

Tecniche di analisi ed interpolazione spaziale nello studio dei sistemi produttivi: il caso della Provincia di Cuneo

Vincenzo Demetrio, Paola Guerreschi

Politecnico e Università di Torino, Dipartimento Interateneo Territorio, viale Mattioli, 39 – 10125 Torino
tel. 011-0907481, fax 011-0907451
vincenzo.demetrio@unito.it, paola.guerreschi@unito.it

Riassunto

Questo lavoro mira a verificare la possibilità di utilizzare, in assenza dell'ipotesi di autocorrelazione dei dati esaminati, gli strumenti dell'analisi spaziale per ottenere previsioni attendibili o perlomeno ragionevoli, sull'universo a partire da un campione rappresentativo di dati a nostra disposizione. Lo studio, infatti, farà riferimento a fenomeni di tipo socio-economico che a differenza di quelli naturali non si caratterizzano per l'interdipendenza dei dati alle diverse scale dimensionali.

In questo quadro l'obiettivo che ci poniamo è certamente di minima poiché siamo consapevoli che la previsione non potrà riguardare l'intera area oggetto d'esame, ma soltanto un suo sott'insieme e non potrà essere precisa, ma soltanto ragionevole. Nondimeno riteniamo che questo tipo di approfondimento possa rappresentare uno stimolo per l'adeguamento di metodi e modelli, propri dell'analisi spaziale, anche alle analisi di tipo socio-economico caratterizzate da ipotesi di partenza sui dati più restrittive.

Abstract

This paper is aimed to verify the possibility of using, in absence of examined data self-correlation hypothesis, the instruments of spatial analysis in order to obtain reliable or at least reasonable forecasts concerning the universe to which the representative sample of data belong. The study, in fact, refers to socio-economic phenomena which differ from those natural for the independence of data at the various dimensional scales.

In this framework our goal is sure of minim since we know that forecast will not refer to the entire area of analysis, but only to a part of it and could not be *precise*, but only *reasonable*. Nevertheless we think that this work may represent a stimulus for the adaptation of methods and models which belong to the spatial analysis to socio-economic investigation generally characterized by more restrictive hypothesis concerning data used.

Introduzione

La comprensione della distribuzione di un certo fenomeno nello spazio, risulta cruciale in molteplici ambiti della conoscenza che spaziano, solo per citarne qualcuno, dalla medicina all'ambiente, dalla geologia all'urbanistica. Le analisi spaziali sono diventate sempre più comuni grazie all'ampia disponibilità di sistemi GIS, talvolta a basso costo, ma sempre di facile utilizzo, che permettono agli utenti anche meno esperti di visualizzare la distribuzione geografica o prevedere l'andamento di un fenomeno partendo soltanto da una banca dati e da una base cartografica. Nondimeno, proprio l'apparente semplicità di utilizzo degli strumenti di analisi spaziale e la facile reperibilità di dati georiferibili di varia natura può condurre a previsioni fuorvianti perché costruite senza tener conto delle più elementari ipotesi circa le proprietà dei dati.

In questo quadro, il lavoro che qui proponiamo, si pone l'obiettivo di offrire un esempio di applicazione di tecniche di analisi e interpolazione spaziale allo studio dei sistemi socio-economici.

L'analisi è stata condotta a partire dai dati desunti dalle interviste realizzate per una recente indagine sul Piemonte (Demetrio e Giaccaria, 2010). L'ambito territoriale selezionato è quello della provincia di Cuneo, che per molti versi continua a essere un'area poco studiata e pertanto poco conosciuta in ambito regionale.

Inizialmente il campione delle imprese a nostra disposizione si componeva di 232 unità, rappresentative dell'intero tessuto produttivo provinciale. Tuttavia, la necessità di mettere a punto una metodologia di analisi adeguata e contemporaneamente di avere superiori possibilità di controllo e verifica dei risultati finali in assenza di prassi consolidate, ci ha spinto a restringere il campo d'azione a una soltanto delle 5 aree omogenee che compongono la provincia di Cuneo: il Saluzzese, limitando di conseguenza il campione a 22 imprese.

Basi cartografiche e geo-referenziazione delle imprese

Dal punto di vista empirico la fase preparatoria dell'analisi si è articolata nelle tre fasi qui elencate:

- *Scelta della cartografia di base.* Tra le alternative possibili (CTR raster, CTRN, ecc...), abbiamo preferito utilizzare i fogli della CTRN in WGS84, poiché permettono di effettuare analisi spaziali rispetto agli elementi della carta stessa (es. viabilità, distanza da punti di interesse, ecc...).
- *Georeferenziazione delle 22 aziende del campione.* Con l'ausilio dell'applicativo Google Earth, avviando una ricerca per ragione sociale o indirizzo, sono state individuate le singole localizzazioni, successivamente restituite in forma di punti. Il codice identificativo di ogni geometria (corrispondente al numero dell'intervista) è stato la chiave che ha permesso di effettuare l'unione con la matrice dei dati.
- *Predisposizione della base dati da analizzare.* Il questionario sottoposto alle imprese intervistate prevedeva l'acquisizione di dati quantitativi (per es. fatturato, export, utile, ecc...) e qualitative (per esempio valutazioni sulle dinamiche cognitive e relazionali). In questo secondo caso è stato necessario trasformare i dati qualitativi in indicatori di tipo numerico. In altri termini i giudizi espressi dalle imprese sono stati convertiti in indicatori idonei per poter essere analizzati con ArcMap. Nella maggior parte dei casi la conversione ha visto l'utilizzazione di valori compresi tra 0 e 4, dove con 0 si è identificato un giudizio "molto negativo" rispetto al quesito posto, con 1 un giudizio "negativo", con 2 un giudizio "ininfluente", con 3 un giudizio "positivo" e infine con 4 un giudizio "molto positivo". Nei casi di risposte del tipo si/no i valori trasposti sono stati di tipo booleano (0, 1).

Origine e natura della banca dati utilizzata

La banca dati utilizzata descrive ciascuna impresa del campione attraverso 35 variabili capaci di fornire un'adeguata comprensione delle dinamiche aziendali (micro) e, per estensione, dei comportamenti territoriali dei sistemi d'impresa (meso).

Il primo gruppo di 9 variabili fa riferimento alla situazione economica dell'impresa, il secondo gruppo di 13 variabili alle dinamiche cognitive e infine il terzo, anch'esso composto di 13 variabili, relativo le dinamiche relazionali. Sui gruppi così costruiti, abbiamo effettuato un'analisi delle componenti principali finalizzata a identificare per ciascun gruppo 3 variabili in grado di spiegare la maggior parte della variabilità interna al gruppo in esame.

Nell'analisi territoriale delle dinamiche economiche, relazionali e cognitive sono state considerate soltanto le componenti principali di ciascun gruppo in modo tale da non dover fare un'eccessivo e impegnativo ricorso alla MapAlgebra.

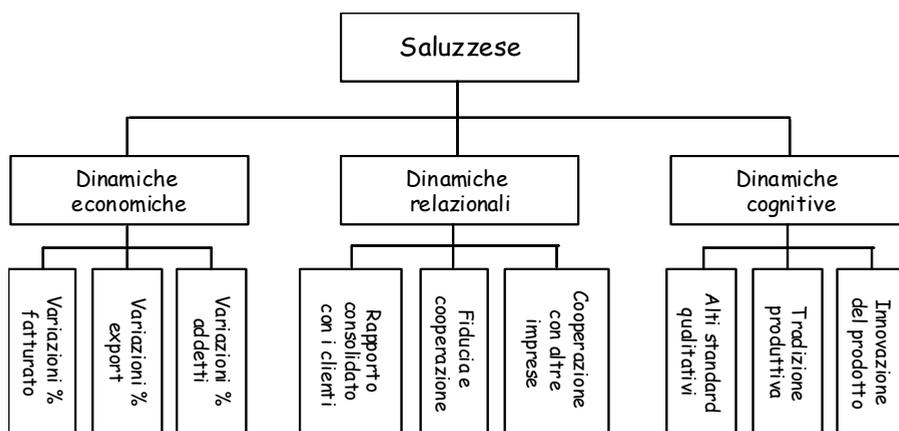


Figura 1 – Selezione delle variabili caratterizzanti le dinamiche di impresa.

Possibilità e limiti di applicazione dei modelli geo-statistici all'analisi economica

L'applicazione di metodi e tecniche geostatistiche allo studio di variabili socio economiche non è immediata o possibile *sic et simpliciter*. Occorre *in primis* controllare la natura dei dati da analizzare, verificandone le caratteristiche d'indipendenza (base per l'inferenza statistica) o di interdipendenza (fondamentale per le analisi spaziali).

Sappiamo che la maggior parte dei fenomeni naturali si connota per un certo grado di dipendenza spaziale nella distribuzione dei dati a tutte le scale dimensionali. Ciò comporta la presenza di strutture spaziali in cui i valori sono auto-correlati. In altre parole, le coppie di punti posti ad una certa distanza sono più simili (autocorrelazione positiva) o meno simili (autocorrelazione negativa) rispetto a quanto sarebbe lecito attendere da coppie di osservazioni associate casualmente (Legendre, 1993). In presenza dell'ipotesi di autocorrelazione i valori associati a punti nello spazio possono essere, almeno parzialmente, prevedibili quando si conoscano valori dei relativi a punti vicini, dato che questi non sono stocasticamente indipendenti gli uni dagli altri. Nello studio dei fenomeni socio-economici invece, l'ipotesi di autocorrelazione dei dati spaziali non è affatto scontata, ma al contrario è una condizione da sottoporre a verifica.

In assenza di autocorrelazione tra i dati analizzati, infatti, è necessario un adeguamento di metodi in grado di offrire una ragionevole attendibilità alle previsioni ottenute. L'utilizzo di dati indipendenti, infatti, mina alle fondamenta la capacità dei metodi di interpolazione geostatistica, quali per esempio (IDW, Kriging, ecc...) di interpretare correttamente la realtà rendendone di fatto fuorvianti le previsioni ottenute.

In questo quadro la costruzione del modello di previsione diventa necessariamente più complessa e riferibile obbligatoriamente a un contesto spaziale più limitato nel quale sia possibile ritrovare fenomeni di interdipendenza tra valori.

Inevitabilmente nel caso di dati non correlati, una delle problematiche alla quale occorre offrire una prima, seppur parziale, soluzione è quella di come individuare un contesto territoriale nel quale sia possibile effettuare previsioni con ragionevole sicurezza.

Analisi dei dati e soluzioni tecniche adottate

Abbiamo già osservato che quando si utilizzino dati correlati, la possibilità di ottenere previsioni realistiche dei valori propri delle localizzazioni in cui non è stata effettuata alcuna misurazione, dipende dalla capacità di costruire un modello, che leghi i valori osservati con quelli incogniti.

Viceversa, nel caso in cui i dati a disposizione siano non correlati (per es. dati socio-economici), invece, i metodi di interpolazione geostatistica non sono in grado di prevedere correttamente la

realità. In questi casi è possibile optare per la costruzione di un modello di zonizzazione in cui l'area oggetto d'esame viene suddivisa in sottoinsiemi disgiunti, sulla base dell'uniformità dei parametri osservati. I valori incogniti sono assunti essere uniformi ai dati noti.

Empiricamente abbiamo ottenuto questo risultato utilizzando le *Voronoi map*¹ (ArcGIS/geostatistical analyst/explore data/voronoi map) calcolate per ciascuna delle 3 variabili principali di ciascun gruppo. I parametri settati per la definizione dei poligoni sono stati rispettivamente: *layer* (il file delle 22 ditte), *attribute* (le 9 variabili) infine *type* (cluster). Nella tabella di attributi del file vettoriale ottenuto sono stati automaticamente aggiunti molti dati quali *mean, mode, median, cluster, standard deviation*.

L'individuazione delle aree in cui i comportamenti delle imprese risultano omogenei (rispetto alle singole variabili) è stata possibile sommando i singoli poligoni caratterizzati dall'appartenenza allo stesso cluster. Naturalmente l'esecuzione di queste operazioni richiede di operare in ambiente raster. Pertanto si è reso necessario effettuare una conversione da polygon a raster (ArcGIS\ArcToolbox\Toolboxes\Conversion Tools.tbx \ PolygonToRaster) andando a definire la dimensione della cella (*sizecell*). Il campo *value* ottenuto indica l'inclusione della singola cella in un cluster definito o in altri termini un comportamento delle imprese tendenzialmente simile.

Abbiamo considerato i valori -1 e 0 associati al cluster come indicativi di un comportamento atipico delle imprese rispetto al contesto territoriale circostante, mentre i valori 1, 2, 3 e 4 come sintomatici di un progressivo aumento della somiglianza negli andamenti delle variabili aziendali: molto debole, debole, forte, molto forte.

Per ottenere la mappa delle dinamiche cognitive, economiche e relazionali e successivamente quella complessiva dell'area, si sono utilizzate le funzioni della Map Algebra.

Affinché si potessero effettuare queste operazioni era necessario ricorrere alla funzione di *reclassify* dei cluster attribuendo agli stessi un punteggio (ArcGIS/ArcToolbox/spatialanalysttools/reclass/reclassify). Al cluster -1 abbiamo quindi attribuito punteggio 0, al cluster 0 punteggio 1, al cluster 1 punteggio 3, al cluster 2 punteggio 5, al cluster 3 punteggio 7, al cluster 4 punteggio 10.

Utilizzando il modulo di Spatial Analyst abbiamo usato la funzione di *Overlay* con il metodo della *Weighted Sum* (ArcGIS/ArcToolbox/spatial analyst tools/overlay/weighted sum). Il *weighted* di ciascun raster è stato impostato pari ad 1.

Attraverso questo procedimento abbiamo ottenuto, per ciascuno dei 3 gruppi di variabili, una rappresentazione del territorio suddiviso in poligoni. Raggruppando i poligoni contigui caratterizzati da andamenti similari delle variabili siamo in grado di identificare aree (quelle più scure in figura 2) in cui ci sia aspetta che le imprese di cui non si hanno informazioni abbiano performance in linea con quelle del campione.

Le tre carte di sintesi oltre a fornire alcuni importanti elementi per la comprensione dei fenomeni in esame, permettono di valutare la possibilità di utilizzare i dati campionari per prevedere il comportamento dell'universo o di parte di esso.

Ciò che qui emerge con maggiore evidenza è che i comportamenti relazionali, cognitivi ed economici delle imprese del campione abbiano andamenti spaziali differenti. Per esempio le imprese operanti nelle aree montane (nord-ovest) hanno dinamiche relazionali (figura 2B) più simili rispetto alle imprese che operano in pianura, mentre le dinamiche cognitive (figura 2C) sono maggiormente coerenti nell'area nord est del Saluzzese, tradizionalmente influenzata dalla vicinanza con l'importante asse industriale Torino-Cuneo. Infine, è senz'altro ben visibile nella figura 2A un'ampia area omogenea dal punto di vista degli andamenti delle variabili economiche.

¹ I poligoni Voronoi sono creati in modo tale che ciascuna area contenga esattamente un punto della banca dati esaminata e tutti i punti del piano che sono più vicini a quel punto piuttosto che ad un altro punto del dataset. Le regioni definite dalla poligonizzazione di Voronoi sono anche dette *aree di prossimità* o *bacini di attrazione*. Tali poligoni servono anche alla ricostruzione di un fenomeno nello spazio a partire da osservazioni campionarie.

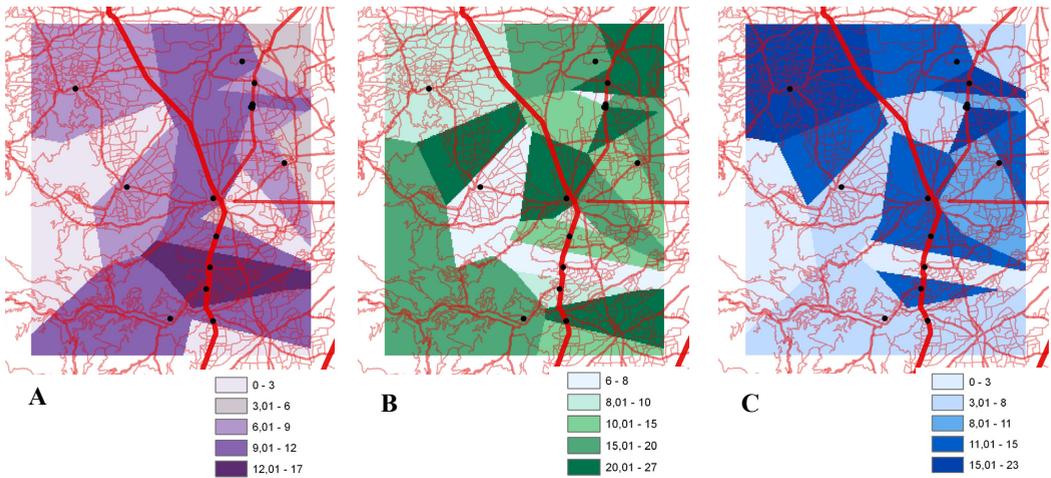


Figura 2 – Le dinamiche di impresa. A - dinamiche economiche; B - dinamiche relazionali; C - dinamiche cognitive.

Individuare aree omogenee dal punto delle dinamiche delle variabili esaminate consente di stimare le informazioni non note semplicemente estendendo l'ipotesi di omogeneità² di comportamento anche alle imprese per le quali non si abbiano rilevazioni e di precisare ampiezza e la direzionalità del fenomeno esaminato, dato che non tutti i fenomeni riguardano sempre l'intera area oggetto di studio. Per esempio alcuni possono interessare solo unità territoriali vicine (si pensi alle zone industriali presenti su di un comune o sul territorio di comuni limitrofi) oppure seguire direzioni precise (legate per esempio alla presenza di assi viari principali) ed essere del tutto assenti in altre zone.

Grazie agli strumenti messi a disposizione da ArcGIS è stato possibile verificare la correttezza dei risultati ottenuti. In particolare, limitandoci alle sole dinamiche economiche, abbiamo valutato l'ampiezza e la direzione del fenomeno utilizzando la funzione *directional distribution* (ArcGIS\ ArcToolbox\ spatial statistic tools\ Measuring Geographic Distributions\ directional distribution (standard deviational ellipse)).

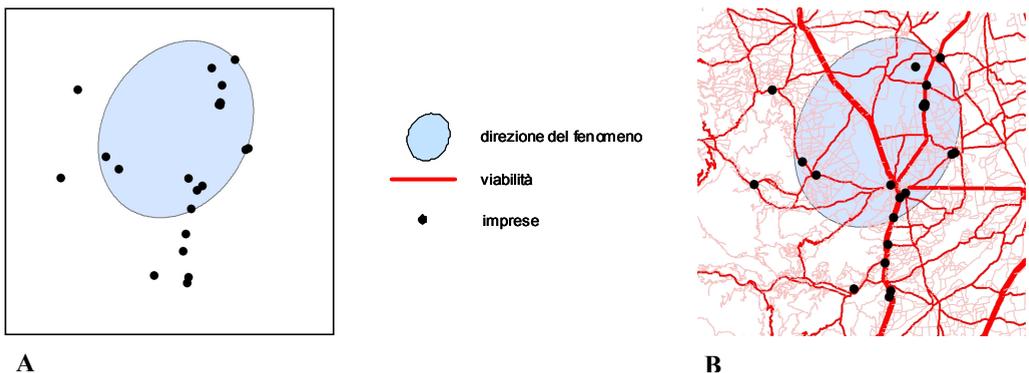


Figura 3 – Ampiezza e distribuzione dei comportamenti economici omogenei.

² Si tratta di un'ipotesi di omogeneità locale

Coerentemente con le attese, sono emerse una chiara concentrazione territoriale e direzione del fenomeno esaminato (figura 3A). È possibile legare le ragioni di tale concentrazione ad alcuni aspetti logistici quali per esempio la viabilità o alla vicinanza ad aree maggiormente performanti quali per esempio la direttrice Torino-Cuneo.

Sovrapponendo l'area di concentrazione con la viabilità tematizzata per ordine di importanza (figura 3B), emerge che le imprese collocate lungo i principali assi viari risultano caratterizzate da dinamiche economiche più simili rispetto alle attività situate lontano dalle strade più importanti. Considerazioni simili valgono anche per le imprese prossime ai confini con la direttrice Torino-Cuneo.

È utile qui ricordare che dinamiche economiche più simili, non significa in generale migliore performance delle imprese clusterizzate. L'analisi, infatti, si limita a individuare l'omogeneità dei comportamenti per individuare quegli ambiti in cui è possibile una stima dei valori non noti nella generale ipotesi di dati non correlati.

Riflessioni conclusive

In questo lavoro abbiamo testato la possibilità di fornire ragionevoli previsioni partendo da un campione rappresentativo di dati indipendenti.

Naturalmente l'assenza di auto-correlazione dei fenomeni ci ha spinto a limitare l'area alla quale riferire la previsione e ad accontentarci di dati tendenziali più che puntuali. Nondimeno pensiamo che la procedura seguita possa rappresentare un primo generico passo verso la definizione di prassi consolidate con le quali affrontare problematiche socio- economiche con gli strumenti propri dell'analisi spaziale.

Nel caso in esame questa procedura ha permesso di evidenziare alcune interessanti ricorrenze spaziali quali:

- la marginalità delle aree montane caratterizzate da dinamiche cognitive simili, basate sullo sfruttamento del bagaglio di conoscenze sedimentato sul territorio più che sull'acquisizione di nuove informazioni dall'esterno;
- la forte tendenza all'interazione delle aree di confine del Saluzzese con la direttrice Torino-Cuneo a nord e con il cuneese a sud;
- il comportamento di tipo "distrettuale" delle imprese collocate lungo i principali assi viari in prossimità della direttrice Torino-Cuneo a verifica della presenza di forme di organizzazione della produzione riconducibili a questo ideal-tipo.

Riferimenti bibliografici

Demetrio V., Giaccaria P. (2010), *Geografia del sistema manifatturiero piemontese*, Carocci, Roma
Legendre P. (1993), "Spatial autocorrelation: Trouble or new paradigm?" in *Ecology*, 74 :1659-1673