse ancora alla ricerca di quella polarità che, rimanendo fuori da questi ambiti, le sarebbe preclusa.

Un'ampia parte di questo volume è dedicata alla pratica, illustrata anche con esempi di come l'analisi spaziale possa fornire conoscenza utile alla gestione del territorio e alla definizione di politiche ambientali e socio-sanitarie. Ben oltre gli esempi illustrati, proprio l'utilità delle applicazioni concrete dovrebbe essere il volano che consente alla disciplina di uscire dalla nicchia in cui la sua stessa vicenda l'ha relegata. Non sono i dati spaziali a stimolare l'analisi, sono i fenomeni, i problemi del nostro mondo a richiedere soluzioni che non possono prescindere da un'analisi spaziale rigorosa e affidabile: al rigore e

all'affidabilità non possono supplire né la grandezza dei dati né la ricchezza delle analitiche.

Durante la stesura di questo libro i fenomeni si sono complicati, i dati sono aumentati, la tecnologia ha continuato a correre; la domanda di analisi rigorosa e competente dei dati spaziali ha continuato a crescere. C'è bisogno di questa analisi, e di geografi disposti a raccoglierne la sfida, magari infilando anche questo volume nella propria cassetta degli attrezzi, per rispondere alla domanda di geografia che ... conta.

Bibliografia

Un elenco particolareggiato esaurisce tutto quello che c'è da dire su un oggetto [...]. La "natura" di un aggregato particolare è determinata dalle sue parti e dal modo con cui le parti sono connesse l'una all'altra. *Se si enumerano le parti nell'ordine appropriato si ha l'oggetto.*

Feyerabend, *Contro il metodo. Abbozzo di una teoria anarchica della conoscenza*, 1995, p. 219.

- Abler R.F. (1993) "Everything in its Place: GPS, GIS, and Geography in the 1990s", in *The Professional Geographer*, 45(2):131-139.
- Adrados C., Girard I., Gendner J.-P., Janeau G. (2002) "Global Positioning System (GPS) Location Accuracy Improvement Due to Selective Availability Removal", in *Comptes Rendus Biologies*, 325(2):165-170.
- Aitken S.C., Michel S.M. (1995) "Who Contrives the "Real" in GIS? Geographic Information, Planning and Critical Theory", in *Cartography and Geographic Information Systems*, 22(1):17-29.
- Akerman J.R. (2017) *Decolonizing the Map. Cartography from Colony to Nation*, Chicago, University of Chicago Press.
- Alexander D.S., Alfonso M.L., Hansen A.R. (2015) "Childhood Obesity Perceptions Among African American Caregivers in a Rural Georgia Community. A Mixed Methods Approach", in *Journal of Community Health*, 40(2):367-378.
- An L., Tsou M.-H., Crook S.E., Chun Y., Spitzberg B., Gawron J.M., Gupta D.K. (2015) "Space-Time Analysis. Concepts, Quantitative Methods, and Future Directions", in *Annals of the American Association of Geographers*, 105(5):891-914.
- Andreatta M. (2021) *Archimede, l'arte della Misura*, Bologna, Il Mulino.
- Anselin L. (1988) *Spatial Econometric. Methods and Models*, New York, Kluwer.
- Anselin L. (1989) *What is Special About Spatial Data? Alternative Perspectives on Spatial Data Analysis*, UC Santa Barbara NCGIA Technical Report (89-4).
- Anselin L. (1995) "Local Indicators of Spatial Association - LISA", in *Geographical Analysis*, 27(2):93-115.
- Anselin L. (2000) "The Link between GIS and Spatial Analysis: GIS, Spatial Econometrics and Social Science Research", in *Journal of Geographical Systems*, 2(1):11-15.
- Anselin L. (2001) "Spatial Econometrics", in Baltagi B., ed., *A Companion to Theoretical Econometrics*, New York, John Wiley and Sons.
- Anselin L. (2002) "Under the Hood. Issues in the Specification and Interpretation of Spatial Regression Models", in *Agricultural Economics*, 27(3):247-267.
- Anselin L. (2005) *Spatial Regression Analysis in R: A Workbook*, University of Illinois.

- Anselin L. (2010) "Thirty Years of Spatial Econometrics", in *Papers in Regional Science*, 89(1):3-25.
- Anselin L. (2012) "From SpaceStat to CyberGIS: Twenty Years of Spatial Data Analysis Software", in *International Regional Science Review*, 35(2):131-157.
- Anselin L. (2020) "Spatial Data Science", in *International Encyclopedia of Geography. People, the Earth, Environment and Technology*, New York, Wiley, 1-6. DOI: 10.1002/9781118786352.
- Anselin L., Murray A.T., Rey S.J. (2000) "Spatial Analysis", in *Criminal Justice 2000: Measurement and Analysis of Crime and Justice*, 4.
- Bagolini B., Dalmeri G. (1987) "I siti mesolitici del Colbricon (Trentino). Analisi spaziale e fruizione del territorio", in *Preistoria alpina*, 23:7-188.
- Bao S., Anselin L., Martin D., Stralberg D. (2000) "Seamless Integration of Spatial Statistics and GIS: The S-PLUS for ArcView and the S+Grassland Links", in *Journal of Geographical Systems*, 2(3):287-306.
- Barnes T.J. (1998) "A History of Regression. Actors, Networks, Machines, and Numbers", in *Environment and Planning A*, 30(2):203-233.
- Barnes T.J. (2004) "Placing Ideas: Genius Loci, Heterotopia and Geography's Quantitative Revolution", in *Progress in Human Geography*, 28(5):565-595.
- Barnes T.J. (2010) "Taking the Pulse of the Dead: History and Philosophy of Geography, 2008-2009", in *Progress in Human Geography*, 34(5):668-677.
- Bauer J.R., Rose K. (2015) "Variable Grid Method. An Intuitive Approach for Simultaneously Quantifying and Visualizing Spatial Data and Uncertainty", in *Transactions in GIS*, 19(3):377-397.
- Beaumont, J.R. (1991) "Spatial Decision Support Systems: Some Comments with Regard to their Use in Market Analysis", *Environment and Planning A*, 23(3):311-317.
- Becker, R.A. (1994) "A Brief History of S", in Dirschedl P., Ostermann R., eds., *Computational Statistics*, Heidelberg, Physica.
- Benedicto J., Dinwiddy S., Gatti G., Lucas R., Lugert M. (2000) *GALILEO: Satellite System Design*. European Space Agency.
- Bergmann L., O'Sullivan D. (2018) "Reimagining GIScience for Relational Spaces", in *The Canadian Geographer/Le Géographe canadien*, 62(1):7-14.
- Bernatsky S., Smargiassi A., Johnson M., Kaplan G.G., Barnabe C., Svenson L., Brand A., Bertazzon S., Hudson M., Clarke A. E. (2015) "Fine Particulate Air Pollution, Nitrogen Dioxide, and Systemic Autoimmune Rheumatic Disease in Calgary, Alberta", in *Environmental Research*, 140:474-478.
- Berry B.J.L. (2004) "Epilogue: Spatial Analysis in Retrospect and Prospect", in Goodchild M.F., Janelle D.J., eds., *Spatially Integrated Social Science*, Oxford, Oxford University Press.
- Berry B.J.L., Marble D.F. (1968), *Spatial Analysis; A Reader in Statistical Geography*, Englewood Cliffs, Prentice-Hall.
- Bertazzon S. (2001), "Il nome del GIS. I sistemi verso una scienza dell'informazione geografica", in *Rivista geografica italiana*, 108(3):409-440.

- Bertazzon S. (2003) “Una definizione di contiguità per l’analisi di regressione spaziale in GISc. Aspetti concettuali e computazionali della dipendenza spaziale”, in *Rivista geografica italiana*, 110(2):247-280.
- Bertazzon S. (2005) “L’arte dell’analisi spaziale: perché saper scegliere i parametri del modello se li puoi semplicemente selezionare dal menu a cascata”, in Capineri C., Randelli F., Tinacci M., eds., *Conoscere il mondo: Vespucci e la modernità. Atti del Convegno Internazionale*. Firenze 28-29 ottobre 2004. *Memorie geografiche* NS 5:407-416.
- Bertazzon S. (2009) “Malattie cardiovascolari e fattori socioeconomici di rischio: un’analisi spaziale empirica su Calgary (Canada)”, in *Rivista geografica italiana*, 116(3):379-400.
- Bertazzon S., Calder-Bellamy C., Shahid R., Couloigner I., Wong R. (2020) “A Preliminary Spatial Analysis of the Association of Asthma and Traffic-Related Air Pollution in the Metropolitan Area of Calgary, Canada”, in *Atmosphere*, 11(10):1066.
- Bertazzon S., Carlon C., Critto A., Marcomini A., Zanetto G. (2000) “Integration of Spatial Analysis and Ecological Risk Assessment in a GIS Environment. The Case Study of the Venetian Lagoon Contaminated Sediments”, in *Proceedings of GIS and Environmental Modeling (GIS/EM4)*. CD-ROM, No. 122.
- Bertazzon S., Couloigner I. (2018) “Exploratory Spatial Analysis of Air Pollution Over a Large Metropolitan Area”, in *Rivista geografica italiana*, 125(4):525-548.
- Bertazzon S., Couloigner I., Mirzaei M. (2018) “Smart cities e inquinamento atmosferico: modelli di regressione spaziale (LUR) su dati da sensori low-cost e volunteered geographic information (VGI)”. in *Atti della XXII Conferenza Nazionale ASITA* (<http://atti.asita.it/ASITA2018/>), Bolzano, 27-29 Novembre 2018, 139-140.
- Bertazzon S., Couloigner I., Mirzaei M. (2021) Spatial Regression Modelling of Particulate Pollution in Calgary, Canada. *Geojournal*. Published online: 05 January 2021.
- Bertazzon S., Couloigner I., Underwood F. (2019) “Spatial Land Use Regression of Nitrogen Dioxide over a 5-Year Interval in Calgary, Canada”, in *International Journal of Geographical Information Science*, 33(7):1335-1354.
- Bertazzon S., Hanes A., Shahid R., Couloigner I. (2021) “The First to Be Hit: Spatio-Temporal Analysis of First-Wave COVID-19 Diffusion in Two Regions of Northern Italy”, in N. Pinto *et Alii*, eds., *Book of Extended Abstracts of European Colloquium on Theoretical and Quantitative Geography 2021*. Manchester, Spatial Policy and Analysis Laboratory, Manchester Urban Institute, University of Manchester, 402-406.
- Bertazzon S., Johnson M., Eccles K., Kaplan G.G. (2015) “Accounting for Spatial Effects in Land Use Regression for Urban Air Pollution Modeling”, in *Spatial and Spatio-temporal Epidemiology*, 14-15:9-21.
- Bertazzon S., Micheletti C. (2006) “Analisi integrate. Rischio ecologico da metalli pesanti e PCBs”, in Guerzoni S., Tagliapietra D., eds., *Atlante della laguna. Venezia fra terra e mare*, Venezia, Marsilio.

- Bertazzon S., Micheletti C., Critto A., Marcomini A. (2006) "Spatial Analysis in Ecological Risk Assessment: Pollutant Bioaccumulation in Clams *Tapes Philipinarum* in the Venetian Lagoon", in *Computers, Environment and Urban Systems*, 30(6):880-904.
- Bertazzon S., O'Hara P.D., Barrett O., Serra-Sogas N. (2014) "Geospatial Analysis of Oil Discharges Observed by the National Aerial Surveillance Program in the Canadian Pacific Ocean", in *Applied Geography*, 52:78-89.
- Bertazzon S., Olson S., Knudtson M. (2010) "A Spatial Analysis of the Demographic and Socio-Economic Variables Associated with Cardiovascular Disease in Calgary (Canada)", in *Applied Spatial Analysis and Policy*, 3(1):1-23.
- Bertazzon S., Shahid R. (2017) "Schools, Air Pollution, and Active Transportation. An Exploratory Spatial Analysis of Calgary, Canada", in *The International Journal of Environmental Research and Public Health*, 14(8):834.
- Bertazzon S., Waters N. (1996) "Immaginazioni GISgrafiche", in *Geotema*, 6:27-33.
- Bivand R., Pebesma E., Gómez-Rubio V. (2008) *Applied Spatial Data Analysis with R*, Berlin, Springer-Verlag.
- Bivand R., Piras G. (2015) "Comparing Implementations of Estimation Methods for Spatial Econometrics", in *Journal of Statistical Software*, 63:1-36.
- Bivand R., Wong D.W.S. (2018) "Comparing Implementations of Global and Local Indicators of Spatial Association", in *TEST*, 27(3):716-748.
- Boffi M. (2004) *Scienza dell'informazione geografica. Introduzione ai GIS*, Bologna, Zanichelli.
- Boots B. (2000) "Using GIS to Promote Spatial Analysis", in *Journal of Geographical Systems*, 2(1):17-21.
- Boots B., Tiefelsdorf M. (2000) "Global and Local Spatial Autocorrelation in Bounded Regular Tessellations", in *Journal of Geographical Systems*, 2(4):319-348.
- Borruso G. (2013) "Cartografia e Informazione Geografica '2.0 e Oltre', Webmapping, Webgis. Un'introduzione", in *Bollettino dell'Associazione Italiana di Cartografia*, 147:7-15.
- Box P. W. (2000) "Specialized Aspects of GIS and Spatial Analysis. Garage Band Science and Dynamic Spatial Models", in *Journal of Geographical Systems*, 2(1):49-54.
- Bressan G. (2021) "Assessing the Positional Accuracy of Perceptual Landscape Data: A Study from Friuli Venezia Giulia, Italy", in *Transactions in GIS*, 25(2):642-671.
- Breusch T.S., Pagan A.R. (1979) "A Simple Test for Heteroscedasticity and Random Coefficient Variation", in *Econometrica*, 47:1287-1294.
- Brotton J. (2013) *La storia del mondo in dodici mappe*, Milano, Feltrinelli (ed. orig. *A History of the World in 12 Maps*, 2013, New York, Viking).
- Brunsdon C. (2016) "Quantitative Methods I. Reproducible Research and Quantitative Geography", in *Progress in Human Geography*, 40(5):687-696.
- Brunsdon C., Comber A. (2020) "Big Issues for Big Data: Challenges for Critical Spatial Data Analytics", in *Journal of Spatial Information Science*, 21:89-98.

- Brunsdon C., Comber A. (2021) "Opening Practice: Supporting Reproducibility and Critical Spatial Data Science", in *Journal of Geographical Systems*, 23(4):477-496.
- Bunge W. (1962) *Theoretical Geography*, Lund Studies in Geography, Series C, n. 1, C.W.K. Gleerup Lund (I ed.).
- Bunge W. (1966) *Theoretical Geography*, Lund Studies in Geography, Series C, C.W.K. Gleerup (II ed.).
- Burrough P.A., McDonnell R. (1986) *Principles of Geographical Information Systems*, Oxford, Oxford University Press.
- Burt J., Barber G., Rigby D.L. (2009) *Elementary Statistics for Geographers*, New York, Guilford Press.
- Cai C., Gao Y., Pan L., Zhu J. (2015) "Precise Point Positioning with Quad-Constellations: GPS, BeiDou, GLONASS and Galileo", in *Advances in Space Research*, 56(1):133-143.
- Calvino I. (1978) *Le città invisibili*, Torino, Einaudi.
- Capel H. (1987) *Filosofia e scienza nella geografia contemporanea*, Milano, Unicopli (ed. orig. *Filosofía y ciencia en la geografía contemporánea*, 1981, Barcellona, Barcanova).
- Carroll L. (2010) *Through the Looking Glass and What Alice Found There*, Old Saybrook, Tantor Media.
- Casetti E. (1972) "Generating Models by the Expansion Method. Applications to Geographical Research", *Geographical Analysis*, 4(1):81-91.
- Chambers J.M. (1998) *Programming with Data. A Guide to the S Language*, Berlin, Springer Science and Business Media.
- Chavarría Arnau A., Valente V. (2009) "Progetto ARMEP: gestione GIS e analisi spaziali", in *Progetto ARMEP: gestione GIS e analisi spaziali*, Archeologia dell'Architettura, XIV, 177-187.
- Chen B.Y., Li Q.Q., Wang D.G., Shaw S.L., Lam W.H.K., Yuan H., Fang Z.X. (2013) "Reliable Space-Time Prisms Under Travel Time Uncertainty", in *Annals of the American Association of Geographers*, 103(6):1502-1521.
- Chen X., Vo H., Aji A., Wang F. (2014) "High Performance Integrated Spatial Big Data Analytics", in *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data*, Dallas, Texas, Association for Computing Machinery, 11-14.
- Chun Y., Griffith D.A. (2016) "Measuring Spatial Dependence", in *International Encyclopedia of Geography. People, the Earth, Environment and Technology*, New York, Wiley, 1-6. DOI: 10.1002/9781118786352
- Clark G.L. (1992) "Commentary: GIS - What Crisis?", in *Environment and Planning A*, 24(3):321-322.
- Clarke K.C., Murray A.T. (2020) "Introduction to the Special Issue Commemorating Professor Waldo Tobler", in *Geographical Analysis*, 52(4):477-479.
- Cliff A.D., Ord J.K. (1970) "Spatial Autocorrelation. A Review of Existing and New Measures with Applications", in *Economic Geography*, 46(sup1):269-292.
- Cliff A.D., Ord J.K. (1981) *Spatial Processes: Models and Applications*, London, Pion.

- Cliff, A., Pred, A., Hägerstrand, T. (1992) "Classics in Human Geography Revisited: Hägerstrand, T. 1967: Innovation Diffusion as a Spatial Process. Chicago: University of Chicago Press. Translation and postscript by Allan Pred", in *Progress in Human Geography*, 16(4):541-544.
- Coccia M. (2000) *Trasferimento tecnologico: analisi spaziale*, Ceris-Cnr.
- Comber A., Wang Y., Lü Y., Zhang X., Harris P. (2018) "Hyper-Local Geographically Weighted Regression. Extending GWR through Local Model Selection and Local Bandwidth Optimization", in *Journal of Spatial Information Science*, 17:63-84.
- Conti Puorger A. (2018) *Analisi geografica e metodi statistici. Una introduzione*, Bologna, Patron.
- Cope M., Elwood S., eds. (2009) *Qualitative GIS: A Mixed Methods Approach*, Thousand Oaks, SAGE.
- Coppock J.T., Rhind D.W. (1991) "The History of GIS", in Maguire D.J., Goodchild M.F., Rhind D., eds., *Geographical Information Systems. Principles and Applications*, Vol. 1, London, Longman Scientific and Technical.
- Corna Pellegrini G., ed. (1987) *Aspetti e problemi della geografia*, 2 Voll., Milano, Marzorati.
- Corna Pellegrini G. Brusa C., eds. (1980) *La ricerca geografica in Italia 1960-1980*, Varese, Ask Edizioni.
- Costa P., Canestrelli E. (1983) *Agglomerazione urbana. Localizzazione industriale e mezzogiorno*, SVIMEZ, Collana monografia n. 49, Milano, Giuffrè.
- Couclelis H. (1999) "Space, Time, Geography", in Longley P.A., Goodchild M.F., Maguire D.J., Rhind D.W., eds., *Geographical Information Systems. Principles and Technical issues*, Second Edition, Vol. 1, New York, Wiley.
- Crampton J. (1995) "The Ethics of GIS", in *Cartography and Geographic Information Systems*, 22(1):84-89.
- Cressie N. (1989) "Geostatistics", in *The American Statistician*, 43(4):197-202.
- Cressie N. (1993) *Statistics for Spatial Data*, New York, Wiley.
- Cressie N., Kornak J. (2003) "Spatial Statistics in the Presence of Location Error with an Application to Remote Sensing of the Environment", in *Statistical Science*, 18(4):436-456.
- Cresswell T. (2014) "Déjà Vu All Over Again. Spatial Science, Quantitative Revolutions and the Culture of Numbers", in *Dialogues in Human Geography*, 4(1):54-58.
- Curry M.R. (1994) "Image, Practice and the Hidden Impacts of Geographic Information Systems", in *Progress in Human Geography*, 18(4):441-459.
- Dangermond J., Goodchild M. F. (2020) "Building Geospatial Infrastructure", *Geospatial Information Science*, 23(1):1-9.
- Dale M.R., Fortin, M.J. (2014) *Spatial Analysis: A Guide for Ecologists*, Cambridge, Cambridge University Press.
- De Albuquerque J.P., Herfort B., Brenning A., Zipf A. (2015) "A Geographic Approach for Combining Social Media and Authoritative Data towards Identifying Useful Information for Disaster Management", in *International Journal of Geographical Information Science*, 29(4):667-689.

- De Smith M.J., Goodchild M.F., Longley P. (2020) *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*, Market Harbourough, Troubador.
- Denis D.J., Docherty K.R. (2007) "Late Nineteenth Century Britain. A Social, Political, and Methodological Context for the Rise of Multivariate Statistics", *Electronic Journal of History of Probability and Statistics*, 3 (www.jehps.net/).
- Diggle P.J. (1979) "On Parameter Estimation and Goodness-of-Fit Testing for Spatial Point Patterns", in *Biometrics*, 35:87-101.
- Diggle P.J. (1983) *Statistical Analysis of Spatial Point Patterns*, London, Academic Press.
- Do Lee W., Qian M., Schwanen T. (2021) "The Association between Socioeconomic Status and Mobility Reductions in the Early Stage of England's COVID-19 Epidemic", in *Health and Place*, 69:102563.
- Dueker K.J. (1974) "Urban Geocoding", in *Annals of the American Association of Geographers*, 64(2):318-325.
- Eco U. (1977) *Come si fa una tesi di laurea*, Milano, Bompiani.
- Eco U. (2004) *La bustina di minerva*, Milano, Bompiani.
- Elwood S., ed. (2010) *Mixed Methods: Thinking, Doing, and Asking in Multiple Ways*, Thousand Oaks, SAGE.
- Esopo (1996) *Favole*, Milano, Mondadori.
- Evans J., van Donkelaar A., Martin R.V., Burnett R., Rainham D.G., Birkett N.J., Krewski D. (2013) "Estimates of Global Mortality Attributable to Particulate Air Pollution using Satellite Imagery", in *Environmental Research*, 120:33-42.
- Favretto A. (2000) *Strumenti per l'analisi geografica. GIS e telerilevamento*, Bologna, Patron.
- Feng X., Murray A.T. (2018) "Spatial Analytics for Enhancing Street Light Coverage of Public Spaces", in *Leukos*, 14(1):13-23.
- Ferro G. (1983) *Geografia e libertà. Temi e problemi di geografia umana*, Bologna, Patron.
- Feyerabend P.K. (1995) *Contro il metodo. Abbozzo di una teoria anarchica della conoscenza*. Milano, Feltrinelli (ed. orig. *Outline of an Anarchist Theory of Knowledge*, 1975, London, Verso).
- Fischer M., Scholten H.J., Unwin D. (1996) *Spatial Analytical Perspectives on GIS*, London, Taylor and Francis.
- Fisher P.F. (1997) "Editorial. Welcome to the International Journal of Geographical Information Science", in *International Journal of Geographical Information Science*, 11(1):1-3.
- Fisher W.D. (1971) "Econometric Estimation with Spatial Dependence", in *Regional and Urban Economics*, 1(1):19-40.
- Florax R., Folmer H., Rey S.J. (2003) "Specification Searches in Spatial Econometrics: The Relevance of Hendry's Methodology", in *Regional Science and Urban Economics*, 33(5):557-579.
- Foresman T.W. (1998) *The History of Geographic Information Systems: Perspectives from the Pioneers*, Upper Saddle River, Prentice Hall.

- Foresman, T. Luscombe, R. (2017) "The Second Law of Geography for a Spatially Enabled Economy" in *International Journal of Digital Earth*, 10(10):979-995.
- Fotheringham A.S. (1991) Some Random(ish) Thoughts on Spatial Decision Support Systems, *Environment and Planning A*, 23(12):1699-1700.
- Fotheringham A.S. (2009) "The Problem of Spatial Autocorrelation and Local Spatial Statistics", in *Geographical Analysis*, 41(4):398-403.
- Fotheringham A.S., Brunson C., Charlton M. (2000) *Quantitative Geography. Perspectives on Spatial Data Analysis*, Thousand Oaks, SAGE.
- Fotheringham A.S., Brunson C., Charlton M. (2003) *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*, New York, John Wiley and Sons.
- Fotheringham A.S., Charlton M., Brunson C. (1998) "Geographically Weighted Regression. A Natural Evolution of the Expansion Method for Spatial Data Analysis", *Environment and Planning A*, 30(11):1905-1927.
- Fotheringham A.S., Crespo R., Yao J. (2015) "Geographical and Temporal Weighted Regression (GTWR)", in *Geographical Analysis*, 47(4):431-452.
- Fotheringham A.S., Rogerson P. (1994) "Spatial Analysis and GIS", in Fotheringham A.S., Rogerson P., eds., *Spatial Analysis and GIS*, London, Taylor and Francis.
- Fotheringham A.S., Wegener M. (1999) *Spatial Models and GIS. New and Potential Models*, London, CRC press.
- Fox C.H., O'Hara P.D., Bertazzon S., Morgan K., Underwood F., Paquet P.C. (2016) "A Preliminary Spatial Assessment of Risk: Marine Birds and Chronic Oil Pollution on Canada's Pacific Coast", in *Science of the Total Environment*, 573:799-809.
- Frank A.U. (2000) "Technology and the Future of GIS and Spatial Analysis", in *Journal of Geographical Systems*, 2(1):99-105.
- Franklin R.S. (2021) "Introduction to the 50th Anniversary Special Issue: Happy Birthday, Geographical Analysis!" in *Geographical Analysis*, 53(1):3-12.
- Franklin R.S., Houlden V., Robinson C., Arribas-Bel D., Delmelle E.C., Demšar U., Miller H.J., O'Sullivan D. (2021) "Who Counts? Gender, Gatekeeping, and Quantitative Human Geography", in *The Professional Geographer*, 73(1):48-61.
- Galilei G. (1977) *Il Saggiatore*, Torino, Einaudi.
- Garfield S. (2016) *Sulle mappe. Il mondo come lo disegnamo*, Milano, Ponte alle Grazie (ed. orig. *On the Map*, 2012, Profile Books).
- Garrison W.L. (1956) "Applicability of Statistical-Inference to Geographical Research", in *The Geographical Review*, 46(3):427-429.
- Gatrell A.C. (1983) *Distance and Space: A Geographical Perspective*, Oxford, Oxford University Press.
- Gebbers R., Adamchuk V.I. (2010) "Precision Agriculture and Food Security", in *Science*, 327(5967):828-831.
- Gelfand A.E., Diggle P., Guttorp P., Fuentes M. (2010) *Handbook of Spatial Statistics*, London, CRC Press.
- Getis A. (1993) "Scholarship, Leadership, and Quantitative Methods", in *Urban Geography*, 14(6):517-525.

- Getis A. (2000) "Spatial Analysis and GIS. An Introduction", in *Journal of Geographical Systems*, 2(1):1-3.
- Getis A. (2007) "Reflections on Spatial Autocorrelation", in *Regional Science and Urban Economics*, 37(4):491-496.
- Getis A. (2008) "A History of the Concept of Spatial Autocorrelation. A Geographer's Perspective", in *Geographical Analysis*, 40(3):297-309.
- Getis A., Aldstadt J. (2004) "Constructing the Spatial Weights Matrix Using a Local Statistic", in *Geographical Analysis*, 36(2):90-104.
- Getis A., Ord J.K. (1992) "The Analysis of Spatial Association by Use of Distance Statistics", in *Geographical Analysis*, 24(3):189-206.
- Gibertoni D., Adja K., Golinelli D., Reno C., Regazzi L., Lenzi J., Sanmarchi F., Fantini M.P. (2021) "Patterns of COVID-19 Related Excess Mortality in the Municipalities of Northern Italy During the First Wave of the Pandemic", in *Health and Place*, 67:102508.
- Giordano A., Cole T. (2018) "The Limits of GIS. Towards a GIS of Place", in *Transactions in GIS*, 22(3):664-676.
- Gong P., Mu L. (2000) "Error Detection through Consistency Checking", in *Geographic Information Sciences*, 6(2):188-193.
- Goodchild M.F. (1986) *Spatial Autocorrelation*, CATMOG (Concepts and Techniques in Modern Geography) n. 47, Norwich, Geo Abstracts, University of East Anglia.
- Goodchild M.F. (1991) "Just the Facts", in *Political Geography Quarterly*, 10(4):335-337.
- Goodchild M.F. (1992) "Geographical Information Science", in *International Journal of Geographical Information Systems*, 6(1):31-45.
- Goodchild M.F. (1996) "I GIS e la ricerca geografica", in *Geotema*, 6:8-18
- Goodchild M.F. (2000) "Spatial Analysts and GIS Practitioners. The Current Status of GIS and Spatial Analysis", in *Journal of Geographical Systems*, 2(1):5-10.
- Goodchild M.F. (2001) "A Geographer Looks at Spatial Information Theory", in Montello D.R., ed., *Foundations of Geographic Information Science - International Conference, COSIT 2001, Proceedings*. Lecture Notes in Computer Science 2205, Berlin, Springer-Verlag, 1-13.
- Goodchild M.F. (2004) "The Validity and Usefulness of Laws in Geographic Information Science and Geography", in *Annals of the American Association of Geographers*, 94(2):300-303.
- Goodchild M.F. (2007) "Citizens as Sensors. The World of Volunteered Geography", in *GeoJournal*, 69(4):211-221.
- Goodchild M.F. (2009) "What Problem? Spatial Autocorrelation and Geographic Information Science", in *Geographical analysis*, 41(4):411-417.
- Goodchild, M.F. (2010) "Towards Geodesign: Repurposing Cartography and GIS?", in *Cartographic Perspectives*, 66:7-22.
- Goodchild M.F. (2013) "Prospects for a Space-Time GIS", in *Annals of the American Association of Geographers*, 103(5):1072-1077.
- Goodchild M.F. (2018) "Reimagining the History of GIS", in *Annals of GIS*, 24(1):1-8.

- Goodchild M.F., Janelle D. (2010) "Toward Critical Spatial Thinking in the Social Sciences and Humanities", in *GeoJournal*, 75(1):3-13.
- Goodchild M.F., Kemp K.K. (1992) "NCGIA Education Activities: the Core Curriculum and beyond", in *International Journal of Geographical Information Systems*, 6(4):309-320.
- Gore A. (2006) *An Inconvenient Truth. The Planetary Emergency of Global Warming and What we Can Do About It*, Emmaus, Rodale.
- Goss J. (1995) "We Know Who You Are and We Know Where You Live: The Instrumental Rationality of Geodemographic Systems", in *Economic geography*, 71(2):171-198.
- Granger C.W. (1969) "Spatial Data and Time Series Analysis", in Scott A.J., ed., *Studies in Regional Science*, London, Pion.
- Granger C.W. (1974) "Aspects of the Analysis and Interpretation of Temporal and Spatial Data", in *Journal of the Royal Statistical Society. Series D (The Statistician)*, 24(3):197-210.
- Granger C.W. (1981) "Some Properties of Time Series Data and their Use in Econometric Model Specification", in *Journal of econometrics*, 16(1):121-130.
- Griffith D.A. (1992) "What is Spatial Autocorrelation? Reflections on the Past 25 Years of Spatial Statistics", in *L'Espace géographique*, 21:265-280.
- Griffith D.A. (2008) "Spatial-Filtering-Based Contributions to a Critique of Geographically Weighted Regression (GWR)", in *Environment and Planning A*, 40(11):2751-2769.
- Griffith D.A. (2013) "Establishing Qualitative Geographic Sample Size in the Presence of Spatial Autocorrelation", in *Annals of the American Association of Geographers*, 103(5):1107-1122.
- Griffith D.A. (2019) "Negative Spatial Autocorrelation. One of the Most Neglected Concepts in Spatial Statistics", in *Stats*, 2(3):388-415.
- Griffith D.A., Amrhein C.G. (1977) *Multivariate Statistical Analysis for Geographers*, Upper Saddle River, Prentice-Hall.
- Griffith D.A., Paelinck J.H. (2007) "An Equation by any other Name Is still the Same: on Spatial Econometrics and Spatial Statistics", in *The Annals of Regional Science*, 41(1):209-227.
- Grinberger A.Y., Novack T. (2021) "The Geographical and Cultural Aspects of Geoinformation", in *Transactions in GIS*, 25(2):577-578.
- Guarrasi V., ed., (1996) *Realtà virtuali: nuove dimensioni dell'immaginazione geografica*, Geotema, 6, pp. 1-101.
- Guarrasi V. (2002) "Eterotopia del paesaggio e retorica cartografica", in *Atlante Virtuale*, Palermo, Università degli Studi di Palermo.
- Gujarati D. (2003) *Basic Econometrics*, Boston, McGraw Hill.
- Haggett P. (2008a) "The Local Shape of Revolution: Reflections on Quantitative Geography at Cambridge in the 1950s and 1960s", in *Geographical Analysis*, 40(3):336-352.
- Haggett P. (2008b) "The Spirit of Quantitative Geography (1)", in *Geographical Analysis*, 40(3):226-228.

- Haining R. (1990) *Spatial Data Analysis in the Social and Environmental Sciences*, Cambridge, Cambridge University Press.
- Haining R. (2003) *Spatial Data analysis. Theory and Practice*, Cambridge, Cambridge University Press.
- Haining R. (2009) "The Special Nature of Spatial Data", in Fotheringham A.S., Rogerson, P., eds., *The SAGE Handbook of Spatial Analysis*, Thousand Oaks, SAGE.
- Halder S., Michel B. (2019) "Editorial – This Is Not an Atlas", kollektiv orangotango+, ed., *This Is Not an Atlas. Transcript-Verlag*, 12-25, <https://library.oapen.org/handle/20.500.12657/27378>
- Harley J. B. (1989) "Deconstructing the Map", in *Cartographica*, 26(2):1-20.
- Harris R., Nicholas T., Souch C., Singleton A., Orford S., Keylock C., Jarvis C., Brunson, C. (2014) "Geographers Count: A Report on Quantitative Methods in Geography" in *Enhancing Learning in the Social Sciences*, 6(2):43-58.
- Harris T., Weiner D. (1998) "Empowerment, Marginalization, and 'Community-Integrated' GIS", in *Cartography and Geographic Information Systems*, 25(2):67-76.
- Harvey D. (1969) *Explanation in Geography*, London, Edward Arnold and St Martin's Press.
- Hepple L.W. (2001) "Multiple Regression and Spatial Policy Analysis. George Udny Yule and the Origins of Statistical Social Science", in *Environment and Planning D*, 19(4):385-407.
- Heuvelink G.B., Burrough P.A., Stein A. (1989) "Propagation of Errors in Spatial Modelling with GIS", in *International Journal of Geographical Information System*, 3(4):303-322.
- Hoek G., Beelen R.M.J., de Hoogh K., Vienneau D., Gulliver J., Fischer P., Briggs D. (2008) "A Review of Land-Use Regression Models to Assess Spatial Variation of Outdoor Air Pollution", in *Atmospheric Environment*, 42(33):7561-7578.
- Hunt D., Stevenson S.A. (2017) "Decolonizing Geographies of Power: Indigenous Digital Counter-Mapping Practices on Turtle Island, in *Settler Colonial Studies*, 7(3):372-392.
- Isard W. (1960) *Methods of Regional Analysis. An Introduction to Regional Science*, Cambridge (Mass.), M.I.T. Press.
- Janssen M., Charalabidis Y., Zuiderwijk A. (2012) "Benefits, Adoption Barriers and Myths of Open Data and Open Government", in *Information Systems Management*, 29(4):258-268.
- Jensen, J.R.(2007) *Remote Sensing of the Environment: An Earth Resource Perspective*, Second edition. Upper Saddle River: Pearson Prentice Hall.
- Jerrett M., McConnell R., Wolch J., Chang R., Lam C., Dunton G., Gilliland F., Lurmann F., Islam T., Berhane K. (2014) "Traffic-Related Air Pollution and Obesity Formation in Children: A Longitudinal, Multilevel Analysis", in *Environmental Health*, 13(1):1-9.
- Johnson S. (2006) *The Ghost Map. The Story of London's Most Terrifying Epidemic—And How It Changed Science, Cities, and the Modern World*, London, Penguin.
- Johnston J., DiNardo J. (1997) *Econometric Methods*, New York, McGraw Hill.

- Johnston R. (2000) "On Disciplinary History and Textbooks: Or Where Has Spatial Analysis Gone?", in *Australian Geographical Studies*, 38(2):125-137.
- Johnston R., Harris R., Jones K., Manley D., Sabel C.E., Wang W.W. (2014) "Mutual Misunderstanding and Avoidance, Misrepresentations and Disciplinary Politics: Spatial Science and Quantitative Analysis in (United Kingdom) Geographical Curricula", in *Dialogues in Human geography*, 4(1):3-25.
- Johnston R., Harris R., Jones K., Manley D., Wang W.W., Wolf L. (2019) "Quantitative Methods I. The World We Have Lost-or where We Started from", in *Progress in Human Geography*, 43(6):1133-1142.
- Johnston R., Jones K. (2020) "A Classic that Wasn't. Statistical Geography and Paths only Later Taken", in *Progress in Human Geography*, 44(2):357-373.
- Jordan G.A., Anastasopoulos P.C., Peeta S., Somenahalli S., Rogerson P.A. (2019) "Identifying Elderly Travel Time Disparities Using a Correlated Grouped Random Parameters Hazard-Based Duration Approach", in *Research in Transportation Business and Management*, 30, 100369.
- Kalbasi R., Janowicz K., Reitsma F., Boerboom L., Alesheikh A. (2014) "Collaborative Ontology Development for the Geosciences", in *Transactions in GIS*, 18(6):834-851.
- Kemp K.K., Goodchild M.F. (1991) "Developing a Curriculum in GIS: the NCGIA Core Curriculum Project", in *Cartographica*, 28(3), 39-54.
- Kennedy M. (1994) "Geographical Information Systems: Principles and Applications, eds. David J. Maguire, Michael F. Goodchild, and David W. Rhind", *Annals of the American Association of Geographers*, 84(1):172-173.
- Kitchin R. (2014) "Big Data, New Epistemologies and Paradigm Shifts", in *Big Data and Society*, 1(1): 1-12.
- Klein L.J., Marianno F.J., Albrecht C.M., Freitag M., Lu S., Hinds N., Shao X., Rodriguez S.B., Hamann H.F. (2015) "PAIRS: A Scalable Geo-Spatial Data Analytics Platform", in *2015 IEEE International Conference on Big Data (Big Data)*, Piscataway, IEEE, 1290-1298.
- Klippel A., Hardisty F., Li R. (2011) "Interpreting Spatial Patterns. An Inquiry into Formal and Cognitive Aspects of Tobler's First Law of Geography", in *Annals of the American Association of Geographers*, 101(5):1011-1031.
- Koyré A. (2000) *Dal mondo del pressappoco all'universo della precisione*, Torino, Einaudi (ed. orig. "Du monde de l'«à-peu-près» à l'univers de la précision", in *Critique*, 1948, n. 28).
- Kraak M.-J., Ormeling F. (2009) *Cartography: Visualization of Geospatial Data*, Upper Saddle River, Prentice Hall.
- Krisztin T., Piribauer P., Wögerer M. (2020) "The Spatial Econometrics of the Coronavirus Pandemic", in *Letters in Spatial and Resource Sciences*, 13(3):209-218.
- Kumar M. (1988) "World Geodetic System 1984. A Modern and Accurate Global Reference Frame", in *Marine Geodesy*, 12(2):117-126.
- Kwan M.P. Schwanen T. (2009) "Quantitative Revolution 2. The Critical (Re)Turn", in *Professional Geographer*, 61(3):283-291.

- Lafreniere D. Gilliland J. (2015) “All the World’s a Stage. A GIS Framework for Recreating Personal Time-Space from Qualitative and Quantitative Sources”, in *Transactions in GIS*, 19(2):225-246.
- Lake R.W. (1993) “Planning and Applied Geography. Positivism, Ethics, and Geographic Information Systems”, in *Progress in Human Geography*, 17(3):404-413.
- Lando F. (1985) “L’analisi delle componenti principali tra il mito dell’oggettivo e la manipolazione cosciente dei dati”, in P. Pagnini, ed., *Geografia del Principe*, Milano, Unicopli.
- Lando F. (2020) *Per una storia del moderno pensiero geografico. Passaggi significativi*, Milano, FrancoAngeli.
- Li J., Qin Q., Han J., Tang L.-A., Lei K.H. (2015) “Mining Trajectory Data and Geotagged Data in Social Media for Road Map Inference”, in *Transactions in GIS*, 19(1):1-18.
- Lin H., Batty M., Jørgensen S.E., Fu B., Konecny M., Voinov A., Torrens P., Lu G., Zhu A.X., Wilson J.P., Gong J., Kolditz O., Bandrova T., Chen M. (2015) “Virtual Environments Begin to Embrace Process-Based Geographic Analysis”, in *Transactions in GIS*, 19(4):493-498.
- Livingstone D.N. (1992) *The Geographical Tradition. Episodes in the History of a Contested Enterprise*, Oxford, Blackwell.
- Liu X., Bertazzon S. (2016) “Fine Scale Spatiotemporal Modelling of Urban Air Pollution”, in Miller J., O’Sullivan D., Wiegand N., eds., *Proceedings of GIScience 2016. The Ninth International Conference on Geographic Information Science*. Montreal, Canada, September 27-30, 2016. Volume 9927 of Lecture Notes in Computer Science, Berlin, Springer-Verlag, 210-224.
- Liu X., Bertazzon S., Villeneuve P.J., Johnson M., Stieb D., Coward S., Tanyingoh D., Windsor J.W., Underwood F., Hill M.D., Rabi D., Ghali W.A., Wilton S.B., James M.T., Graham M., McMurtry M.S., Kaplan G.G. (2020). “Temporal and Spatial Effect of Air Pollution on Hospital Admissions for Myocardial Infarction: A Case-Crossover Study”, *Canadian Medical Association Open Access Journal*, 8(4):E619-E626.
- Longley P.A. (2000) “The Academic Success of GIS in Geography. Problems and Prospects”, in *Journal of Geographical Systems*, 2(1):37-42.
- Longley P.A. Batty M. (1997) *Spatial Analysis: Modeling in a GIS Environment*, New York, John Wiley and Sons.
- Maguire D. J. (1991) “An Overview and Definition of GIS”, in Maguire D.J., Goodchild M.F., Rhind D., eds., *Geographical Information Systems. Principles and Applications*, Vol. 1, London, Longman Scientific and Technical.
- Maguire D.J., Goodchild M.F., Rhind D. (1991) *Geographical Information Systems. Principles and Applications*, London, Longman Scientific and Technical.
- Matheron G. (1963) “Principles of Geostatistics”, in *Economic Geology*, 58(8):1246-1266.
- Mathur H., Bertazzon S. (2009) “Rasterizing Census Geography: Definition and Optimization of a Regular Grid”, in Sester M., Bernard L., Paelke V., eds., *Advances in GIScience. Proceedings of the 12th AGILE Conference*. Lecture

- Notes in Geoinformation and Cartography, Hannover, Germany 2-5 June 2009, Berlin, Springer-Verlag, 251-269.
- McCann P., Mudambi R. (2005) "Analytical Differences in the Economics of Geography: The Case of the Multinational Firm", in *Environment and Planning A*, 37(10):1857-1876.
- McHaffie P.H. (1995) "Manufacturing Metaphors. Public Cartography, the Market, and Democracy", in J. Pickles, ed., *Ground Truth. The Social Implications of Geographic Information Systems*, New York, Guilford Press.
- Medina R.M., Hepner G.F. (2011) "Advancing the Understanding of Sociospatial Dependencies in Terrorist Networks", in *Transactions in GIS*, 15(5):577-597.
- Migliaccio F., Carrion D. (2016) *Sistemi informativi territoriali: principi e applicazioni*, Torino, UTET Università.
- Miller H.J. (2004) "Tobler's First Law and Spatial Analysis", in *Annals of the American Association of Geographers*, 94(2):284-289.
- Miller H.J. (2007) "Place-Based versus People-Based Geographic Information Science", in *Geography Compass*, 1(3):503-535.
- Miller H.J. (2018) "Geographic Information Science II. Mesogeography: Social Physics, Giscience and the Quest for Geographic Knowledge", in *Progress in Human Geography*, 42(4):600-609.
- Miller H.J., Goodchild M.F. (2015) "Data-Driven Geography", in *GeoJournal*, 80(4):449-461.
- Mirzaei M., Bertazzon S., Couloigner I., Farjad B. (2021) "Assessing the Potential of Artificial Intelligence (Artificial Neural Networks) in Predicting the Spatiotemporal Pattern of Wildfire-Generated PM_{2.5} Concentration", in *Geomatics*, 1(1):18-33.
- Monnat S.M. Brown D.L. (2017) "More than a Rural Revolt. Landscapes of Despair and the 2016 Presidential Election", in *Journal of Rural Studies*, 55:227-236.
- Monti A. (2007) "GIS, spazio antropizzato e strategie comportamentali: analisi spaziali per la simulazione e l'interpretazione dei comportamenti umani", in *Atti della 11a Conferenza Nazionale ASITA*, <http://atti.asita.it/Asita2007.html>.
- Morrill R. (1993) "Geography, Spatial Analysis, and Social Science", in *Urban Geography*, 14(5):442-446.
- Morrison J. (1991) "The Organizational Home for GIS in the Scientific Professional Community", in Maguire D.J., Goodchild M.F., Rhind D., eds., *Geographical Information Systems. Principles and Applications*, Vol. 1, London, Longman Scientific and Technical.
- Mundy B.E. (2000) *The Mapping of New Spain: Indigenous Cartography and the Maps of the Relaciones Geográficas*, Chicago, University of Chicago Press.
- Murgante B., Borruso G., Balletto G., Castiglia P., Dettori M. (2020) "Why Italy first? Health, Geographical and Planning Aspects of the COVID-19 Outbreak", in *Sustainability*, 12(12):5064.
- Murray A.T. (2010) "Quantitative Geography", in *Journal of Regional Science*, 50(1):143-163.
- Murray A.T. (2018) "Evolving Location Analytics for Service Coverage Modeling", in *Geographical Analysis*, 50(3):207-222.

- Murray A.T. (2020) “Significance Assessment in the Application of Spatial Analytics”, in *Annals of the American Association of Geographers*, 111(6):1740-1755.
- Murray A.T., Grubestic T.H. (2019) “Evolving Regional Analytics in a Rural World”, in *International Regional Science Review*, 42(5-6):374-399.
- Muscarà L., ed. (1998) *GIS e geografia politica*. Atti del convegno tenutosi a Trieste il 27 gennaio 1997, Padova, CEDAM.
- Namgyal J., Tenzin T., Checkley S., Lysyk T.J., Rinchen S., Gurung R.B., Dorjee S., Couloigner I., Cork S.C. (2021) “A Knowledge, Attitudes, and Practices Study on Ticks And Tick-Borne Diseases in Cattle among Farmers in a Selected Area of Eastern Bhutan”, in *PloS one*, 16, e0247302.
- O’Sullivan D., Bergmann L., Thatcher J.E. (2018) “Spatiality, Maps, and Mathematics in Critical Human Geography. Toward a Repetition with Difference”, in *The Professional Geographer*, 70(1), 129-139.
- Obermeyer N.J., Pinto J.K. (1994) *Managing Geographic Information Systems*, New York, Guilford Press.
- Odd-Ambrosetti M., Turco A., Zanetto G. (1983) “Spazi imprecisi e regionalizzazione”, in *Ricerche Economiche*, 37(4):547-571.
- Onsrud H.J., Rushton G. (1995) *Sharing Geographic Information*, New Brunswick, Center for Urban Policy Research.
- Openshaw S. (1977) “A Geographical Solution to Scale and Aggregation Problems in Region-Building, Partitioning, and Spatial Modelling”, in *Transactions of the Institute of British Geographers*, n.s. 2:459-472.
- Openshaw S. (1983) *The Modifiable Aerial Unit Problem* CATMOG (Concepts and Techniques in Modern Geography) n. 38, Norwich, Geo Abstracts, University of East Anglia.
- Openshaw S. (1991) “A View on the GIS Crisis in Geography, or, Using GIS to Put Humpty-Dumpty back together again”, in *Environment and Planning A*, 23(5):621-628.
- Openshaw S. (1992) “Further Thought on Geography and GIS. A Reply”, in *Environment and Planning A*, 24(4):463-466.
- Openshaw S. (1994) “GIS Crime and Spatial Analysis”, in *Proceedings of GIS and Public Policy Conference*, Coleraine, Ulster Business School, 22-35.
- Openshaw S. (1998) “Some Trends and Future Perspectives for Spatial Analysis in GIS”, in *Geographic Information Sciences*, 4(1-2):5-13.
- Openshaw S., Albanides S. (1999) “Applying Geocomputation to the Analysis of Spatial Distributions”, in Longley P.A., Goodchild M.F., Maguire D.J., Rhind D.W., eds., *Geographical Information Systems. Principles and Technical issues*, Second Edition, Vol. 1, New York, Wiley.
- Openshaw S., Taylor P.J. 1979 “A million or so correlation coefficients: Three experiments on the modifiable areal unit problem”, in N. Wrigley, ed. *Statistical Applications in the Spatial Sciences*, London, Pion.
- Ord J.K., Getis A. (1995) “Local Spatial Autocorrelation Statistics: Distributional Issues and an Application”, in *Geographical Analysis*, 27(4):286-306.
- Paelnick J., Klaassen L. (1979) *Spatial Econometrics*, Farnbrough, Saxon House.

- Pagnini P., ed. (1985) *Geografia per il principe. Teoria e misura dello spazio geografico, Omaggio ad Eliseo Bonetti*, Milano, Unicopli.
- Pandey V., Kipf A., Neumann T., Kemper A. (2018) "How Good are Modern Spatial Analytics Systems?", in *Proceedings of the VLDB Endowment*, 11:1661-1673.
- Parkinson B.W., Enge P., Axelrad P., Spilker Jr. J.J. (1996) *Global Positioning System. Theory and Applications*, Reston, American Institute of Aeronautics and Astronautics.
- Pesaresi C. (2017) *Applicazioni GIS: principi metodologici e linee di ricerca: esercitazioni ed esemplificazioni guida*, Torino, Utet Università.
- Pesaresi C. (2020) "A Geographical and Crosscutting Look at the COVID-19 Pandemic in an International Framework. Introduction", in *J-READING-Journal of Research and Didactics in Geography*, 2:13-19.
- Pesaresi C., Marta, M. (2014) "Applicazioni GIS per l'analisi dell'urbanizzazione nella provincia di Napoli. Un'analisi multitemporale in aree esposte a elevato rischio vulcanico", *Bollettino dell'Associazione Italiana di Cartografia*, 150:34-51.
- Pequet D.J. (1994) "It's about Time. A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems", in *Annals of the American Association of Geographers*, 84(3):441-461.
- Pickles J. (1991) "Geography, GIS, and the Surveillant Society", In *Papers and Proceedings of Applied Geography Conferences*, 14(8):80-91.
- Pickles J., ed. (1995a) *Ground Truth. The Social Implications of Geographic Information Systems*, New York, Guilford Press.
- Pickles J. (1995b) "Representations in an Electronic Age. Geography, GIS, and Democracy", in J. Pickles, ed., *Ground Truth. The Social Implications of Geographic Information Systems*, New York, Guilford Press.
- Pickles J. (1997) "Tool or science? GIS, Technoscience, and the Theoretical Turn", *Annals of the American Association of Geographers*, 87(2):363-372.
- Pocewicz A., Nielsen-Pincus M., Brown G., Schnitzer R. (2012) "An Evaluation of Internet Versus Paper-based Methods for Public Participation Geographic Information Systems (PPGIS)", in *Transactions in GIS*, 16(1):39-53.
- Poorthuis A., Zook M. (2020) "Being Smarter about Space. Drawing Lessons from Spatial Science", in *Annals of the American Association of Geographers*, 110(2):349-359.
- Popper K.R. (1994) *Congetture e confutazioni. Lo sviluppo della conoscenza scientifica*, Bologna, Il Mulino (ed. orig. *Conjectures and Refutations*, 1969, London, Routledge and Kegan).
- Popper K.R. (2009) *Poscritto alla Logica della scoperta scientifica. Il realismo e lo scopo della scienza*, vol. I, Milano, Il Saggiatore (ed. orig. *Realism and the aim of Science from the Postscript to the logic of scientific discover*, 1982, London, Hutchinson).
- Porta S., Crucitti P., Latora V. (2006), "The Network Analysis of Urban Streets: A Dual Approach", in *Physica A: Statistical Mechanics and its Applications*, 369(2):853-866.

- Preston B., Wilson M.W. (2014) "Practicing GIS as Mixed Method. Affordances and Limitations in an Urban Gardening Study", in *Annals of the American Association of Geographers*, 104(3):510-529.
- Primi A., Marchioro C. (2020) "Time-Geography Approach During the COVID-19 Emergency in Italy: A Constraints Study Aat National and Local Level", in *J-READING-Journal of Research and Didactics in Geography*, 2(9):97-112.
- Pumain, D. ed. (1991) *Spatial Analysis and Population Dynamics: 6th European Colloquium of Theoretical and Quantitative Geography; Held at Chantilly (France) in September 1989*, New Barnet, John Libbey.
- Reynolds R.B. (1956) "Statistical Methods in Geographical Research", in *Geographical Review*, 46(1):129-132.
- Ripley B.D. (1981) *Spatial Statistics*, New York, Wiley.
- Ripley B.D. (1988) *Statistical Inference for Spatial Processes*, Cambridge, Cambridge University Press.
- Rogerson P.A. (2014) *Statistical Methods for Geography. A Student's Guide*, Thousand Oaks, SAGE.
- Rogerson P.A. (2019) "I Dream of Gini. Measures of Population Concentration and their Application to US Population Distribution", in *Population, Place, and Spatial Interaction*, 40:1-17.
- Rogerson P.A., Fotheringham A.S. Pellegrini P. (1995) *NCGLA Research initiative 14: GIS and spatial analysis*, Closing Report, Buffalo, University of Buffalo.
- Rosenkrantz L., Schuurman N., Bell N., Amram O. (2021), "The Need for GIScience in Mapping COVID-19", in *Health and Place*, 67:102389.
- Russo M. (2008) "Introduzione", in Russo M., ed., *L'industria meccanica in Italia. Analisi spaziale delle specializzazioni produttive 1951-2001*, Roma, Carocci.
- Salmond J.A., Tadaki M., Dickson M. (2017) "Can Big Data Tame a 'Naughty' World?", in *The Canadian Geographer/Le Géographe canadien*, 61(1):52-63.
- Samet H. (1984) "The Quadtree and Related Hierarchical Data Structures", in *ACM Computing Surveys (CSUR)*, 16(2):187-260.
- Sanna V. (2012) "Analisi Spaziale in ambiente GIS: Un caso di studio". *Atti della XLIV Riunione scientifica: la preistoria e la protostoria della Sardegna*, Cagliari, Barumini, Sassari 23-28 novembre 2009, Vol. 2-3, 1195-1200.
- Santini L., Miracco A.M., Santucci A. (2018) "Urban Regenerations. Application of Multi Criteria Spatial Analysis for the Redevelopment of the Military Barracks in the Historic Centre of Pisa", in *J-READING-Journal of Research and Didactics in Geography*, 2(7):13-26.
- Šavrič B., Patterson T., Jenny B. (2019) "The Equal Earth Map Projection", in *International Journal of Geographical Information Science*, 33(3):454-465.
- Scheider S., Nyamsuren E., Krüger H., Xu H. (2020) "Why Geographic Data Science is not a Science", in *Geography Compass*, 14(11):e12537.
- Schuurman N. (2000) "Trouble in the Heartland. GIS and its Critics in the 1990s", in *Progress in Human Geography*, 24(4):569-590.
- Schuurman N. (2006) "Formalization Matters. Critical GIS and Ontology Research", in *Annals of the American Association of Geographers*, 96(4):726-739.

- Schuurman, N. (2009) "Critical GIScience in Canada in the New Millennium", in *The Canadian Geographer/Le Géographe canadien*, 53(2):139-144.
- Shahid R., Bertazon S. (2015) "Local Spatial Analysis and Dynamic Simulation of Childhood Obesity and Neighbourhood Walkability in a Major Canadian City", in *AIMS Public Health*, 2(4):616.
- Sheppard E. (1993a) "Automated Geography: What Kind of Geography for What Kind of Society?", in *The Professional Geographer*, 45(4):457-460.
- Sheppard E. (1993b) "GIS and Society. Ideal and Reality", in *Proceedings of NCGIA Geographic Information and Society Workshop. Friday Harbor, Washington, November 11-14, 1993*.
- Sheppard E. (1995) "GIS and Society: towards a Research Agenda", in *Cartography and Geographic Information Systems*, 22(1):5-16.
- Sheppard, E. (2001) "Quantitative Geography: Representations, Practices, and Possibilities" in *Environment and Planning D: Society and space*, 19(5):535-554.
- Sheppard E. (2005) "Knowledge Production through Critical GIS. Genealogy and Prospects", in *Cartographica*, 40(4):5-21.
- Sheppard E., Couclelis H., Graham S., Harrington J., Onsrud H. (1999) "Geographies of the Information Society", in *International Journal of Geographical Information Science*, 13(8):797-823.
- Sletto B. (2009) "Indigenous Cartographies", in *Cultural Geographies*, 16(2):147-152.
- Smith N. (1992) "History and Philosophy of Geography. Real Wars, Theory Wars", in *Progress in Human Geography*, 16(2):257-271.
- Stein R., Mann C., Stewart III C., Birenbaum Z., Fung A., Greenberg J., Kawsar F., Alberda G., Alvarez R., Atkeson L. (2020) "Waiting to Vote in the 2016 Presidential Election. Evidence from a Multi-County Study", in *Political Research Quarterly*, 73(2):439-453.
- Stimson R. J. (2008) "A Personal Perspective from Being a Student of the Quantitative Revolution", in *Geographical Analysis*, 40(3):222-225.
- Sui D. (2001) "Terra Incognita and Limits of Computation: whither GIScience?", in *Computers, Environment and Urban Systems*, 6(25):529-533.
- Sui D. (2004) "Tobler's First Law of Geography. A Big Idea for a Small World?", in *Annals of the American Association of Geographers*, 94(2):269-277.
- Sun K., Chen J., Viboud C. (2020) "Early Epidemiological Analysis of the Coronavirus Disease 2019 Outbreak Based on Crowdsourced Data: A Population-Level Observational Study", in *The Lancet Digital Health*, 2(4):e201-e208.
- Tallone O., Lando F. (2009) "Industrializzazione e terziarizzazione. I processi di diffusione nell'area veneto-friulana", in F. Lando, ed., *La geografia dell'industrializzazione nel secondo dopoguerra. La situazione nazionale ed il caso veneto-friulano*, Padova, CEDAM.
- Tambassi T. (2018) "Ontologia della geografia e rappresentazioni cartografiche: uno spunto critico", in *Bollettino della Società Geografica Italiana*, Serie 14, 1(1):19-26.
- Tattara G., Volpe M. (2001) "I distretti industriali: definizioni e storia", in Tattara G., ed., *Il piccolo che nasce dal grande. Le molteplici facce dei distretti industriali veneti*, Milano, FrancoAngeli.

- Taylor P. (1990) "GKS", in *Political Geography Quarterly*, 9(3):211-212.
- Taylor P., Johnston R. (1995), "Geographic Information Systems and Geography", in J. Pickles, ed., *Ground Truth. The Social Implications of Geographic Information Systems*, New York, Guilford Press.
- Taylor P., Overton M. (1991) "Commentary. Further Thoughts on Geography and GIS", in *Environment and Planning A*, 23(8):1087-1094.
- Thatcher J., Bergmann L., O'Sullivan D. (2018) "Speculative and Constructively Critical GIS", in *The Canadian Geographer/Le Géographe canadien*, 62(1):4-6.
- Theil H. (1971) *Principles of Econometrics*, New York, Wiley.
- Tobler W. (1970) "A Computer Movie Simulating Urban Growth in the Detroit Region", in *Economic Geography*, 46(sup1):234-240.
- Tolomeo C., 1574. *Geografia*. Venezia, Giordano Ziletti (Traduzione di G. Ruscelli) (<http://websrv.archeo.unict.it:8080/items/show/203?c=0&m=0&s=0&cv=88&xywh=109%2C20%2C654%2C568>)
- Tomlinson R. (1984) "Geographic Information Systems: A New Frontier", in *The Operational Geographer*, 5:31-35.
- Tomlinson R. (1988) "The Impact of the Transition from Analogue to Digital Cartographic Representation", in *The American Cartographer*, 15(3):249-262.
- Turco A. 1984. *Regione e regionalizzazione*, Milano, FrancoAngeli.
- Unwin D. (1975) *An Introduction to Trend Surface Analysis*, CATMOG (Concepts and Techniques in Modern Geography) n. 5, Norwich, Geo Abstracts, University of East Anglia.
- Unwin D. (1986) *Analisi Spaziale*, Milano, FrancoAngeli.
- Unwin D. (1996) "GIS, Spatial Analysis and Spatial Statistics", in *Progress in Human Geography*, 20(4):540-551.
- Upton G., Fingleton B. (1985) *Spatial Statistics by Example*, Chichester, Wiley.
- Vagaggini V., Dematteis G. (1976) *I metodi analitici della geografia*, Firenze, La Nuova Italia.
- Vallega A. (1984) "Dalla regione alla regionalizzazione: avanzamento teorico e nodi concettuali", in Turco A., ed., *Regione e regionalizzazione*, Milano, FrancoAngeli.
- van Meeteren M. (2019) "Statistics Do Sweat. Situated Messiness and Spatial Science", in *Transactions of the Institute of British Geographers*, 44(3):454-457.
- van Meeteren M., Poorthuis A. (2018) "Christaller and "Big Data". Recalibrating Central Place Theory via the Geoweb", in *Urban Geography*, 39(1):122-148.
- Wainwright J., Bryan J. (2009) "Cartography, Territory, Property: Postcolonial Reflections on Indigenous Counter-Mapping in Nicaragua and Belize", in *Cultural Geographies*, 16(2):153-178.
- Walden A., Guttorp P. (1992) *Statistics in the Environmental and Earth Sciences*, London, Hodder Education.
- Wan, Y., Shi W., Gao L., Chen P., Hua Y. (2015) "A General Framework for Spatial Data Inspection and Assessment" in *Earth Science Informatics*, 8(4): 919-935.
- Warren S. (2020) "Introduction to the Special Section: Teaching critical GIS/Teaching GIS critically", in *The Canadian Geographer/Le Géographe canadien*, 64(4):467-470.

- Waters N. (2017) "GIS: History", in *International Encyclopedia of Geography. People, the Earth, Environment and Technology*, New York, Wiley, 1-6. DOI: 10.1002/9781118786352
- Westerholt R., Mocnik F.-B. Comber A. (2020) "A place for Place. Modelling and Analysing Platial Representations", in *Transactions in GIS*, 24(4):811-818.
- Wilson M. (2015) "New Lines? Enacting a Social History of GIS", in *The Canadian Geographer/Le Géographe canadien*, 59(1):29-34.
- Wilson M. (2017) *New Lines. Critical GIS and the Trouble of the Map*, Minneapolis, University of Minnesota Press.
- Wolf L., Murray A. (2016) "Spatial Analysis", in *International Encyclopedia of Geography. People, the Earth, Environment and Technology*. New York, Wiley, 1-6. DOI: 10.1002/9781118786352
- Wright D., Goodchild M., Proctor J. (1997) "Demystifying the Persistent Ambiguity of GIS as 'Tool' versus 'Science'", in *Annals of the American Association of Geographers*, 87(2):346-362.
- Wright D., Wood R., Sylvander, B. (1998) "ArcGMT: a Suite of Tools for Conversion between Arc/INFO® and Generic Mapping Tools (GMT)" in *Computers and Geosciences*, 24(8):737-744.
- Wyly E. (2009) "Strategic Positivism", in *The Professional Geographer*, 61(3):310-322.
- Wyly E. (2014) "The New Quantitative Revolution", in *Dialogues in Human geography*, 4(1):26-38.
- Zanetto G. (1979) "Il potenziale: da modello a strumento", in *Rivista geografica italiana*, 86(3):298-320.
- Zanetto G. (1982) "Analisi dei sistemi in geografia regionale", in *Rivista geografica italiana*, 89(2):362-372.
- Zanetto G. (1986) "Teoria della regionalizzazione e verifica quantitativa", in *Regione e regionalizzazione* in Turco A., ed., *Regione e regionalizzazione*, Milano, FrancoAngeli.
- Zanetto G., Lando F. (1980) "Mestre: analisi tipologica di una struttura urbana", in *Bollettino della Società Geografica Italiana*, 117:213-251.
- Zhang Z., Manjourides J., Cohen T., Hu Y., Jiang Q. (2016) "Spatial Measurement Errors in the Field of Spatial Epidemiology", in *International Journal of Health Geographics* 15:21.
- Zhu A.X., Lu G., Liu J., Qin C.Z., Zhou C. (2018) "Spatial Prediction Based on Third Law of Geography", in *Annals of GIS*, 24(4):225-240.
- Zucchini F. (2013) *Repubblica dei veti (La): Un'analisi spaziale del mutamento legislativo in Italia*, Milano, EGEA.

Ringraziamenti

To write a book is like making a Pinocchio. It runs away on its own.
Hägerstrand, *Progress in Human Geography*, 16(4), p. 543

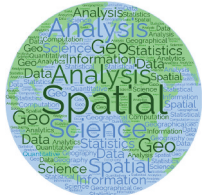
Prima d'ora, non avevo mai scritto un libro, e non pensavo neppure di volerlo fare. Anzi, ripensandoci ora, credo fosse il libro che voleva venire al mondo, uscendo dalla materia grezza, e per questo è stato lui a guidare la mia mano che impugnava lo scalpello. È stato un po' come liberare da un ciocco di legno un burattino capriccioso che appena gli fai la bocca ti chiede di fargli le gambe per correre dove vuole lui.

A questo mio burattino riconosco tenacia e pazienza. Scriverlo è stato un viaggio lungo, interrotto sia da impegni prevedibili, come la didattica, che con cadenza periodica s'impone sopra ogni altro progetto, sia da eventi imprevisti, come la pandemia, che ci ha sconvolto piani, priorità, stato d'animo. Ma anche durante le interruzioni il burattino non smette il suo lavoro, e quando ci rimetti mano ti stupisci di quanto sia cresciuto e maturato.

Questo libro *parla* di quello che da tanti anni *faccio*. E tra parlare e fare, si sa, c'è di mezzo il mare. Uno straordinario mare che ho attraversato leggendo, perché scrivere un libro come questo significa soprattutto leggere, leggere anche ciò che pensavi di sapere e imparare anche ciò che credevi di conoscere. Leggere è un viaggio, imparare è un privilegio, scrivere è condividere, e questi mesi sono stati anche una splendida avventura.

In questo viaggio non sono mai stata sola, e voglio esprimere la mia gratitudine a quanti mi sono stati vicini e a quanti mi hanno guidato inconsapevolmente, come fanno le stelle coi marinai. Ringrazio l'Editore e i Revisori, le cui critiche costruttive mi hanno aiutato a migliorare il manoscritto. Sono riconoscente ai colleghi di Firenze per il sostegno nelle fasi iniziali, a quelli di Calgary per la pazienza e

generosità durante questo mio impegno, e a tutti i colleghi, studenti e amici per tutte le discussioni da cui sono emersi spunti di riflessione. Ringrazio soprattutto i colleghi che hanno collaborato agli studi presentati nel capitolo 7, e in modo particolare Mojgan Mirzaei ed Alison Hanes per i loro contributi alla realizzazione delle Figure 4.1, 7.7 e 7.8. Alla preziosa Isabelle Couloigner va molto più che la mia riconoscenza per aver rivisto le formule del capitolo 2, ma anche per aver sostenuto altri progetti durante questi mesi. Non riesco ad esprimere tutta la mia gratitudine all'ineguagliabile Fabio Lando per la paziente lettura e rilettura del manoscritto, per i commenti saggi, discreti, incoraggianti e ironici e, soprattutto, per aver creduto in me. Infine, grazie a tutta la mia famiglia, che è stata sempre con me sostenendomi in tanti modi, dalla fornitura di “servizi bibliotecari” alla sopportazione del mio impegno e delle mie distrazioni: so che c'eravate sia quando vi informavate su questo mio lavoro, sia, soprattutto, quando non osavate chiedermene; e al mio Lorenzo, con cui condivido tante letture, per aver saputo ridere con me tutte le volte che mi chiedevi “come va col libro?”, e io ti parlavo di quello che stavo scrivendo, anziché di quello che stavo leggendo con te.



L'ANALISI SPAZIALE

La Geografia che... conta

In questo inizio di secolo siamo circondati da fenomeni complessi, che si manifestano con intensità variabile e caratteristiche locali nello spazio geografico, interagendo con altri fenomeni su scale diverse: il cambiamento climatico, le migrazioni di massa, l'inquinamento, la produzione alimentare, le epidemie. Nel 1854 il medico inglese John Snow disponeva solo di una penna e di uno stradario su cui segnare i decessi da colera, eppure con quella rudimentale analisi spaziale contribuì ad individuare l'origine del focolaio epidemico, arginando il contagio e fornendo un apporto straordinario alla medicina. Oggi disponiamo di tecnologia e di un diluvio di dati, ma soprattutto di teoria, concetti e strumenti per condurre analisi spaziali affidabili, capaci di contribuire alla comprensione dei fenomeni e alla gestione del territorio.

Cresciuta oltreoceano, l'analisi spaziale, che fu e rimane la ragion d'essere del GIS, è una parte qualificante della moderna geografia quantitativa e continua a stimolare sviluppi della statistica che affrontino gli effetti dei processi e dai dati spaziali, perché *spaziale* è *speciale*. Conosciuta e praticata da pochi, sporadicamente presente nella nostra letteratura, l'analisi spaziale è una disciplina giovane eppure matura, sfuggente, forse elitaria per non aver ceduto a tentazioni populiste e commerciali, ma capace di teoria e di critica oltre che di statistica e di matematica, di metodi rigorosi che si spingono ben oltre paradigmi riduttivamente quantitativi, pronta al dialogo con le scienze sociali, mediche, ambientali.

Questo è dunque lo scopo di questo libro: presentare, discutere, rendere più accessibile questa disciplina ancora troppo poco nota, che anche nell'epoca dei *big data* e della tecnologia pervasiva non persegue la popolarità con *analitiche* redditizie e grafiche futuristiche, accettando invece l'errore, i limiti dei dati, cui antepone teoria, rigore e verifica di ipotesi. Per questo, l'analisi spaziale potrebbe, oggi, non solo contribuire alla gestione del territorio, ma anche riaffermare l'utilità e l'importanza, se non la necessità, della geografia.

Stefania Bertazzon, professore ordinario di geografia presso l'Università di Calgary (Canada), insegna analisi spaziale nei corsi post-laurea. I suoi studi producono modelli analitici spaziali affidabili per sanità pubblica e gestione ambientale, anche in collaborazione con networks e agenzie governative. I suoi lavori sono pubblicati in riviste internazionali, non solo di geografia.