

Studio dei flussi migratori in Italia mediante analisi di autocorrelazione spaziale

Grazia Scardaccione, Francesco Scorza, Giuseppe Las Casas, Beniamino Murgante

Laboratorio di Ingegneria dei Sistemi Urbani e Territoriali
Università degli Studi della Basilicata, 10, Viale dell'Ateneo Lucano, 85100, Potenza
nome.cognome@unibas.it

Abstract. Nel corso degli ultimi decenni, il fenomeno della migrazione sta assumendo dimensioni tali da suscitare interesse crescente a livello di opinione pubblica, tanto da diventare oggetto di numerosi studi e ricerche. La migrazione è un fenomeno complesso che richiede un sistema di analisi che va oltre le considerazioni demografiche ed economiche. Partendo da questa riflessione, con questo lavoro, ci si propone di studiare la distribuzione spaziale della presenza straniera in Italia, al fine di individuare le diverse linee di demarcazione geografica sulla base di diverse interpretazioni. La tradizionale analisi statistica suggerisce alcuni indici convenzionali che permettono di quantificare il fenomeno della migrazione: il Quoziente di Localizzazione e l'Indice di Dissimilarità. Tali indici sono stati confrontati con misure di autocorrelazione spaziale a livello globale e locale, l'Indice di Moran e l'indicatore LISA, (Local Indicator of Spatial Association). Questo studio contribuisce all'identificazione delle aree di destinazione dei flussi migratori in Italia, applicando l'approccio geostatistico a dati riguardanti variabili elementari facilmente reperibili, in modo da progettare una metodologia riproducibile e che possa essere estesa anche a casi diversi da quello italiano.

Keywords: Immigrazione straniera, Quoziente di Localizzazione, Indice di Dissimilarità, Analisi Statistica Spaziale, Indicatori di Autocorrelazione Spaziale Globale e Locale.

1. Introduzione

L'analisi dei flussi migratori è un campo di applicazione che coinvolge numerose discipline (dalla geografia, all'economia, alle scienze sociali) in quanto il fenomeno migratorio rappresenta un fattore chiave nell'evoluzione delle dinamiche demografiche a differenti scale, con ripercussioni su economia, cultura, territorio e ambiente (Sjaastad 1962) (Walker 2009) (Mincer 1978).

Questo lavoro contribuisce all'identificazione delle aree di destinazione dei flussi migratori in Italia, sviluppando un approccio combinato tra indicatori tradizionali e tecniche di geostatistica, partendo da variabili elementari (popolazione residente, iscritti e cancellati, presenza straniera), in modo da progettare una metodologia riproducibile.

È stato possibile individuare dei clusters territoriali di cui tener conto per decisioni politiche efficaci in tema di immigrazione e di problematiche di integrazione sociale, sicurezza, opportunità lavorative ad essa legate.

Lo studio ha riguardato il territorio nazionale, data la consistenza del fenomeno migratorio nel nostro territorio, soprattutto negli ultimi decenni in cui l'Italia ha vissuto un'inversione di tendenza, passando da origine a destinazione dei flussi migratori.

2. Dati e metodologia

La ricerca è stata sviluppata rispetto all'intero territorio nazionale assumendo la municipalità come entità statistica-geografica minima. Questa scelta metodologica risponde alla necessità di rappresentare nel dettaglio un fenomeno che tradizionalmente viene descritto a scala nazionale per macro aggregazioni.

Per le finalità del lavoro sono stati selezionati indicatori sintetici per misurare l'incidenza degli stranieri sulla popolazione residente in Italia. Tali indicatori sono stati costruiti sulla base di dati disponibili rispetto al caso in studio, ma anche reperibili nella maggior parte dei contesti nazionali.

Il dataset di riferimento è quello disponibile presso l'Istituto Nazionale di Statistica (ISTAT 1999, 2007). Le variabili di riferimento sono di tipo elementare: popolazione residente, popolazione straniera presente, numero di iscritti e cancellati. Per ciascuna variabile sono state considerate preliminarmente le serie storiche. In questo paper si riportano le elaborazioni relative agli anni 1999 e 2007 in quanto rappresentativi delle più recenti tendenze.

2.1. Il fenomeno migratorio: tecniche di analisi tradizionali

L'approccio metodologico adottato ha previsto inizialmente l'uso di indicatori statistici tradizionali. In particolare, sono stati calcolati l'Indice di Efficacia dei movimenti migratori (I_e), per ottenere informazioni sulla dinamica degli spostamenti, il Quoziente di Localizzazione (QL) e l'Indice di Dissimilarità (D), per ottenere misure di localizzazione e di segregazione.

Il quoziente di localizzazione misura l'intensità della modalità di un fenomeno relativo alla micropartizione territoriale rispetto all'intensità della stessa modalità riferita all'intera area cui appartiene.

L'Indice di Dissimilarità misura la distribuzione di due differenti gruppi etnici, pesandone la compatibilità (o incompatibilità) di localizzazione residenziale nell'ambito delle unità spaziali considerate.

In genere l'indice di segregazione è l'indice maggiormente usato nelle ricerche internazionali che riguardano la mobilità geografica, in particolar modo l'indice di dissimilarità sviluppato da Duncan e Duncan (Duncan e Duncan 1955, (a)). Esso permette di cogliere la divisione sociale presente nello spazio intra-urbano e quindi di risalire alle cause delle scelte localizzative che concorrono a favorire l'insediamento degli immigrati in una certa area piuttosto che in un'altra.

2.1.1. Indice di efficacia migratoria

L'indice di efficacia I_e ha permesso di rilevare il peso dei flussi migratori sulle variazioni demografiche. Esso è definito come:

$$I_e = \left[\frac{(I - D)}{(I + D)} \right] 100 \quad (1)$$

dove I sono gli iscritti all'anagrafe, D i cancellati; (I-D) rappresenta il saldo migratorio. Valori prossimi a zero indicano che l'interscambio migratorio di un comune non produce una variazione significativa della popolazione, valori prossimi a 100 segnalano che i flussi in ingresso sono molto maggiori rispetto a quelli in uscita e infine, valori prossimi a -100 mostrano che prevalgono i flussi di emigrazione.

2.1.2. Quoziente di localizzazione

Il quoziente di localizzazione (QL) fornisce una stima del grado di specializzazione della singola unità statistica ad accogliere la popolazione straniera.

Esso fornisce una misura della concentrazione relativa di un gruppo in un determinato comune, in relazione alla incidenza media dello stesso gruppo a livello nazionale.

L'indicatore è definito come:

$$QL = \frac{(x_i / y_i)}{(X / Y)} \quad (2)$$

dove x_i rappresenta il numero degli stranieri totali nell'area i-esima (il comune), X la popolazione straniera nell'intera area di studio (il territorio nazionale), y_i la popolazione residente nella area i-esima, Y la popolazione residente nell'intera area di studio.

Il quoziente di localizzazione fornisce la misura della concentrazione relativa di un gruppo in una determinato comune in relazione all'incidenza media dello stesso gruppo a livello nazionale. Un valore di QL inferiore a 1 indica la debole presenza di un gruppo in una certa area, un valore uguale a 1 indica che la concentrazione nel comune è simile a quella nazionale ed, infine, un valore superiore a 1 indica una situazione di sovrarappresentazione relativa di un gruppo in una certo comune (Cristaldi 2002).

2.1.3. Indice di dissimilarità

L'indice di dissimilarità nella formulazione di Duncan e Duncan (Duncan e Duncan 1955, (b)). fornisce una stima del grado di segregazione di due gruppi di popolazione presenti nell'area di studio. Esso descrive una concentrazione spaziale dei gruppi di popolazione.

L'indice è definito come:

$$D = \frac{1}{2} \sum_{i=1}^k |x_i - z_i| 100 \quad (3)$$

dove x_i rappresenta il rapporto tra il numero di residenti del gruppo nella zona i-esima e la popolazione complessiva di quel gruppo nell'intera area di studio, z_i rappresenta il rapporto tra il numero degli stranieri nella zona i-esima e il numero degli stranieri totali nell'area di studio, k è il numero delle zone in cui è divisa l'area di studio. L'indice varia tra 0 e 100. Valori prossimi a 0 indicano una bassa dissimilarità. Alti valori di questo indice indicano che la coesistenza di due gruppi nella stessa area è quantitativamente limitata.

Gli indici di segregazione, anche se ampiamente presenti in letteratura (Massey e Denton 1988), non forniscono indicazioni in merito alla distribuzione spaziale del fenomeno, hanno il limite di una forte dipendenza spaziale (Wong 1997) e producono risultati fortemente dipendenti dal sistema di zonizzazione scelto (O' Sullivan e Wong 2007).

2.2. Tecniche di analisi spaziale applicate allo studio del fenomeno migratorio.

L'identificazione di cluster rappresentativi della concentrazione territoriale della immigrazione è stata effettuata attraverso l'applicazione di tecniche di analisi spaziale. Sono state applicate al dataset misure di autocorrelazione spaziale tramite l'Indice I di Moran, il corrispondente scatterplot di Moran e il Local Indicator of Spatial Association (LISA).

Il primo indice fornisce una misura sintetica della possibile autocorrelazione spaziale (Moran 1948), lo scatterplot di Moran costituisce una tecnica di rappresentazione grafica delle relazioni spaziali, consentendo di investigare possibili raggruppamenti locali, il LISA permette di tener conto degli effetti locali del fenomeno (Anselin 1988; Anselin 1995).

2.2.1. Moran's I statistic

L'Indice di Moran (Moran 1948) può essere formalizzato come segue:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x})(x_j - \bar{x})w_{ij}}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (4)$$

dove x_i è la variabile osservata nelle n ripartizioni spaziali e \bar{x} è la media di tale variabile; w_{ij} è l'elemento generico della matrice di contiguità, esso denota l'effetto della unità territoriale j sulla unità i ; S_0 è la somma di tutti gli elementi della matrice di contiguità. La matrice di contiguità è una matrice quadrata di pesi, di ordine n (numero di zone distinte), generalmente simmetrica e rappresenta lo schema delle connessioni tra unità spaziali e delle loro intensità" (Badaloni e Vinci 1988). L'indice I può assumere valori che ricadono fuori del range $-1+1$. Inoltre, nel caso di autocorrelazione nulla (distribuzione casuale nello spazio) il suo valore non è 0 ma è pari a $-1/(n-1)$ (O'Sullivan, Unwin, 2002). Per cui se:

- $I < -1/(n-1)$ l'autocorrelazione è negativa
- $I = -1/(n-1)$ l'autocorrelazione è nulla
- $I > -1/(n-1)$ l'autocorrelazione è positiva

All'indice di Moran è possibile associare un utile grafico, il Moran Scatterplot, che fornisce informazioni complementari e integrative.

2.2.2. Moran Scatterplot

Il software GeoDa (Anselin 2003) permette di calcolare la I di Moran e di costruire il Moran Scatterplot che consente di analizzare l'autocorrelazione spaziale attraverso un grafico cartesiano in cui sull'asse delle ascisse si riporta la variabile x , cioè la variabile in esame espressa in deviazioni rispetto alla media, e su quello delle ordinate il ritardo spaziale di questa variabile.

In questo tipo di rappresentazione il primo ed il terzo quadrante rappresentano aree a correlazione positiva (High- High, Low - Low) mentre il secondo ed il quarto quadrante rappresentano aree a correlazione negativa.

La significatività della correlazione spaziale misurata attraverso Moran I e il Moran Scatterplot è fortemente dipendente dalla estensione dell'area di studio. Territori ampi, come quello italiano, possono nascondere patterns di diffusione spaziale al loro interno molto differenti. L'indice di Moran tende a mediare le variazioni locali all'interno di una unica misura di autocorrelazione.

L'indice di autocorrelazione locale LISA (Local Indicator of Spatial Association) consente, invece, di considerare effetti locali legati al fenomeno.

2.2.3. Indicatore locale di associazione spaziale(LISA)

LISA rappresenta una misura statistica in grado di soddisfare due condizioni: per ogni osservazione, tale indicatore misura il valore dell'aggregazione spaziale di valori simili attorno all'osservazione e la somma delle misure LISA per tutte le osservazioni è proporzionale ad un corrispondente indicatore globale di associazione spaziale (Anselin 1988; Anselin 1995).

Il valore dell'indice è calcolato secondo la formula:

$$I_j = \frac{\sum_j w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2} \quad (5)$$

Il valore dell'indice permette di valutare, per ciascuna localizzazione, la similitudine di ciascuna osservazione con quella del proprio intorno. Ne emergono in tal modo 5 scenari: localizzazioni con alti valori del fenomeno ed alto livello di similitudine con l'intorno (high - high), definiti "hot spot"; localizzazioni con bassi valori del fenomeno ed alto livello di similitudine con l'intorno (low - low), definiti "cold spot"; localizzazioni con alti valori del fenomeno e basso livello di similitudine con l'intorno (high - low), definiti potential "spatial outliers"; localizzazioni con bassi valori del fenomeno e basso livello di similitudine con l'intorno (low - high), definiti potential "spatial outliers"; localizzazione prive di significative autocorrelazioni.

Il LISA consente in modo efficace di associare ad ogni unità territoriale una misura del livello di associazione spaziale rispetto al suo intorno, permettendo di evidenziare il tipo di concentrazione spaziale per l'individuazione di cluster spaziali.

3. La distribuzione spaziale degli immigrati in Italia

Il caso di studio riguarda la distribuzione spaziale della presenza straniera in Italia. Sono stati calcolati gli indicatori descritti nei precedenti paragrafi assumendo la municipalità come unità statistica minima. Le elaborazioni si riferiscono ai dati più recenti: 1999, 2007.

La Fig. 1 mostra il valore dell'Indice di Efficacia dei movimenti migratori al 2007.

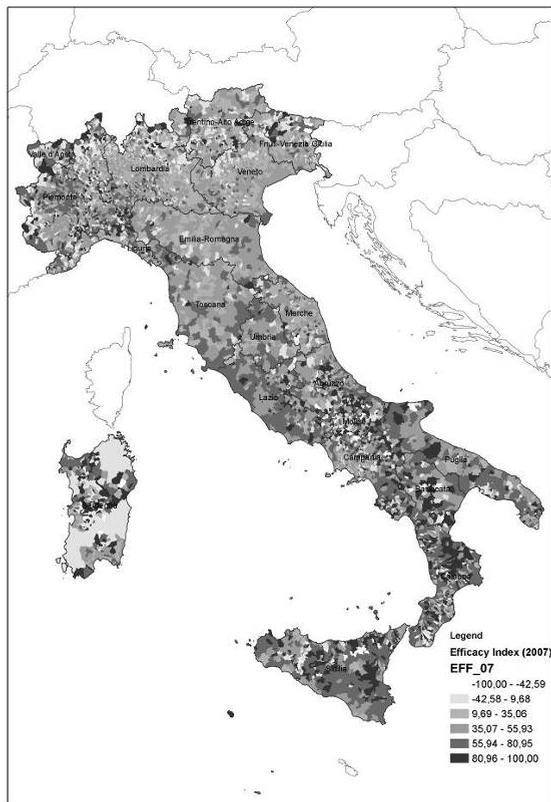


Fig. 1. Indice di Efficacia dei movimenti migratori al 2007

In Fig 1 è possibile osservare un comportamento eterogeneo del fenomeno migratorio tale da non consentire di identificare i cluster di origine e di destinazione dei flussi migratori. Si può osservare, inoltre, che i comuni montani hanno una marcata tendenza a generare migrazione, confermando la tendenza allo spopolamento di queste zone.

Attraverso la determinazione del Quoziente di Localizzazione, poi, è stato possibile stabilire che il maggior grado di specializzazione a ospitare la popolazione straniera riguarda le zone centrali e nord-orientali dell'Italia (Fig. 2). L'indice di dissimilarità ha permesso, invece, di misurare l'eterogeneità della struttura della popolazione straniera (Fig. 3).

Al fine di individuare cluster rappresentativi della concentrazione degli stranieri sono state utilizzate tecniche di analisi di autocorrelazione spaziale.

Il coefficiente I di Moran esprime la correlazione tra il numero degli stranieri/popolazione in un determinato luogo e numero degli stranieri/popolazione di unità spaziali vicine.

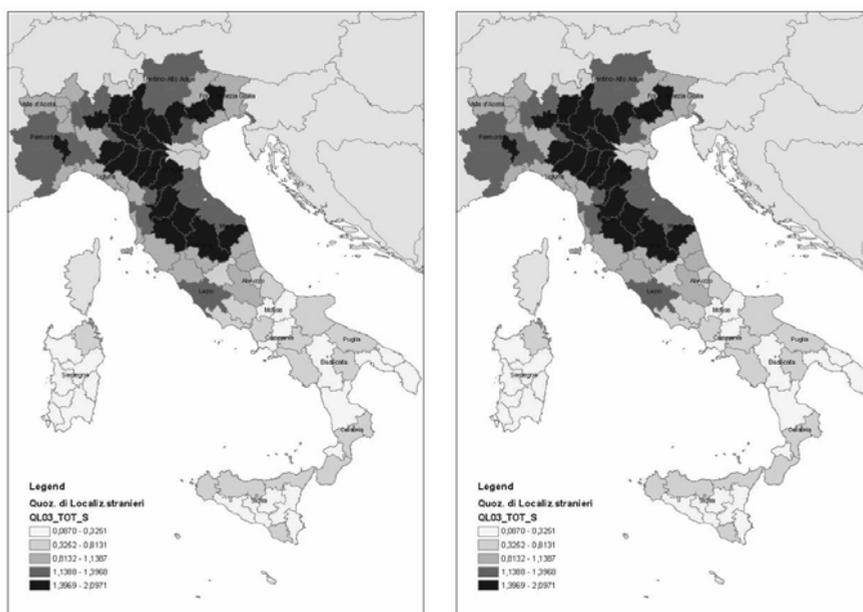


Fig. 2. Quoziente di Localizzazione al 2007 (a); Indice di Dissimilarità al 2007 (b)

Per tener conto delle connessioni e delle loro intensità, in questo studio, si è definita una matrice di contiguità W , dove $w_{ij} = 1$ se l'area i confina con l'area j , cioè se, nel nostro caso, due comuni sono confinanti, e $w_{ij} = 0$ negli altri casi. Si è scelto, come criterio di contiguità, la Queen Contiguity secondo la quale sono contigue quelle celle che nella griglia regolare hanno un lato ed un angolo in comune.

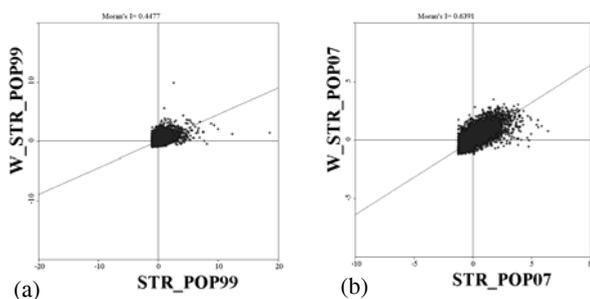


Fig. 3 Moran Scatterplot rispetto alla variabile Stranieri/Popolazione al 1999(a), 2007(b)

Alla I di Moran è stato associato lo scatterplot di Moran (Fig.3) relativamente ai dati del 1999 e del 2007. I risultati ottenuti tramite tale rappresentazione sono riportati, in ambiente GIS, su una mappa, Fig.4, dove si distinguono le singole municipalità rispetto ai gradi di

correlazione significativi: ‘High-Hight’ e ‘Low-Low’ (I e III quadrante dello scatterplot di Moran). Il risultato della rappresentazione è una clusterizzazione geografica strutturata essenzialmente su due raggruppamenti principali: uno che include i comuni del centro e Nord-Est dell’Italia (correlazione ‘High-Hight’) e uno che raggruppa i comuni dell’Italia Meridionale e insulare (correlazione ‘Low-Low’). Lo Scatterplot di Moran evidenzia, inoltre, i possibili comuni “anomali”, che rappresentano dli outliers e che richiedono, pertanto, un’adeguata investigazione.

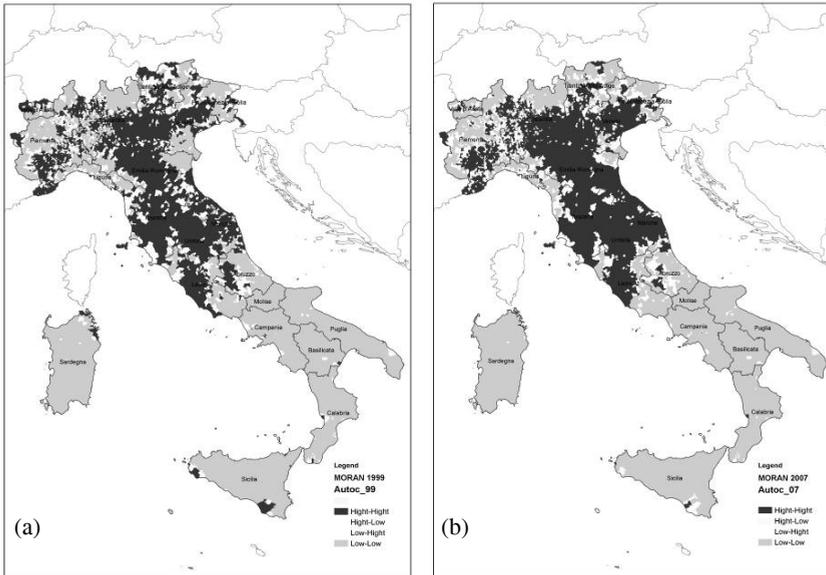


Fig. 4. Rappresentazione dei comuni rispetto alla distribuzione sullo scatterplot di Moran al 1999 (a) e al 2007 (b)

Il Moran Scatterplot non dà informazioni sulla significatività dei raggruppamenti spaziali, perciò è stato calcolato, con l’ausilio di GeoDa, l’indicatore LISA. Per il calcolo del Lisa si è fatto riferimento alla stessa matrice dei pesi costruita per la costruzione dello scatterplot di Moran. Esso, ha permesso di misurare, per ciascun comune, l’interdipendenza con gli altri comuni, di indicarne la tipologia (positiva o negativa) e la sua significatività. Si sono riportati i risultati ottenuti in ambiente GIS, ottenendo le mappe di Fig.6, dove si evidenziano i cluster relativamente alla concentrazione spaziale degli stranieri. Seppure con diversi livelli di significatività, emergono tre agglomerazioni: il primo cluster, si riferisce ad una relazione positiva di tipo ‘High-Hight’ e, geograficamente concentrato nella porzione nord-orientale, aree che si caratterizzano da livelli di benessere crescenti e quindi aree a forte richiamo per gli stranieri per la possibilità di occupazione che esse offrono, il secondo cluster, sempre di tipo ‘High-Hight’, interessa la parte centrale del territorio nazionale e trova una spiegazione ancora una volta negli elevati livelli di reddito ed occupazione, il terzo cluster, di tipo ‘Low-Low’, comprende i comuni dell’Italia Meridionale e Insulare, notoriamente caratterizzati da basso reddito e scarse opportunità occupazionali.

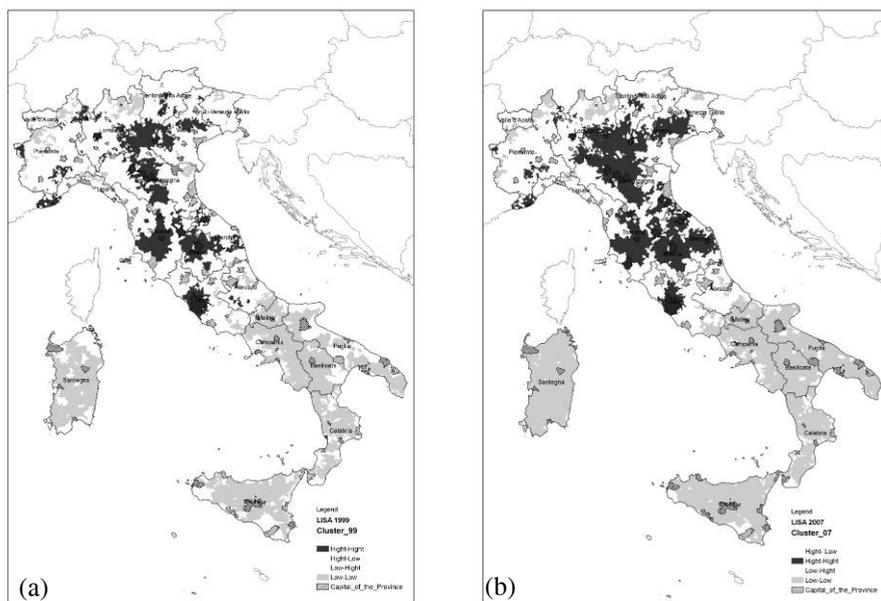


Fig. 6. Rappresentazione dei comuni ripetuto al LISA al 1999, (a) al 2007 (b)

Dal confronto del LISA alle due date considerate le aree che presentano correlazione del tipo Alto – Alto e Basso-Basso tendono ad espandersi interessando anche altri comuni limitrofi.

4. Conclusioni

Nel corso degli ultimi anni l'attenzione rivolta al fenomeno migratorio è diventata sempre maggiore in concomitanza con l'accrescimento della dimensione della presenza straniera dovuta all'intensificazione dei flussi migratori e al dibattito politico su diversi aspetti dell'immigrazione e della presenza straniera nel nostro Paese.

L'obiettivo del presente lavoro è stato quello di indagare sulla struttura spaziale della presenza di stranieri in Italia al fine di esaminare, tra le diverse interpretazioni presenti in letteratura, i divari geografici della presenza di stranieri. A tal fine, attraverso il ricorso alla Spatial Analysis, è stata verificata l'esistenza di autocorrelazione tra i comuni italiani relativamente agli anni 1999, 2007. I cluster così individuati si candidano a essere aree bersaglio di politiche settoriali volte a superare problemi di integrazione sociale, sicurezza e opportunità di lavoro.

Alla luce dei risultati emersi sembra plausibile affermare un'interpretazione di equilibrio dei divari regionali della migrazione: aree caratterizzate dalla stessa performance (alta presenza di stranieri o bassa presenza di stranieri) tendono ad aggregarsi ed, inoltre, effetti positivi nella presenza di stranieri in un comune si espandono ai comuni vicini, determinando

i movimenti degli stranieri dai comuni contigui a quelli nei quali l'effetto si è prodotto. Uno spostamento, dunque, degli stranieri da un'area all'altra esiste, ma non a lungo raggio: questo elemento caratterizzante ben si concilia con la storica distribuzione spaziale degli stranieri in Italia (divario nord-sud).

Bibliografia

- Anselin L (2003) *GeoDa 0.9 User's Guide*, Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics and CSISS, University of Illinois
- Anselin L (1988) *Spatial Econometrics: Methods and Models*, Boston, MA: Kluwer Academic
- Anselin L (1995) Local Indicators of Spatial Association-LISA. *Geographical Analysis* 27: 93-115
- Badaloni M, Vinci E (1988) Contributi all'Analisi dell'Autocorrelazione Spaziale, *Metron*, 46.
- Cristaldi F (2002) Multiethnic Rome: toward residential segregation?. *GeoJournal* 58: 81-90
- Duncan OD, Duncan B (1955) (a) A Methodological Analysis of Segregation Indexes. *American Sociological Review* 20: 210-217
- Duncan OD, Duncan B (1955) (b) Residential Distribution and Occupational Stratification". *American Journal of Sociology* 60: 493-503
- Massey DS, Denton NA (1988) The dimensions of residential segregation. *Social Forces* 67: 281-315
- Mincer J (1978) Family Migration Decisions. *Journal of Political Economy* 86:749 -773
- Moran PAP (1948) The interpretation of statistical map. *Journal the Royal Statistical society*,B, 243-251
- O'Sullivan D, Unwin DJ (2003) *Geographic Information Analysis*. John Wiley & Sons Chichester, New York
- O'Sullivan D, Wong DWS (2007) A Surface-Based Approach to Measuring Spatial Segregation. *Geographical Analysis* 39: 147-168Walker JR. Internal Migration. University of Wisconsin-Madison. http://www.ssc.wisc.edu/~walker/research/palgrave_6.pdf Last access 8/11/2009
- Sjaastad L (1962) The Costs and Returns of Human Migration. *Journal of Political Economy* 70: 80-89
- Wong DWS (1997) Spatial Dependency of Segregation Indices. *The Canadian Geographer* 41:128-136
- ISTAT, Rapporto Annuale. La Situazione del Paese al 2007 e al 1999