

L'impiego di algoritmi AMOEBA per lo studio delle variazioni temporali di un fenomeno economico: prime evidenze generalizzabili da un caso di studio*

The use of AMOEBA algorithms for studying time-variations of an economic phenomenon: first generalizable evidences from a case study

GIAN PIETRO ZACCOMER, LUCA GRASSETTI

Università degli studi di Udine, gianpietro.zaccomer@uniud.it

Riassunto

Il presente lavoro utilizza la spesa sanitaria come pretesto per illustrare come analizzare, su scala comunale, non il valore assoluto ma la variazione temporale di una qualsiasi voce di un bilancio regionale come, ad esempio, quella delle spese sostenute per il restauro, la manutenzione e la conservazione dei beni culturali. Infatti, ogni fenomeno economico diffuso sul territorio risente in modo pesante, anche nelle sue variazioni, delle relazioni esistenti tra unità territoriali considerate "vicine". In questo intento l'aspetto più importante, dal punto di vista dell'analisi geografica, è superare definitivamente la tradizionale tendenza della teoria economica classica a trascurare l'influenza del territorio sui fenomeni economici: è quindi necessario predisporre nuovi strumenti analitici capaci di sfruttare l'informazione sulla struttura territoriale al fine di spiegare in modo più realistico le variazioni temporali di tali fenomeni. Uno dei metodi tradizionalmente utilizzati per analizzare la variazione temporale di un fenomeno è quello della *Shift-Share Analysis* che, per più di un mezzo secolo, è stata considerata dalla letteratura, statistica ed economica, uno strumento territoriale finché la ricerca geografica, in un'ottica sistemica, ha puntualizzato l'idea che non basta considerare dati riferiti ad unità amministrative, o celle diversamente costruite, per parlare di "modello territoriale", ma è invece necessario tenere conto anche delle relazioni che si vengono a formare tra di esse. Allo stato attuale, la struttura relazionale viene ipotizzata a priori dallo stesso ricercatore. In questo contributo si vuole mostrare come sia possibile rimuovere tale ipotesi, rendendo il modello sottostante più generale

Abstract

Every economic phenomenon presenting a spatial diffusion is strongly affected, even in its variations, by the relationships existing among neighbouring territorial units. The present work represents an example of the analyses of the changes in the different kinds of expenditure in the regional budget (as, for instance, the restoration, maintenance, and preservation of cultural heritage). In particular, the analysis focuses on the healthcare expenditures observed at the municipality level. This work aims at overcoming the traditional tendency of classical economic theory to neglect the influence of the territory on economic phenomena. In particular, new analytical tools capable of exploiting information on the territorial structure must be used in the practice in order to explain the temporal variations of these phenomena. A classical approach to the analysis of the temporal variations of an economic phenomenon is the so-called Shift-Share Analysis. For over half a century, this methodology has been considered by the statistical and the economic literatures as a tool for territorial analysis. The geographic literature pointed out that also the relationships among the territorial units must be considered in order to define a "territorial model". Typically, the relational structure (neighbouring system) is assumed ex-ante by the researchers. The present contribution aims to show how this hypothesis can be removed. In particular, the use of a modified version of the AMOEBA algorithm (A Multidirectional Optimum Ecotope-Based Algorithm) in the Spatial Shift-Share Analysis is proposed to account for the spatial structure in accordance with the well-known Tobler's "First Law of Geography".

* Il lavoro è frutto dell'impegno comune degli autori, ma la stesura finale va attribuita per il riassunto e il paragrafo 1 a L. Grassetto, mentre le restanti parti del lavoro a G.P. Zaccomer.

di quanto già proposto in letteratura, individuando la struttura territoriale direttamente in sede di analisi attraverso un originale adattamento dell'algoritmo AMOEBA (*A Multidirectional Optimum Ecotope-Based Algorithm*) che assume come guida la ben nota "Prima legge della geografia" di Tobler.

Parole chiave

Analisi di scomposizione spaziale, matrice geografica, pesi spaziali, Friuli Venezia Giulia, spesa sanitaria

Keywords

Spatial Shift-Share Analysis, geographical matrix, spatial weights, Friuli Venezia Giulia, healthcare expenditure

1. Il quadro teorico di riferimento: l'evoluzione del ruolo del territorio

Il problema dell'analisi di una variazione temporale di un qualsiasi fenomeno economico è un argomento ben noto nella letteratura econometrica, a partire da quella sulle serie storiche. Il problema nasce quando si vuole analizzare fenomeni che si manifestano su un territorio. Considerando la sua natura sistemica, il territorio influisce pesantemente sulla concretizzazione dei fenomeni economici, attraverso l'insieme delle sue relazioni orizzontali e verticali, quindi anche sulle loro variazioni temporali. Sin dalla metà del secolo scorso è stato puntualizzato come sia fondamentale considerare nell'analisi di un fenomeno economico le relazioni che si vengono a formare all'interno del territorio interessato. Questa idea centrale della geografia moderna non solo ha prodotto un proprio *corpus* di modelli, ben noto nell'ambito della manualistica di settore (cfr. Bencardino e Prezioso, 2006; Conti, 2012), ma lentamente ha anche contagiato le scienze affini, quali quelle economiche e statistiche.

Il caso della *Shift-Share Analysis* (SSA) è un valido esempio di quanto appena detto poiché, partendo da studi demografici legati all'industrializzazione inglese degli anni Quaranta, solo agli inizi del nuovo millennio assume una connotazione realmente territoriale

diventando *Spatial Shift-Share Analysis* (SSSA), ossia una metodologia che permette finalmente di inserire nel modello le relazioni territoriali come ipotesi del modello stesso. L'evoluzione storica dalla SSA alla SSSA, e il suo attuale quadro di riferimento, sono oggi facilmente delineabili grazie al minuzioso capitolo di ricostruzione della letteratura internazionale ad opera di Haynes e Parajuli presente nell'*Handbook of research methods and applications in economic geography* del 2015. Tale lavoro, partendo dall'embrionale idea proposta da Jones nel *Report of the Royal Commission on the Distribution of the Industrial Population* del 1940 e passando per la "versione classica" del modello di Dunn del 1960, esplicita la "rivoluzionaria" *spatial extension* proposta da Nazara e Hewings nel 2004, che pone fine alla ben nota tendenza di ignorare le relazioni territoriali sottostanti al fenomeno indagato tipica della teoria economica precedente. Il quadro di riferimento proposto dagli Autori termina con il lavoro di Zaccomer e Mason del 2011. In altre parole, il lavoro di Haynes e Parajuli (2015) riorganizza e sintetizza circa 70 anni di letteratura specifica su questo tipo di analisi.

Rimandando a questa ampia rassegna per ogni approfondimento teorico di base, sembra comunque necessario ricordare che la SSSA viene considerata nella letteratura, anche geografica, come una metodologia capace di scomporre la variazione temporale di un qual-

siasi fenomeno misurabile in componenti elementari, interpretabili dal ricercatore, tenendo conto delle relazioni esistenti tra le unità territoriali coinvolte nello studio. Se l'impostazione classica di Dunn non considerava per niente le relazioni territoriali, postulando indirettamente una ugual interazione tra unità territoriali, Nazara e Hewings (2004) propongono l'idea di introdurre una *matrice geografica* all'interno del modello classico attraverso l'utilizzo di *tassi (di variazione) di vicinato* che misurano la variazione temporale del solo vicinato. Per *vicinato* di una unità territoriale si intende il suo "insieme di unità vicine" che possono essere semplicemente contigue dal punto di vista fisico (se confinano sono "vicine", altrimenti no) oppure connesse in altri modi più complessi. Per il prosieguo di questo lavoro, è fondamentale sottolineare che la matrice geografica incorpora l'ipotesi fatta dal ricercatore sulla struttura dei vicinati non solo mostrando qual è l'insieme di unità "vicine", ma quantificando pure l'entità delle relazioni attraverso i suoi elementi detti anche *pesi spaziali*¹.

L'idea originale di Nazara e Hewings (2004), solo abbozzata dal punto di vista formale, viene successivamente approfondita in Zacommer (2006, 2008), dove vengono proposti alcuni modelli originali, e in Zacommer e Mason (2008) dove si introduce in letteratura un primo studio comparativo tra diverse matrici geografiche, e quindi sugli effetti delle ipotesi poste a priori dal ricercatore sulle componenti del modello *Shift-Share*. Infine, in Zacommer e Mason (2011) vengono studiate, in modo approfondito, le caratteristiche formali dei modelli proposti in precedenza fornendo una prima base formale all'idea proposta da Nazara e Hewings su cui

1 Cfr. Unwin (1986) per una sua prima introduzione geocartografica a quelle che vengono chiamate *matrici di adiacenza*, mentre per una più approfondita trattazione nell'ambito del concetto di autocorrelazione spaziale, qui successivamente utilizzato, si rimanda a Zani (1993); infine, cfr. Zacommer e Mason (2008) per un'ampia rassegna sulle diverse possibilità per costruire tale matrice. Si ricorda che per essere utilizzata all'interno della SSSA, la matrice geografica deve essere di tipo standardizzato per riga, così come richiesto da Nazara e Hewings (2014), ossia i suoi elementi sono pesi spaziali (che si possono leggere anche in termini percentuali) la cui sommatoria di riga è unitaria. Per fare un esempio, se un vicinato è composto da due sole unità con pesi 0,3 e 0,7, questo implica che la prima unità influirà per il 30% e la seconda del 70% sulla manifestazione, misurata presso l'unità territoriale a cui è riferito il vicinato, del fenomeno economico oggetto di studio.

la ricerca internazionale, in particolare quella spagnola, risulta particolarmente attiva.

Dopo il 2011 l'ulteriore avanzamento metodologico, di natura squisitamente geografica, si concentra nel fatto di non lasciare più al ricercatore il compito di selezionare, prima dell'analisi e in modo discrezionale, la matrice geografica ma piuttosto di cercare di perseguire il più possibile la *Prima legge della geografia* senza perdere quella base teorica appena conquistata. Tale principio enuncia che "*everything is related to everything else, but near things are more related than distant things*" (cfr. Tobler, 1970, p. 236): lo strumento utilizzato correntemente in letteratura per misurare la sua veridicità empirica è la *correlazione spaziale* (o *autocorrelazione spaziale* se riferita ad un unico fenomeno come il caso qui trattato; cfr. Zani, 1993, pp. 165-177), ossia una correlazione misurata proprio tenendo conto delle relazioni territoriali formalizzate attraverso la matrice geografica.

Come evoluzione dell'ultimo lavoro citato da Haynes e Parajuli (2015), Zacommer e Grassetto (2014) proprio sulle pagine di questo Bollettino introducono in via sperimentale l'algoritmo AMOEBA (*A Multidirectional Optimum Ecotope-Based Algorithm*) proposto da Aldstadt e Getis (2004, 2006), attraverso una sua variante originale, esplicitando il ruolo fondamentale dell'analisi cartografica come passo intermedio della SSSA. In altre parole, la metodologia proposta diventa ancora "più geografica" rispetto a quanto proposto da Nazara e Hewings (2004) poiché non solo è necessario introdurre la matrice geografica nel modello per tenere conto delle relazioni territoriali, ma non è più necessario ipotizzare a priori la struttura delle relazioni territoriali: la distribuzione spaziale di ogni vicinato prodotta dall'algoritmo è un risultato dell'analisi stessa. Essa può quindi essere studiata graficamente al fine di comprendere meglio il processo di scomposizione nelle componenti elementari che il modello permette di definire.

Il modello di SSSA, che sarà utilizzato per la scomposizione del tasso di variazione della spesa sanitaria regionale del Friuli Venezia Giulia (FVG), racchiude in sé il senso etimologico della parola "analisi" nella sua accezione di "scomposizione di un tutto nelle parti che lo costituiscono". Esso presenta tre livelli territoriali (comunale, di vicinato e regionale) e viene così formalizzato (cfr. Zacommer, 2008):

$$g_{r..} = g_{...} + (\tilde{g}_{r..} - g_{...}) + \left[\sum_{i=1}^I (\tilde{g}_{ri} - \tilde{g}_{r..}) \frac{a_{ri}}{a_{r..}} + \sum_{f=1}^F (\tilde{g}_{r.f} - \tilde{g}_{r..}) \frac{a_{r.f}}{a_{r..}} + C_r \right] + \sum_{i=1}^I \sum_{f=1}^F (g_{rif} - \tilde{g}_{rif}) \frac{a_{rif}}{a_{r..}} \quad (1)$$

= TEND + (CFR) + [INTRA] + LOC

dove il contatore r è relativo all'unità territoriale di riferimento (in questo caso un comune del FVG), mentre i contatori i e f sono relativi alle due variabili che fungono da *chiavi di scomposizione* (come le classi di età e il genere del singolo cittadino che effettua la spesa). I tassi di vicinato, tutti segnalati nella (1) come \tilde{g} , tengono conto della matrice geografica quindi sono tassi territoriali riferiti solo alle unità "vicine" dell'unità territoriale di riferimento. Considerando ora le singole componenti della (1), si definisce: a) la *componente tendenziale* TEND il solo tasso di variazione complessivo regionale $g_{...}$ del FVG che funge da *baseline*, ossia da punto di riferimento per la scomposizione; b) la *componente di confronto* CFR, compresa nella prima parentesi tonda, raffronta il tasso del vicinato con quello della regione permettendo di capire se il fenomeno considerato varia nel vicinato in modo simile, o dissimile, dalla *baseline* regionale; c) la *componente di confronto intra-vicinato* INTRA esprime gli effetti complessivi di vicinato delle due chiavi di scomposizione ed è ottenuta come somma degli elementi all'interno della parentesi quadra (non essenziali per questo lavoro, per una loro definizione si rimanda a Zaccomer e Mason, 2011) ed esprime, nel suo complesso, la variazione del fenomeno dentro il vicinato; d) la *componente locale* LOC, ultimo termine della scomposizione, raccoglie tutti i contributi locali, in questo caso comunali, diversi dai precedenti. In estrema sintesi, una variazione (di un fenomeno economico) comunale viene considerata come influenzata dal fatto di trovarsi in una regione ben precisa, di avere un proprio vicinato che si comporta in un determinato modo e di un insieme di fattori comunali diversi da quelli modellizzati attraverso le due chiavi di scomposizione. Finalmente un'interpretazione territoriale di tale variazione.

Per quanto riguarda la costruzione della matrice geografica, in questo lavoro saranno utilizzare le classiche proposte legate alla semplice contiguità territoriale di primo ordine (del tipo binario "confina o non confina") e a quella di Cliff e Ord (1981), che tiene conto non solo

della contiguità territoriale ma anche della distanza fisica tra baricentri e della porzione di perimetro in comune tra unità territoriali (informazioni che possono essere ricavate solo disponendo di una cartografia regionale dettagliata su base GIS) che serviranno successivamente per il confronto con quelle ottenute attraverso l'algoritmo AMOEBA originale e sue declinazioni.

Le caratteristiche dell'algoritmo AMOEBA sono già state ampiamente trattate in Zaccomer e Grassetto (2014), ma si ricorda che tale algoritmo si basa su una logica di ottimizzazione di un indicatore di autocorrelazione spaziale locale, in particolare su quello proposto da Getis e Ord (1992, 1995). Come anticipato, con questo algoritmo la matrice geografica non viene più ipotizzata a priori dal ricercatore, ma viene prodotta dall'algoritmo ricercando, passo per passo (prima si considerano i vicini diretti, ossia quelli di primo ordine, poi i vicini dei vicini, ossia quelli di secondo ordine, e così via) quale configurazione spaziale del vicinato massimizza l'autocorrelazione. In altre parole, il vicinato non è più semplicemente fissato a priori, ma è costituito da quell'insieme di unità territoriali che presentano, in questo caso, variazioni di spesa simili. Questo permette di affermare che il vicinato, grazie a questo algoritmo, viene costruito ricercando l'omogeneità di comportamento, in linea quindi con la formulazione generale della legge di Tobler dove si postula che unità territoriali "vicine" dovrebbero avere manifestazioni altamente correlate tra di loro: l'algoritmo AMOEBA va proprio alla ricerca di quella configurazione spaziale del vicinato più omogenea nel senso che presenta variazioni temporali simili.

Nell'algoritmo AMOEBA proposto da Getis e Ord rimangono individuati due distinti problemi: il primo è legato ai suoi elevati tempi di esecuzione, che rendono molto onerosa l'applicazione della SSSA nel caso di un elevato numero di unità territoriali; il secondo riguarda un aspetto tecnico sulla costruzione della matrice geografica che vanifica, di fatto, l'apporto delle unità

territoriali individuate all'ultimo passo di aggregazione dell'algoritmo. In Zacommer e Grassetto (2014, p. 61) viene proposto un algoritmo che, recependo il richiamo di Griffith (1996) al *principio di parsimonia* anche nella formalizzazione delle relazioni territoriali, è stato denominato AMOEBA *parsimonioso* (AMOEBAp) che risolve il problema dell'efficienza computazionale. Una sua prima evoluzione, definito AMOEBAp *non zero*, permette di risolvere anche il secondo problema citato.

In questo lavoro, accanto a tali proposte già note in letteratura, si vogliono proporre due nuove declinazioni dell'AMOEBa. La prima è un ulteriore affinamento che risolve il mancato rispetto dell'ipotesi sottostante alla distribuzione dell'indicatore (Ord e Getis, 1995, p. 296) nel caso di vicinati non numerosi: la declinazione così ottenuta sarà qui indicata come AMOEBAp *non zero std* che, allo stato attuale della ricerca, rappresenta la versione più sofisticata dell'algoritmo originale. La seconda proposta vuole fornire più un punto di riferimento al fine dei successivi confronti, che un affinamento dello strumento, poiché evita *in toto* il problema della diversa definizione dei pesi applicando la stessa logica utilizzata per la costruzione delle matrici di contiguità, ossia si attribuiscono pesi tutti uguali, quindi non decrescenti, indipendentemente dall'ordine territoriale delle unità del vicinato: l'algoritmo in questo caso sarà indicato come AMOEBAp *pesi uguali*. Si avranno quindi a disposizione quattro varianti dell'algoritmo per cercare di trarre conclusioni generali sulla loro applicabilità futura ad ogni caso concreto, ossia ad ogni voce di un bilancio regionale.

2. L'analisi del comportamento degli algoritmi AMOEBA nel caso di studio

L'analisi della variazione della spesa sanitaria complessiva è stata condotta sui dati della Regione Autonoma Friuli Venezia Giulia (FVG) prodotti dal *Sistema Informativo SocioSanitario Regionale* (SISSR) e, successivamente, trattati attraverso il linguaggio R (R Core Team, 2013) ed alcune sue librerie. Sembra fondamentale sottolineare che i dati qui utilizzati sono relativi a singoli cittadini, quindi soggetti alle restrittive norme sulla *privacy*, ma il modello di *Shift-Share* permette proprio di analizzare dati di questo tipo per restituire informa-

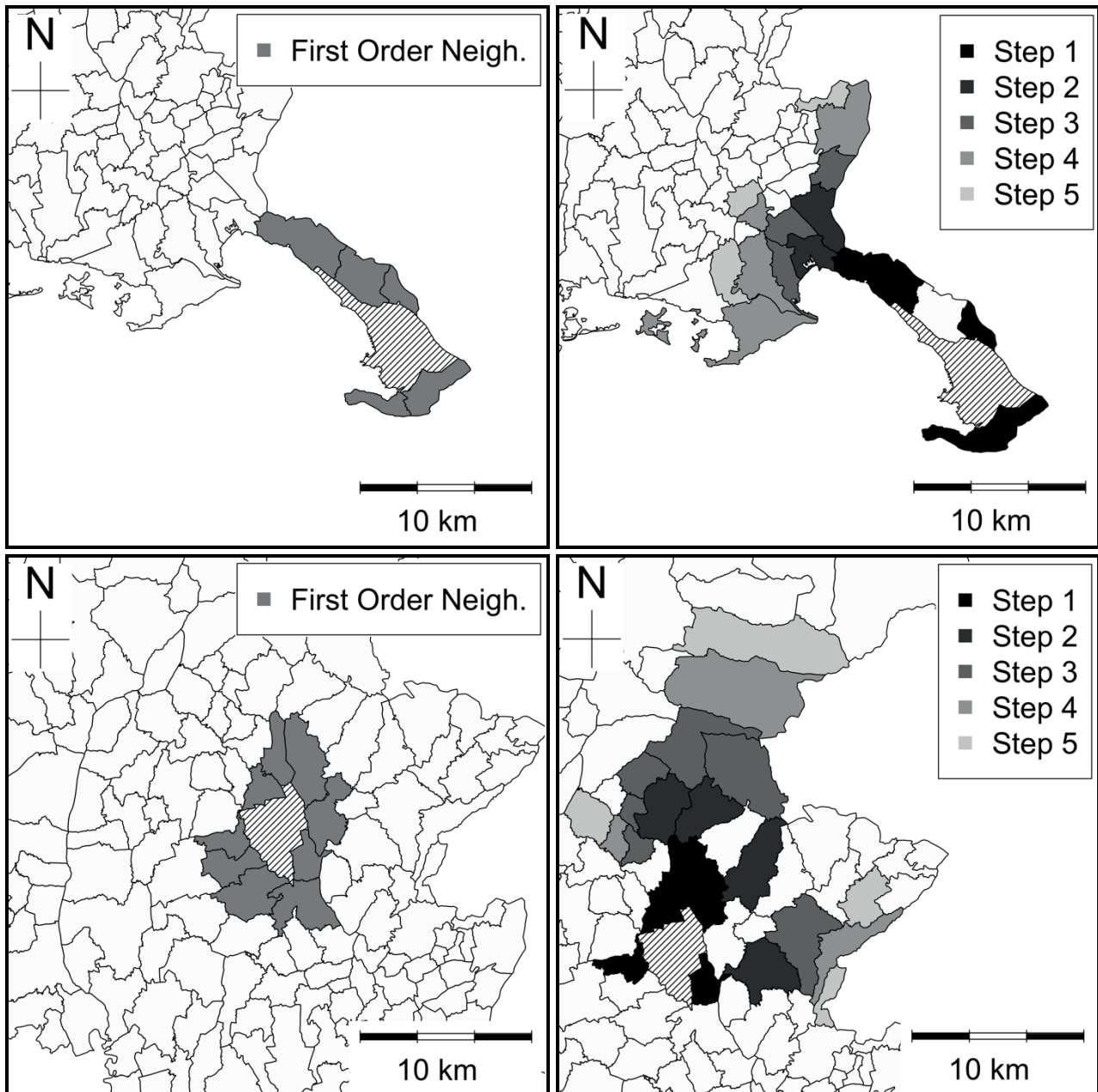
zioni aggregate a livello comunale e di vicinato. Tali dati sono già stati studiati, da diversi punti di vista, in Grassetto e Rizzi (2014) e Piergentili *et al.* (2015) per cui il presente contributo si concentra unicamente sullo studio della loro variazione temporale: l'amministrazione regionale ha fornito i dati relativi agli anni 2010 e 2012 per cui i tassi di variazione qui utilizzati sono tutti di tipo biennale. Si sottolinea poi che il fenomeno economico oggetto di studio è la variazione della spesa sanitaria totale a livello pro-capite, passaggio necessario, poiché le strutture demografiche dei comuni coinvolti sono ben diverse, e che ha richiesto l'informazione sulla popolazione residente di fonte ISTAT per gli anni considerati.

Per quanto riguarda la struttura dei dati, per ogni comune si conosce non solo la spesa sanitaria complessiva, ma anche la sua suddivisione per: *genere* (maschio o femmina); *classi di età* (da 0 a 16, da 17 a 30, da 31 a 50, da 51 a 70, oltre i 70 anni compiuti); *tipologia della spesa sanitaria* (classificata in farmaceutica, ambulatoriale e ospedaliera). Un simile dettaglio conoscitivo permette di avere a disposizione ben tre chiavi di scomposizione da selezionare per applicare la (1). Sulla base di questi dati, per ogni singolo comune del FVG e per ogni coppia di chiavi di scomposizione, sono state portate a termine sei analisi per ogni comune (ossia circa 1300 scomposizioni) relative a due ipotesi territoriali classiche e quattro diverse versioni dell'algoritmo AMOEBA.

In questo lavoro saranno illustrati dapprima l'aspetto cartografico, passo preliminare legato alla costruzione dei vicinati e, successivamente, gli aspetti più quantitativi ottenuti dal modello (1). A titolo meramente esemplificativo, si concentrerà l'attenzione prima di tutto sulla scomposizione del tasso di variazione biennale, che utilizza come chiavi di scomposizione l'età del cittadino e il tipo di spesa che esso effettua, per quanto riguarda le due realtà comunali più rilevanti, ossia Udine e Trieste, ma verranno comunque proposti risultati generali derivati dall'osservazione di tutte le scomposizioni relative ai 217 comuni del FVG, segnalando in seguito esempi più problematici.

L'analisi empirica mostra come per i capoluoghi presentati gli algoritmi AMOEBA, originale e parsimonioso, portino alla stessa configurazione del vicinato così come mostrato in Figura 1, dove i capoluoghi sono evi-

FIGURA 1 – Carta dei vicinati dei comuni capoluogo di Trieste (in alto) e Udine (in basso). A sinistra sono riportati i vicinati basati sulla contiguità di primo ordine (*First Order Neigh.*), a destra le soluzioni convergenti per gli algoritmi AMOEBA e AMOEBAp. I comuni di Trieste e Udine sono rappresentati dal poligono tratteggiato. Se a sinistra il loro vicinato di primo ordine, costituito dai soli comuni direttamente confinanti, è segnalato con un'unica tonalità di grigio, a destra il vicinato è riportato in diverse tonalità poiché la scala cromatica (i comuni direttamente confinanti in nero, i confinanti dei confinanti in grigio scuro, e così via, a scalare verso tonalità di grigio sempre più chiare) corrisponde non solo ai diversi passi (*step*) dell'algoritmo, ma il suo massimo valore esprime l'ordine territoriale del vicinato



FONTE: elaborazione propria su dati SISR-Regione Autonoma Friuli Venezia Giulia

denziati dai poligoni tratteggiati. Tale primo risultato è fondamentale poiché ottenuto sulle maggiori realtà urbane del FVG che non presentano dati mancanti o tassi anomali: questo conferma, dal punto di vista della configurazione spaziale del vicinato, la sostanziale aderenza delle declinazioni degli algoritmi parsimoniosi alla versione originale di Getis e Ord quando non ci siano grossi problemi sui dati.

Entrando nel dettaglio, il confronto cartografico tra vicinati di tipo contiguo (*First Order Neigh.* nella Figura 1) e di tipo AMOEBA mette in evidenza come, nel caso di Trieste, il vicinato di primo tipo restituisca sostanzialmente la sua provincia, mentre quello basato sull'autocorrelazione spaziale tende a sconfinare in quella di Gorizia. Nel caso di Udine il suo vicinato "circolare", basato sulla mera contiguità, nel caso AMOEBA si spezza privilegiando uno sviluppo soprattutto verso nord che giunge fino a Chiusaforte. Nella Figura 1, il vicinato ottenuto dagli algoritmi AMOEBA non viene indicato nello stesso modo, ma vengono evidenziati i diversi passi dell'algoritmo con una scala cromatica di diverse tonalità di grigio. Il numero massimo di tali passi corrisponde all'*ordine territoriale del vicinato*, ossia i vicini di primo ordine sono solo quelli direttamente confinanti, quelli di secondo ordine sono i confinanti dei confinanti e così via. Dalla Figura 1 si evince come, sia nel caso di Trieste sia di Udine, l'ordine territoriale massimo raggiunto è cinque o, per dirla con il linguaggio della *teoria dei grafi* (cfr. Unwin, 1986, p. 147-159), bisogna fare cinque passi per andare da un capoluogo al comune del suo vicinato più lontano.

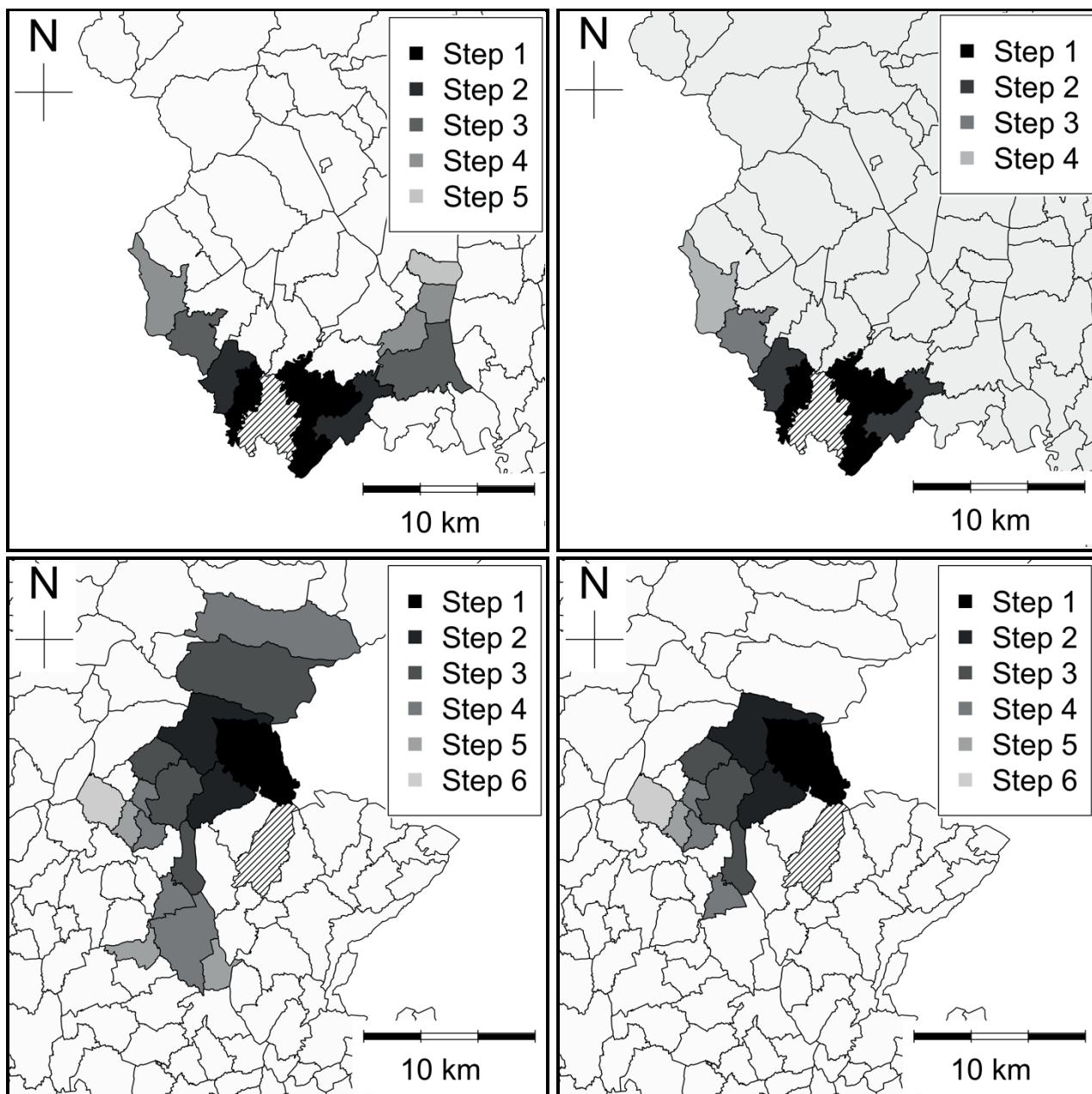
Per completare la descrizione dell'analisi cartografica per ogni vicinato di ciascun comune del FVG, bisogna concentrare ovviamente l'attenzione sui 20 comuni (meno del 10% di quelli considerati) per cui le due versioni dell'algoritmo AMOEBA, originale e parsimonioso, non convergono. In sintesi, la loro analisi ha messo in evidenza come si tratti generalmente di piccole realtà in termini assoluti, che portano a variazioni anomale nei tassi pro-capite. Per fornire un esempio concreto, il più chiaro possibile ai soli fini esplicativi, si propongono i casi di Pasiano di Pordenone (PN) e di Faedis (UD) i cui vicinati AMOEBA sono proposti in Figura 2. Nel caso di Pasiano il vicinato dall'algoritmo classico tende a svilupparsi rispetto ad un "arco orizzontale", con conca-

vità verso nord, più allargato rispetto a quello ottenuto dalla sua declinazione parsimoniosa che risulta troncato a est raggiungendo un ordine territoriale più basso (4 contro 5).

Il caso di Faedis è ancora più intrigante di quello di Pasiano perché permette di mettere in evidenza diversi aspetti: il vicinato individuato dall'algoritmo classico tende ad avvolgersi su se stesso soprattutto verso sud, mentre l'algoritmo parsimonioso contribuisce comunque a migliorare il rispetto della legge di Tobler poiché questo risulta più compattato su Faedis (sia a nord che a sud). Interessante notare come, nonostante quest'ultimo vicinato risulti geograficamente più limitato, l'ordine territoriale rimanga il medesimo ossia sei. Si ottiene quindi un'ulteriore evidenza empirica: non sempre vicinati più compatti portano ad ordini territoriali più bassi. L'esempio di Faedis risulta emblematico anche da un altro punto di vista, ossia dimostra il rischio che si corre a porre l'ipotesi a priori che il suo vicinato sia costituito dai comuni contigui del primo ordine. Infatti, questa si rivela un'ipotesi priva di fondamento empirico poiché, analizzando i dati relativi ai singoli comuni, Faedis è circondato da comuni che non si comportano nello stesso modo, dal punto di vista della variazione della spesa sanitaria pro-capite, tranne quello individuato al primo passo dell'algoritmo (indicato in nero nella Figura 1).

Questo dimostra come l'analisi cartografica dei vicinati risulti un elemento fondamentale della nuova proposta di SSSA, infatti ora la configurazione del vicinato è un prodotto stesso dell'analisi e non più un'ipotesi del ricercatore: solo attraverso l'analisi della configurazione spaziale del vicinato è poi possibile passare all'analisi dei risultati numerici. Ovviamente si potrebbe obiettare che un simile modo di procedere allunghi i tempi richiesti dall'analisi poiché, in questo caso, per ogni analisi regionale bisogna considerare più di duecento immagini cartografiche: questo è vero, ma solo dopo è possibile interpretare i risultati numerici in modo critico, individuando subito dal punto di vista grafico i casi che risultano problematici. Prima di questa proposta il vicinato era sostanzialmente ipotizzato nella SSSA (se non addirittura escluso dalla SSA) senza porre alcun accento sull'analisi cartografica, quindi interpretando direttamente i risultati del modello in modo acritico, almeno dal punto di vista geografico.

FIGURA 2 - Carta dei vicinati dei comuni di Pasiano di Pordenone (in alto) e Faedis (in basso) con soluzioni divergenti per gli algoritmi AMOEBA (a sinistra) e AMOEBAp (a destra). I comuni di Pasiano e Faedis sono rappresentati dal poligono tratteggiato, mentre il vicinato degli stessi in tonalità di grigio poiché la scala cromatica corrisponde (i comuni direttamente confinanti in nero, i confinanti dei confinanti in grigio scuro, e così via, a scalare verso tonalità di grigio sempre più chiare) corrisponde non solo ai diversi passi (step) dell'algoritmo, ma il suo massimo valore esprime l'ordine territoriale del vicinato



FONTE: elaborazione propria su dati SISSR-Regione Autonoma Friuli Venezia Giulia

È possibile infine entrare nel dettaglio dell'analisi numerica, ossia ricongiungersi al percorso di analisi precedente che ignorava l'analisi cartografica dei vicinati, partendo dalla constatazione che in FVG, nel biennio considerato, la spesa sanitaria pro-capite è aumentata di poco più del 21%. Questo costituisce la *baseline*, ossia il valore di riferimento, incorporata nella componente tendenziale, a cui la spesa comunale dovrà fare riferimento o, in altre parole, è quel valore di partenza a cui ogni comune deve rapportarsi per il fatto di appartenere al FVG. Il modello (1) serve proprio per capire come dal valore regionale del 21% si passi poi a valori diversi per ciascun comune e questo tenendo conto che ciascun comune ha "intorno" un proprio vicinato che, nel caso degli algoritmi AMOEBA, viene costruito sulla base dell'omogeneità di comportamento sempre in termini di variazione della spesa sanitaria pro-capite.

I dati delle scomposizioni relativi ai due capoluoghi provinciali sono riportati in Tabella 1. Dalla lettura

verticale dell'ultima colonna della tabella si può notare come i due comuni restituiscano due tassi di variazione ben diversi: se Trieste, con il 16,2%, presenta un tasso di variazione inferiore a quello regionale, Udine fa registrare un valore pari a 23,1%, ossia di due punti percentuali superiore a quello regionale. Nella loro scomposizione si ritrovano comunque diverse analogie. Osservando i dati della tabella per colonna è possibile notare che le componenti di confronto CFR, intravicinato complessiva INTRA e locale LOC portano ad un contributo sempre dello stesso segno, quello che varia è solo l'entità di tale contributo. Quindi, dal punto della spiegazione della scomposizione, il ragionamento evidenziato dalla SSSA è il medesimo: partendo dal valore regionale, le componenti di confronto e intravicinato sono sempre positive indicando il fatto che i vicinati dei due capoluoghi contribuiscono ad aumentare la variazione di spesa rispetto alla *baseline*, ma solo un coacervo di fattori comunali, diversi dalle due chiavi di scomposi-

TABELLA 1 – Risultati numerici delle scomposizioni ottenute dal modello (1) per sei tipologie di matrici geografiche. Le componenti necessarie ai fini della presente analisi sono: TEND (tendenziale pari al solo tasso di variazione g_r del Friuli Venezia Giulia); CFR (confronto); INTRA (intravicinato complessiva); LOC (locale); g_c (tasso di variazione comunale pari a TEND + CFR + INTRA + LOC). Ai fini della coerenza e del confronto con quanto proposto in Zaccomer e Grassetti (2014), si riporta anche il dettaglio numerico della componente INTRA, ossia della parte di scomposizione all'interno della parentesi quadrata della (1), inerente le chiavi di scomposizione ETÀ (intravicinato relativa all'età della popolazione in classi), SPE (intravicinato relativa alla tipologia di spesa sanitaria); C_r (intravicinato relativa alla connessione tra età della popolazione e tipologia di spesa) ricordando che $INTRA = ETÀ + SPE + C_r$.

Matrice geografica	TEND	CFR	ETÀ	SPE	C_r	INTRA	LOC	g_c
	Trieste (cod. ISTAT 32006)							
Contiguità torre I ord.	21,03	0,97	0,44	0,28	0,57	1,29	-7,09	16,21
Cliff e Ord		1,89	0,55	0,29	0,70	1,54	-8,26	
AMOEBA originale		1,12	-0,06	0,81	-0,22	0,54	-6,48	
AMOEBA _p non zero		1,22	-0,07	0,86	-0,22	0,57	-6,61	
AMOEBA _p non zero std		1,20	-0,11	1,11	-0,24	0,76	-6,78	
AMOEBA _p pesi uguali		2,96	-0,30	1,92	-0,19	1,44	-9,22	
Udine (cod. ISTAT 30129)								
Contiguità torre I ord.	21,03	5,97	-0,08	2,17	0,22	2,31	-6,21	23,10
Cliff e Ord		6,48	-0,07	2,18	0,20	2,30	-6,72	
AMOEBA originale		11,94	-0,08	2,88	0,32	3,11	-12,98	
AMOEBA _p non zero		12,00	-0,09	2,88	0,32	3,11	-13,04	
AMOEBA _p non zero std		13,99	-0,13	2,98	0,26	3,11	-15,03	
AMOEBA _p pesi uguali		16,83	-0,20	2,71	0,21	2,72	-17,48	

FONTE: elaborazione propria su dati SISR-Regione Autonoma Friuli Venezia Giulia

zione usate nel modello (1) che si ricorda essere l'età della popolazione e il tipo di spesa sanitaria, permettono ai due capoluoghi osservati di diminuire il proprio tasso di variazione finale. La grande differenza tra i due casi consiste nel fatto che, mentre nel caso di Trieste tali fattori riescono a compensare in modo più che ampio gli effetti degli incrementi del vicinato, per Udine invece questo non accade lasciando, quindi, il suo tasso di variazione ad un valore superiore di quello regionale.

Leggendo invece la Tabella 1 in senso orizzontale, ossia per diversa tipologia della matrice geografica, si individuano sei scomposizioni diverse di cui le prime due basate, rispettivamente, sul criterio di contiguità di primo ordine e sul criterio Cliff e Ord (che hanno la medesima configurazione vista in Figura 1), le altre quattro sono di tipo AMOEBA: l'algoritmo originale e le tre declinazioni parsimoniose che differiscono unicamente nel calcolo dei pesi (quindi non nella configurazione spaziale presentata sempre in Figura 1). Tale selezione, tra tutte le possibilità di costruzione, è motivata dal fatto che tutte queste matrici geografiche condividono per costruzione un criterio di contiguità: Cliff e Ord è una raffinamento di questo criterio che tiene conto anche di altri elementi cartografici come la distanza tra baricentri (cfr. Unwin, 1986, pp. 96-101) e la porzione di perimetro in comune, mentre gli algoritmi AMOEBA, come è immediato vedere dalle loro cartografie, cominciano a costruire il vicinato a partire da comuni confinanti: il caso di Faedis è graficamente emblematico perché viene selezionato un unico comune confinante, senza questo l'algoritmo AMOEBA non sarebbe stato in grado di costruire un vicinato.

Cercando di distillare un risultato generale da tutte le circa 1300 scomposizioni eseguite per questo lavoro, si è empiricamente osservato che se il vicinato è ben costruito, ossia non ipotizzato a priori senza nessuna evidenza empirica successiva, le componenti di confronto CFR e intravicinato INTRA tendono spesso a contribuire in modo più rilevante alla scomposizione del tasso di variazione impedendo che tutta la spiegazione ricada unicamente sulla componente locale LOC. In altre parole, se il vicinato è ben costruito il suo ruolo nella scomposizione risulta più evidente: questo accade per gli algoritmi AMOEBA soprattutto nel caso di Udine, dove il contributo all'innalzamento della spesa sanitaria

pro-capite dei comuni "vicini" è ben messo in evidenza dalla Tabella 1, mentre i risultati per Trieste tendono a convergere, dove solo la scomposizione di tipo AMOEBAp a pesi uguali porta ad una scomposizione più estrema, ma si ricorda che quest'ultima è basata su un'ipotesi semplicistica posta ai soli fini di controllo. Quindi, nel caso di Trieste gli algoritmi AMOEBA tendono a confermare, in linea di massima, i risultati ottenuti attraverso le matrici più tradizionali.

3. Elementi di sintesi e risultati generali derivati dal caso di studio

L'obiettivo dichiarato di questo contributo era di duplice natura. Il primo consisteva nel fornire una soluzione generale al problema dell'analisi delle variazioni temporali di un fenomeno economico gravante su un territorio, quale quello di una voce di un bilancio regionale. Questo a patto di disporre di dati estremamente dettagliati, quali quelli relativi a singoli soggetti (come i cittadini) o singoli oggetti (come i beni portatori di particolari valori artistico-culturali) che solo un'Amministrazione Regionale è in grado di fornire. La via proposta è la *Spatial Shift-Share Analysis*, tecnica molto nota ed utilizzata in diversi campi scientifici da quello socio-economico a quello ambientale ed energetico fino alla modellizzazione delle catastrofi naturali.

L'analisi della letteratura ha messo in evidenza come per decenni questo tipo di analisi fosse considerata "territoriale" in modo del tutto scorretto poiché il modello sottostante (Dunn e sue evoluzioni) non teneva in alcun modo conto delle relazioni territoriali che influiscono, anche drasticamente, sulle manifestazioni di un fenomeno economico, e quindi anche sulle sue variazioni nel tempo. Solo agli inizi del nuovo millennio, Nazara e Hewings hanno suggerito di inserire nel modello una matrice geografica che permetta di tenere conto formalmente delle relazioni tra le unità territoriali (siano esse amministrative del tipo NUTS o progettate a tavolino) come ipotesi sottostante allo modello stesso.

Il secondo obiettivo di questo contributo, che si muove in linea con un precedente lavoro relativo all'occupazione regionale proposto su questo Bollettino nel 2014, è proprio quello di cercare di rimuovere la neces-

sità di formulare ipotesi a priori sul tipo di relazione territoriale esistente tra unità territoriale e suo vicinato. In altre parole, la costruzione del vicinato, ossia dell'insieme delle unità considerate "vicine", è ora implicita nella stessa *Spatial Shift-Share Analysis*: non più un'ipotesi a priori, ma un risultato intermedio dell'analisi stessa.

Per fare questo però è necessario individuare un criterio guida presente in letteratura che sia utile allo scopo: la Prima legge della geografia di Tobler mette l'accento sulla correlazione spaziale, ossia esplicita il fatto che, in generale, nelle unità territoriali "vicine" il fenomeno deve comporsi più o meno nello stesso modo; quindi, in altre parole, le unità presenti nel vicinato devono avere un comportamento omogeneo ovviamente dal punto di vista del fenomeno economico studiato. La via proposta per implementare questa legge nell'analisi è stata quella di adattare l'algoritmo AMOEBA, introdotto da Getis e Ord poco dopo la proposta di Nazara e Hewings e utilizzato nella letteratura GIS per trovare soluzioni a problemi di perimetrazione territoriale basate sull'omogeneità di comportamento. Un simile passo richiede l'analisi cartografica del risultato dell'algoritmo, ossia per ogni vicinato prodotto in ogni scomposizione. Ed ecco esplicitato il valore geografico e cartografico di questo contributo poiché una simile procedura, di una complessità estremamente più elevata di quella proposta originariamente, non solo porta a superare definitivamente la tradizionale esclusione del territorio ad opera degli economisti del secolo scorso, ma rimuove la necessità di formulare ipotesi a priori sul tipo di relazioni territoriali e rende necessaria un'analisi cartografica come passo fondamentale del procedimento di scomposizione. Solo grazie ad essa, come dimostrato in sede empirica, possono essere non solo facilmente individuati i casi problematici, ma anche interpretati i risultati numerici del modello di scomposizione. In altre parole, da una semplice scomposizione numerica si passa ad una vera e propria analisi (di tipo multilivello poiché considera regione, vicinato e comune) che, per quantità e qualità delle informazioni sulle geometrie e per complessità della procedura automatica, può essere svolta solo grazie ad un linguaggio di programmazione che si appoggia ad una cartografia su base GIS. In sostanza, se l'analisi della bibliografia ha messo in evidenza come il problema qui discusso possa essere ora, a pieno titolo,

risolto attraverso la *Spatial Shift-Share Analysis* poiché questa ha conquistato una dignità geografica prima non posseduta, è anche possibile affermare che questo ampio discorso permette di garantire direttamente a monte la qualità dei risultati ottenuti.

L'analisi empirica, come da subito dichiarato, è stata condotta sulla variazione della spesa sanitaria regionale pro-capite non con l'intento di studiare il fenomeno in sé, peraltro già studiato in altre pubblicazioni, ma con quello di estrapolare ulteriori risultati scientifici di carattere generale, ed estensibili a future applicazioni di ricerca, dalle 1300 scomposizioni portate a termine. Infatti, essendo la nuova proposta agli albori, questo contributo presenta un carattere fortemente sperimentale soprattutto per quanto concerne l'evoluzione e l'affinamento dell'algoritmo AMOEBA.

Il primo risultato generale trovato è relativo al fatto che una "buona costruzione" del vicinato, ossia una sua maggiore omogeneità, porta ad una enfattizzazione del ruolo giocato dal vicinato all'interno della scomposizione, mentre una costruzione di bassa qualità, come ad esempio un'ipotesi a priori non suffragata dalla realtà empirica, comporta il fatto che tutto il contenuto informativo venga semplicemente scaricato sulla componente locale riducendo drasticamente la capacità esplicativa del modello stesso.

In secondo luogo, le scomposizioni basate sull'algoritmo AMOEBA, che come detto soddisfano pienamente la precedente osservazione, si dimostrano un'ottima alternativa alle ipotesi a priori basate sulla contiguità o sulla proposta di Cliff e Ord soprattutto quando il caso in esame non è problematico, così come dimostrato empiricamente dal caso di Udine e Trieste. Nel caso invece in cui ci siano problemi di mancanza di dati oppure di variazioni anomale, generalmente legate a piccole realtà territoriali, l'analisi cartografica permette di capire come si sviluppa spazialmente il vicinato. Il caso qui discusso di Faedis è emblematico poiché l'algoritmo AMOEBA individua un unico comune confinante di primo ordine, ossia l'unico che si comporta in modo analogo sempre in termini di variazione della spesa sanitaria pro-capite. Questo permette all'algoritmo comunque di procedere nella sua costruzione del vicinato, ma vi sono alcuni comuni del FVG, come ad esempio Gorizia, per cui l'algoritmo non riesce nemmeno ad individuare un

suo comune confinante che si comporti in modo analogo per cui questo fallisce poiché non è capace di individuare un vicinato (l'insieme rimane quindi vuoto). In questi casi, che si ritrovano soprattutto nei comuni ai bordi della regione dove il numero dei loro confinanti è limitato, l'unico modo per procedere con l'analisi *Shift-Share* è purtroppo quello di ricorrere alle ipotesi a priori tradizionali, come la semplice contiguità, con tutti i rischi che questa ipotesi comporta. L'esperienza qui svolta porta a concludere che, se il fenomeno analizzato è ben diffuso sul territorio, questi casi dovrebbero verificarsi in numero piuttosto limitato e, quindi, sembra più corretto segnalare che l'analisi non è praticabile piuttosto che assumersi il rischio di un'ipotesi a priori non suffragata dalla realtà empirica dei dati.

Ritornando al lato sperimentale di questo lavoro, sempre sulla base di questa esperienza, dal confronto tra gli algoritmi AMOEBA originale e parsimonioso è emerso che quest'ultima proposta, portando a vicinati più ridotti, comunque diversi dai quelli ottenuti con la mera contiguità di primo ordine, non solo riduce i tempi di calcolo (aspetto da non sottovalutare quando i comuni da analizzare sono molti), ma implica anche minori violazioni della legge di Tobler. L'osservazione empirica ha comunque messo in evidenza che, nella maggioranza dei casi, le due versioni dell'algoritmo convergono allo stesso vicinato. In caso contrario, generalmente si tratta di unità territoriali in cui vi è un problema di mancanza di dati oppure di entità ridotte dal punto di vista del fenomeno considerato, come può capitare soprattutto per i comuni di montagna anche per effetto del loro spopolamento.

Proprio da quest'ultima osservazione è possibile trarre un ultimo risultato generale. Infatti, se l'algo-

ritmo parsimonioso era nato con l'intento di migliorare, per alcuni aspetti, l'algoritmo originale ora invece è stato evidenziato un aspetto notevole non atteso: la divergenza delle due configurazioni di vicinato può essere utilizzata come cartina di tornasole che permette di evidenziare subito, e in modo automatico, la presenza di casi difficili da analizzare dove le scomposizioni ottenute sono fortemente dipendenti dalle ipotesi fatte a livello di costruzione del vicinato. Tale risultato fornisce un grande contributo operativo, e questo soprattutto nel caso di regioni con una fitta partizione comunale, perché indica immediatamente al geografo su quali comuni porre maggiore attenzione prima dell'interpretazione numerica dei dati.

A conclusione di questo lavoro si vuole puntualizzare che, allo stato attuale, la ricerca su questo tipo di analisi è ancora in corso, quindi la SSSA risulta ancora suscettibile di ulteriori avanzamenti e affinamenti. Dal punto di vista meramente operativo rimane ancora troppo macchinoso dover continuamente operare tra GIS e linguaggio di programmazione esterno. Grazie all'attuale sviluppo di GIS *open source*, un rilevante passo futuro potrebbe essere quello di integrare la *Spatial Shift-Share Analysis* nei suoi *plug-in* di analisi, così come già accaduto per tecniche come la *Kernel Density Estimation*, molto utilizzata dai geografi per costruire cartografie di svariati tipi, come ad esempio quella inerente ai valori di fruizione dei beni culturali utile per la predisposizione di un Piano Paesaggistico Regionale, non solo relative a partizioni amministrative ma anche a *spazi di rete* (cfr. Borruso, 2008). Solo con questo passo sarà possibile "liberarla" dal contesto accademico e diffonderla anche tra i *practitioner*, pubblici e privati, dei GIS.

Bibliografia

- Aldstadt J. e Getis A. (2004), "Constructing the spatial weights matrix using a local statistics", *Geographical Analysis*, 36, 2, pp. 90-104.
- Aldstadt J. e Getis A. (2006), "Using AMOEBA to create a spatial weights matrix and identify spatial clusters", *Geographical Analysis*, 38, pp. 327-343.
- Bencardino F. e Prezioso M. (2006), *Geografia economica*, McGraw Hill, Milano.
- Borruso G. (2008), "Network density estimation: a GIS approach for analysing point patterns in a network space", *Transactions in GIS*, 12, 3, pp. 377-402.
- Cliff A.D. e Ord J.K. (1981), *Spatial process, Models and Applications*, Pion, London.
- Conti S. (2012), *I territori dell'economia. Fondamenti di geografia economica*, UTET, Milano.
- Dunn E.S. (1960), "A statistical and analytical technique for regional analysis", *Paper and Proceedings of the Regional Science Association*, 6, pp. 97-112.
- Getis A. e Ord J.K. (1992), "The analysis of spatial association by use of distance statistics", *Geographical Analysis*, 24, pp. 189-206.
- Grassetti L. e Rizzi L. (2014), *A spatial approach to the analysis of individual health care expenditures: the case of the Italian region of Friuli-Venezia Giulia*, working paper DIES, 4, Università degli Studi di Udine, ottobre 2014.
- Griffith D.A. (1996), *Some guidelines for specifying the geographic weights matrix contained in Spatial Statistical Models*, in: Arlinghaus S.L. (ed.), *op. cit.*, pp. 65-82.
- Haynes K.E. e Parajuli J. (2015), *Shift-share and its new extensions*, in: Karlsson C., Andersson M. e Norman T. (eds), *Handbook of research methods and applications in economic geography*, Edward Elgar Pub., Cheltenham, pp. 83-117.
- Jones J.H. (1940), "The Report of the Royal Commission on the Distribution of the Industrial Population", *Journal of the Royal Statistical Society*, 103, 3, pp. 323-343.
- Nazara S. e Hewings G.J.D. (2004), "Spatial structure and taxonomy of decomposition in shift-share analysis", *Growth & Change*, 35, pp. 476-490.
- Ord J.K. e Getis A. (1995), "Local spatial autocorrelation statistics: distributional issues and an application", *Geographical Analysis*, 27, 4, pp. 286-306.
- Piergentili P., Simon G., Paccagnella O., Grassetti L., Rizzi L. et Samani F. (2015), "Risk-adjusted models of costs referable to general practitioners based on administrative databases in the Friuli Venezia Giulia region in Northern Italy", *Archives of community medicine & public health*, 1, 1, pp. 12-21.
- R Core Team (2013), "R: A language and environment for statistical computing, R Foundation for Statistical Computing", Vienna, (<http://www.R-project.org>, data di ultima consultazione: 23/10/2014).
- Tobler W.R. (1970), "A computer movie simulation urban growth in the Detroit region", *Economic Geography*, 46, pp. 234-240.
- Unwin S. (1986), *Analisi spaziale. Un'introduzione geocartografica*, F. Angeli, Milano.
- Zaccomer G.P. (2006), "Shift-share analysis with spatial structure: an application to Italian industrial districts", *Transition Studies Review*, 13, 1, Springer, pp. 213-227.
- Zaccomer G.P. (2008), "Spatial structure in a shift-share decomposition: new results for the Italian industrial districts case", *Transition Studies Review*, 15, 1, Springer, pp. 111-123.
- Zaccomer G.P. e Mason P. (2008), "A definition of neighborhoods based on local labor systems: a regional application on employment data", *Statistica & Applicazioni*, 6, 2, pp. 215-236.

Zaccomer G.P. e Mason P. (2011), "A new spatial shift-share decomposition for the regional growth analysis: a local study of the employment based on Italian Business Statistical Register", *Statistical Methods & Applications*, 3, pp. 329-356.

Zaccomer G.P. e Grassetto L. (2014), "La cartografia come strumento di interpretazione dei risultati di un modello di scomposizione spaziale: nuove proposte con applicazione al caso dell'occupazione in Friuli Venezia Giulia", *Bollettino dell'Associazione Italiana di Cartografia*, 152, pp. 56-72.

Zani S. (a cura di) (1993), *Metodi statistici per le analisi territoriali*, F. Angeli, Milano.