



SAPIENZA  
UNIVERSITÀ EDITRICE

ANNALI DEL DIPARTIMENTO DI METODI  
E MODELLI PER L'ECONOMIA  
IL TERRITORIO E LA FINANZA

2019

**Direttore Responsabile - Director**

Alessandra De Rose

**Direttore Scientifico - Editor in Chief**

Roberta Gemmiti

**Curatrici del numero - Managing Editors**

Adriana Conti Puorger, Cristina Giudici

**Comitato Scientifico - Editorial Board**

Maria Giuseppina Bruno (Sapienza Università di Roma)

Adriana Conti Puorger (Sapienza Università di Roma)

Alessandra Faggian (The Ohio State University)

Francesca Gargiulo (Sapienza Università di Roma)

Roberta Gemmiti (Sapienza Università di Roma)

Cristina Giudici (Sapienza Università di Roma)

Ersilia Incelli (Sapienza Università di Roma)

Antonella Leoncini Bartoli (Sapienza Università di Roma)

Isabella Santini (Sapienza Università di Roma)

Marco Teodori (Sapienza Università di Roma)

Catherine Wihtol de Wenden (CERI-Sciences Po-CNRS Paris).

Copyright © 2019

**Sapienza Università Editrice**

Piazzale Aldo Moro 5 – 00185 Roma

[www.editricesapienza.it](http://www.editricesapienza.it)

[editrice.sapienza@uniroma1.it](mailto:editrice.sapienza@uniroma1.it)

Iscrizione Registro Operatori Comunicazione n. 11420

ISSN: 2385-0825

Pubblicato a dicembre 2019



Quest'opera è distribuita  
con licenza Creative Commons 3.0  
diffusa in modalità *open access*.

# ZONIZZAZIONE E DIREZIONI PRIVILEGIATE: RIFLESSIONI GEOGRAFICHE A PARTIRE DAL FRIULI VENEZIA GIULIA

*Abstract:* The aim of this paper is to offer some empirical reflections on the construction of homogeneous regions through the use of classical clustering algorithms. After defining what a territorial analysis is and evaluating the different ways of conceiving and identifying a region, the article focuses on the empirical aspects of the automatic procedure of classification employed to seek homogeneity with respect to selected indicators. Starting from different zoning experiences developed through multivariate analysis methods, using as reference the Italian Friuli Venezia Giulia region, the paper will explain the decisions, which have a subjective and qualitative nature, that a researcher must take into account, such as the possibility of introducing the contiguity constraint in the analysis. Finally, the article will discuss the use of “direction indicators”, if these exist, to solve partially the issue of contiguity among territorial units. Some examples, relating to the refuelling across the Slovenian border and to the Regional Landscape Plan, will be proposed.

*Keywords:* analisi territoriale, indicatori socio-economici, analisi dei cluster, vincolo di contiguità.

## 1. Aspetti preliminari

Il presente lavoro vuole proporre alcune riflessioni, di natura prettamente qualitativa, sul problema della *zonizzazione*, o *riorganizzazione territoriale* secondo alcuni autori (Marra, 1990), su base regionale nel senso amministrativo del termine. Tali riflessioni saranno dedotte partendo da un coacervo di esperienze di ricerca condotte presso l’Università di Udine nell’arco degli ultimi 25 anni e riguardanti la regione Friuli Venezia Giulia (FVG). Per citare le più importanti, queste sono legate ai progetti: di costruzione di un modello per la programmazione energetica regionale degli anni Novanta (Gonano, 1994; Gonano e Zaccomer, 1994); di perimetrazione dei distretti industriali e di approfondimento sull’entrata in crisi di questo modello economico (Grandinetti, 1999; Zaccomer, 2006, 2008); di prima applicazione della legge regionale riguardante i prezzi alla pompa dei carburanti per autotrazione (Zaccomer, 2011) e successiva modifica (Zaccomer, 2012). Infine, sono riconducibili anche ad altre esperienze concernenti lo studio di svariati fenomeni socio-economici, come quelli sulla demografia delle imprese attive e sull’occupazione da esse generata, che hanno interessato e interessano tuttora il FVG, non da ultimo al fatto che la regione è stata di recente coinvolta in un processo di costruzione di un *Piano Paesaggistico Regionale* (PPR) che, per la prima volta nel panorama nazionale, ha previsto al suo interno anche la presenza di un’analisi socio-economica (Zaccomer, 2018a).

Tutte le esperienze appena citate sono sempre state portate avanti con l’idea di tener conto, in modo adeguato, del territorio all’interno dell’eventuale modellizzazione statistica e ognuna di esse, alla fine, ha sempre rimarcato la necessità di affilare gli strumenti quantitativi, com’è stato effettivamente fatto in alcune di esse, per catturare meglio la realtà dei fatti socio-economici e per leggere i risultati dei modelli stimati alla luce del più ampio ventaglio d’informazione qualitativa da essi traslasciato. A tal riguardo, il monito di Alvaro (2011, p. 229) risulta più che mai attuale: «L’errore non nasce quando, per quantificare la realtà, si è costretti a semplificarla; l’errore nasce quando, dopo averla semplificata, si è portati a ritenere che la realtà sia semplice».

Prima di introdurre le riflessioni dedotte su base empirica, pare necessario entrare nel merito di alcuni aspetti di carattere definitorio per introdurre meglio il contesto di riferimento.

### 1.1. Dall'analisi spaziale a quella territoriale

Agli inizi degli anni Ottanta, la modellizzazione statistica dei dati georeferenziati si trovava ancora in fase evolutiva, anche se erano già state gettate le basi per quella odierna<sup>1</sup>, ma non era così matura da aver assunto un ruolo indipendente all'interno del proprio settore disciplinare. La sua evoluzione ha portato a quel *corpus* teorico e di strumenti che oggi prende il nome di *statistica spaziale*, ossia quel ramo della statistica che tratta le manifestazioni dei fenomeni tenendo esplicitamente conto della loro posizione geografica e che trova applicazione in tutte quelle scienze dove lo spazio riveste un ruolo fondamentale. È innegabile affermare che al suo sviluppo hanno contribuito diverse discipline applicate quali quelle fisiche, matematiche, mediche, biologiche, ecc., comprese quelle idee innovative che, alcuni decenni prima, erano nate in seno al pensiero geografico, soprattutto nel suo versante economico, che oggigiorno vengono richiamate con il termine *new geography* o, nella sua traduzione italiana, *geografia quantitativa*, dove i modelli di Von Thünen, Weber e Christaller, riferiti rispettivamente al settore agricolo, industriale e dei servizi, ne costituiscono sia i precursori sia i capisaldi<sup>2</sup>.

Ricordando che nel linguaggio geografico i termini “spazio” e “territorio” hanno un significato ben diverso, non sarà sfuggito al lettore come in quegli anni non vi era ancora una netta distinzione tra l'analisi dei dati spaziali e quella portata a termine su dati territoriali, anzi vi era una certa tendenza a considerare la seconda soltanto come un caso particolare della prima.

Solo con i primissimi anni Novanta questa distinzione entra definitivamente nel linguaggio tecnico anche ufficiale. Infatti, nel 1990, a livello nazionale in ambito statistico pubblico<sup>3</sup>, Zani e Napolitano (1992) propongono di distinguere i due casi fino ad allora confusi: se il primo tipo di analisi è relativo agli studi quantitativi di manifestazioni puntuali o insistenti su griglie costruite dal ricercatore, siano esse regolari<sup>4</sup>, come quelle che utilizzano celle quadrate o esagonali, o irregolari come quelle ottenute dal diagramma di Voronoi o dalla triangolarizzazione di Delaunay (Bonfatti, 1988), il secondo è relativo ai soli studi basati su partizioni amministrative, o su porzioni di esse, già precostituite (Zani, 1993), in particolare su quelle di tipo NUTS<sup>5</sup> per cui, nel caso italiano, sull'usale suddivisione in regioni, province e comuni e, più di recente, anche in città metropolitane.

In linea con questa proposta, Conti Puorger (2018, p. 23) chiarisce che se «il dato si riferisce a un definito spazio o unità spaziale costruita *ad hoc*, si parla di serie spaziali e di statistica spaziale e lo scopo dell'analisi è principalmente quello di rilevare, nella struttura osservata del fenomeno, proprio l'incidenza, l'effetto del fattore spazio». Sull'altro versante, la «definizione di serie territoriali e statistica territoriale è riferita alla trattazione dei dati di unità preesistenti» per cui lo spazio, in senso lato, è lo scenario di riferimento, ma sono le relazioni territoriali, nella loro diversa forma e intensità, ad agire sul fenomeno oggetto di studio. Tali relazioni possono ancora essere modellizzate, al pari di quanto avviene nella statistica spaziale, attraverso la costruzioni di matrici geografiche che permettono anche di introitare espressamente nell'analisi la “prima legge della Geografia” di Tobler (1970) sulla correlazione delle “cose

<sup>1</sup> Lo stato dell'arte dell'epoca è ben rappresentato in Ripley (1981), Griffith (1988) e in Upton e Fingleton (1985, 1989) e nelle bibliografie, divise per argomenti, dei contributi presenti in Zani (1993) che fornisce una sorta di stato dell'arte dell'epoca.

<sup>2</sup> Cfr. Romagnoli (2002), pp. 24-25, mentre sul legame tra analisi spaziale e geografia quantitativa si veda l'introduzione della traduzione italiana, a cura di P. Landini e G. Massimi, del volume di Unwin (1986).

<sup>3</sup> Più precisamente durante la sessione su “Rappresentazioni grafiche spaziali e sistemi geografici informativi” delle Giornate di studio su “Avanzamenti metodologici e statistiche ufficiali” indette congiuntamente dall'Istituto Nazionale di Statistica e dalla Società Italiani di Statistica, Roma 13-14 dicembre 1990.

<sup>4</sup> La definizione proposta inizialmente prevedeva solo il caso di celle regolari, ma è evidente che la vera discriminante della definizione fornita dagli autori è la possibilità del ricercatore di poter costruire il reticolato, in funzione delle esigenze di studio, o di doverlo “subire” *tel quel*, scegliendo al massimo (dipende sempre dalla disponibilità dei dati) la scala geografica, o le scale geografiche, su cui condurre lo studio.

<sup>5</sup> Acronimo di *Nomenclature des Unités Territoriales Statistiques* che corrisponde alla partizione territoriale ufficiale a fini statistici adottata da tutti i paesi dell'Unione Europea.

vicine” che però ai tempi della globalizzazione, con l’importante riduzione dell’attrito delle distanze e della facilità di comunicazione telematica<sup>6</sup>, sembra essere sempre meno stringente.

Nel secondo caso, dal punto di vista geografico, sussiste un “primato del territorio” non solo nel senso della sua capacità di modificare pesantemente le manifestazioni dei fenomeni socio-economici che, perfino *ceteris paribus*, possono risultare diverse da territorio a territorio ma anche in quello di rendere, con il complesso insieme delle sue relazioni orizzontali e verticali, certe modellistiche proposte dalla letteratura statistica solo parzialmente capaci di fornire una completa lettura del fenomeno sotto osservazione. Un simile punto di vista è molto chiaro in ambito geografico ma non è certo comune tra gli statistici che, sicuramente negli anni Ottanta, ma ancora tutt’oggi considerano il territorio soprattutto come un vincolo, o peggio un’ipotesi, da inserire in un modello. Tanto meno lo è tra la gran parte gli economisti ortodossi che, per tradizione, hanno per lungo tempo trascurato, e continuano a sottovalutare, il ruolo del territorio all’interno della loro modellizzazione.

Un esempio perfettamente calzante di quanto appena affermato è costituito dall’evoluzione della *Shift-Share Analysis* che, nata negli anni Quaranta e sviluppatasi un ventennio più tardi in seno all’ambito economico – così come ben descritto in Haynes e Parajuli (2015) – è stata sin dall’inizio considerata uno strumento di analisi territoriale per il solo fatto lavorare con unità amministrative di tipo NUTS. Solo più recentemente, con la critica di Nazara e Hewings (2004), questa supposizione è stata messa in crisi sollevando la necessità di introdurre una matrice geografica, che esprima il tipo di struttura territoriale, all’interno del modello quantitativo al fine di tenere in considerazione anche gli effetti dovuti alle relazioni tra le unità territoriali. Ma anche gli autori che hanno seguito quest’idea, hanno spesso finito per adottare un particolare tipo di matrice meramente per ipotesi, in pratica senza una valida giustificazione *ex ante* e senza cercare di indagare *ex post* la robustezza dei dati ottenuti dai modelli, ossia senza provare a capire come questa ipotesi territoriale impatti sui risultati finali<sup>7</sup>.

Tornando alle analisi spaziali e territoriali, si ricorda che sono entrambe affette dal noto problema della dipendenza areale già sollevato negli anni Cinquanta da Kendall e Yule (1950, p. 313), la cui affermazione “our results depend on our units” è stata ripresa svariate volte in letteratura, e da Robinson (1956), ma successivamente approfondito dai lavori di Openshaw (1984) – che l’ha reso noto come *Modifiable Areal Unit Problem* (MAUP) –, di Openshaw e Taylor (1979) e di Arbia (1989). Tale problema getta un’ombra di soggettività sui risultati ottenuti. Secondo Tobler (1989, p. 178), la vera essenza del MAUP consiste nel fatto che le unità spaziali “can be aggregated in a immensely large number of different ways producing a distribution of results. Currently there is no technology for determining what might be considered the ‘best’ aggregation for any particular purpose”. Sono più recentemente nate alcune proposte che, per opportuni modelli, cercano di indirizzare il ricercatore nella scelta appropriata del livello di aggregazione, come in Arbia e Petrarca (2011). Focalizzando ora il problema dal punto di vista geografico il MAUP fa nascere l’esigenza di replicare le analisi territoriali, ovviamente quando possibile, per diversi livelli di scala, o in altri termini viene richiamata la necessità di un *approccio transcalare*<sup>8</sup> per verificare se, ad esempio, le evidenze ritrovate a livello comunale sono poi confermate a quello provinciale.

Sulla base di quanto appena detto, per terminare gli aspetti definitori, si rileva come il primato del territorio permetta di legittimare il passaggio terminologico da un’analisi spaziale, meramente statistica concentrata soprattutto alla modellizzazione del fenomeno tenendo conto del generico “fattore spazio”, a un’analisi territoriale a maggiore contenuto geografico, figlia di un moderno approccio transcalare, che miri non solo alla piena comprensione del fenomeno alla luce delle relazioni presenti sul territorio che lo supporta, ma anche all’integrazione delle informazioni non considerate direttamente dal modello statistico. Quest’ultima definizione rende evidente il salto metodologico rispetto alla definizione degli anni Novanta,

---

<sup>6</sup> Oggi, nel modo dell’*Internet of things*, grazie all’*additive manufacturing* anche gli oggetti (anche complessi come i nanosatelliti) possono essere concepiti, digitalizzati, spediti e riprodotti in un luogo diverso da quello dell’invio.

<sup>7</sup> Un primo tentativo di valutazione *ex post* degli effetti dovuti ai diversi metodi di costruzione delle matrici geografiche, all’interno della *shift-share analysis*, è stato portato avanti, ancora per il FVG, in Zaccomer e Mason (2008, 2011), mentre più recentemente in Zaccomer (2018b) si è cercato di valutarli *ex ante*, ossia unicamente sulla base delle geometrie costruite via GIS e, quindi, prima del loro inserimento in un qualsiasi modello statistico.

<sup>8</sup> Quanto meno nella sua versione debole, ossia di *multiscalarità*, così come definita da Bonaverò (2005, p. 5).

riscontrato anche nel caso della scomposizione *shift-share*, ossia di considerare un'analisi come territoriale per il solo fatto di lavorare su informazioni concernenti unità amministrative<sup>9</sup>.

## 1.2. Le possibili definizioni e delimitazioni di una regione

Senza voler entrare nell'ambito dell'articolata discussione sulle mutazioni subite dal significato attribuito al termine "regione" nell'arco dell'evoluzione del pensiero geografico, a cominciare dal suo periodo determinista, dal punto di vista concettuale vi sono alcune differenze da tenere sempre in considerazione poiché queste si riversano ineluttabilmente sulla fase operativa della sua individuazione.

Facendo riferimento alla classificazione di regione geografica fornita dal Dematteis (1989), un primo modo di delimitazione, caro alle discipline economiche e urbaniste, deriva direttamente da un'attività di pianificazione, per cui si parla anche di *regione di piano*, che può essere amministrata nel senso specificato dalla carta costituzionale, oppure può riguardare ambiti territorialmente più ristretti, come nel caso dei bacini idrografici, dove si mettono in atto particolari politiche o interventi. In via generale, trattasi di partizioni del territorio derivate dal sistema giuridico, per cui si parla anche di *regione istituzionale*, o che dovranno essere individuate *ex novo*; in quest'ultimo caso è consigliabile che questo tipo di regione sia individuato con criteri ben precisi, meglio se univoci, sempre in linea con gli obiettivi di pianificazione perseguiti.

Un secondo metodo di delimitazione parte dal concetto di *regione funzionale, gerarchica o a rete*, la cui individuazione avviene sulla base dell'interazione tra i suoi elementi, in particolare ricorrendo all'analisi dei loro flussi. Tali flussi possono essere di diversa natura, quindi non solo antropici (come i flussi migratori e quelli supportati dalle reti di trasporto, quali quelli generati dal pendolarismo casa-lavoro e dagli *shopping trip*), ma anche di merci (come esportazioni o importazioni) e servizi (si pensi alle cosiddette utilities legate all'erogazione di servizi quali acqua, energia elettrica e gas) o, ancora, di informazioni codificate che viaggiano a livello telematico (collegamenti telefonici e informatici di vario tipo compresi quelli che supportano l'attività monetaria e finanziaria).

Le analisi di questo tipo partono sempre dalla costruzione di una matrice quadrata, non per forza simmetrica, dei flussi e si differenziano sostanzialmente per il tipo di misura statistica quali quelle di centralità, autocontenimento e di intensità dell'interazione tra i suoi elementi (Martini in Zani, 1993, pp. 126-132). Validi esempi di questo tipo di regione sono le aree gravitazionali dei centri commerciali, i sistemi locali del lavoro (ISTAT, IRPET, 1989; ISTAT, 1997) e i distretti industriali da essi derivati (Grandinetti, 1999). Spesso in questi studi vengono utilizzati non solo misure statistiche, ma anche concetti della matematica o della ricerca operativa (ISTAT, 1992; Bachi, 1999), specialmente quando si stanno studiando le reti (Borruso, 2011). Uno degli esempi più noti nella letteratura geografica, relativo ai contatti telefonici tra diverse città, è il pionieristico lavoro di Nystuen e Dacey (1961) che hanno mostrato come sia possibile ricorrere alla Teoria dei grafi per lo studio di questo tipo di problematiche.

L'ultimo tipo di criterio è quello che ricerca l'omogeneità, ossia la somiglianza rispetto a una lista di caratteristiche palesata attraverso la selezione di una batteria di indicatori. In questo caso, l'idea è quella di una regione omogenea, dove l'omogeneità è qui intesa in un senso statistico più ampio rispetto all'accezione bidimensionale impiegata nell'ambito del possibilismo vidaliano. Gli strumenti quantitativi più utilizzati per delimitare questo tipo di regioni sono quelli che rientrano nel "cassetto" dell'analisi multivariata, in particolare la *Cluster Analysis*. In questo caso, le regioni individuate dipenderanno non solo dagli indicatori, ma anche da alcune decisioni che il ricercatore deve prendere in sede di analisi. Il suo vantaggio principale è che, una volta prese queste decisioni, si avvia un processo di costruzione territoriale trasparente, e soprattutto, ripetibile.

Tornando per un momento al caso della regione di piano, è importante che questa coincida o derivi da una regione funzionale o omogenea, sempre in funzione degli obiettivi del piano. Un valido esempio è

---

<sup>9</sup> La valorizzazione dell'analisi geografica, soprattutto dal punto di vista cartografico, all'interno della *shif-share analysis* come elemento imprescindibile che permetta di evitare di scegliere *ex ante* la matrice geografica, in qualità di mera ipotesi del modello, è stato proposto in Zaccomer e Grassetti (2014; 2016) sempre in riferimento al FVG.

costituito dagli Ambiti di Paesaggio del PPR del Friuli Venezia Giulia individuati sulla base di ampie riflessioni geo-morfologiche, storiche e socio-economiche che privilegiano ancora l'aspetto della contiguità areale, fattore che nella visione sistemica dell'attuale società dell'informazione non è più così stringente, così come messo in luce ancora da Dematteis (1989).

## 2. La zonizzazione multivariata tramite indicatori

Le riflessioni che ora si vogliono proporre concernono gli esercizi di zonizzazione che bisogna portare a termine quando una normativa, oppure un obiettivo di ricerca, richiede l'individuazione di aree omogenee a livello mesoregionale<sup>10</sup>. Questo esercizio, dal punto di vista statistico, viene generalmente risolto con l'individuazione di uno specifico raggruppamento di unità territoriali ottenuto sulla base di una *batteria di indicatori* che misurano, in modo indiretto, il fenomeno socio-economico sotto osservazione<sup>11</sup>. In via del tutto generale, la selezione degli indicatori avviene in conformità a una teoria assodata in letteratura, a una prassi scientifica consolidata oppure, caso comune quando si indagano fenomeni innovativi, a una decisione collegiale presa da una commissione di esperti nominata ad hoc.

Considerando il caso di una qualsiasi regione italiana (NUTS 2), suddivisa nei rispettivi comuni (LAU 2), il problema consiste nel trovare una diversa partizione, in conformità a quella amministrativa già esistente, costruita in modo da raggruppare comuni considerati omogenei dal punto di vista della batteria. È immediato constatare che un problema così posto, rispetto alle definizioni prima date, ricade appieno nell'ambito dell'analisi territoriale, in questo caso di natura socio-economica.

Come già anticipato, per dare soluzione a un simile problema, è prassi ricorrere alla *Cluster Analysis* (CLA), che è una metodologia multivariata che permette di raggruppare unità territoriali, in questo caso i comuni, in gruppi territoriali, ossia in aree omogenee più ampie, rispettando alcuni criteri impliciti nella scelta sulla tipologia di analisi. Infatti, i risultati che essa restituisce non dipendono solo dalla batteria di indicatori selezionata a monte per l'analisi, ma anche da alcune fondamentali scelte operate del ricercatore quali: 1) la tipologia di analisi, 2) la strategia e 3) la metrica utilizzata dall'algoritmo di aggregazione (per i primi tre punti si veda Anderberg, 1973; Gordon, 1999; Cerioli, Zani, 2007); 4) se prevista, l'implementazione nello stesso algoritmo di vincoli territoriali (Gordon, 1996; Zani, 1993).

Considerando la prima di queste scelte, ossia quale tipo di CLA utilizzare, l'elemento fondamentale che guida tale scelta è la conoscenza a priori del numero dei gruppi territoriali in quanto, ad esempio, già fissato da una normativa: in questo caso l'esercizio di riorganizzazione consiste solo nell'individuazione della loro struttura. In questo caso può essere utilizzata un'analisi di classificazione non gerarchica, che mira a restituire un'unica soluzione che rispetta un criterio di ottimalità per un prefissato numero di gruppi. La più nota di queste analisi si fonda sul *metodo delle k-medie* che, attraverso una procedura iterativa, mira a soddisfare un criterio di *coesione interna* (che si realizza quando le unità appartenenti a uno stesso gruppo sono il più possibile simili tra di loro) basato sulla devianza nei gruppi. Tale metodo non è però adatto per capire come, dal punto di vista geografico, i singoli comuni vanno ad agglomerarsi per formare il gruppo, ma soprattutto presenta un problema nel momento in cui la classificazione finale può essere influenzata dall'ordine in cui sono ordinati i comuni (in questo caso il codice ISTAT) nella matrice degli indicatori (Anderberg, 1973). L'effetto di questo problema sulla classificazione diventa di difficile valutazione soprattutto se esistono aree di sovrapposizione del fenomeno, ossia se questo non si comporta in modo netto sul territorio formando gruppi ben separati (MacQueen, 1967), così come emergerà fra poco rispetto al caso del FVG.

Generalmente, nella gran parte dei casi non ci si trova in una simile situazione per cui anche il numero di gruppi da utilizzare fa parte del problema da risolvere. In questa eventualità è consigliabile un'analisi gerarchica di tipo aggregativo poiché essa non fornisce solo una singola partizione, ma una gerarchia di partizioni tra cui scegliere (da tanti gruppi quanti sono i comuni a un gruppo composto da tutti i comuni)

---

<sup>10</sup> Le idee di partenza per questo articolo sono state presentate per la prima volta alla 7ma Giornata Italiana della Statistica organizzato dall'ISTAT il 26 ottobre 2017, presso la Camera di Commercio di Udine, mediante un contributo invitato dal titolo *La riorganizzazione territoriale su base regionale: aspetti teorici ed esperienze condotte nel Friuli Venezia Giulia*.

<sup>11</sup> In questo lavoro non si vuole entrare nel merito della definizione e della costruzione degli indicatori, e degli indici, per cui si rimanda direttamente all'ampia trattazione presente in Strassoldo, Mattioli e Schifini (1996).



la cui rappresentazione grafica viene chiamata dendrogramma (per la sua definizione formale si veda Cerioli, Zani, 2007). Sezionando il dendrogramma nel punto di maggiore salto è possibile individuare una soluzione ottimale che non solo soddisfa la citata coesione interna, ma anche la separazione esterna (che invece richiede che i gruppi devono essere i più possibili dissimili tra di loro). Da un punto prettamente geografico, il dendrogramma risulta fondamentale anche per capire, sulla base dei relativi cartogrammi, come si aggregano spazialmente le unità territoriali, aspetto di primaria importanza ai fini dell'interpretazione spaziale dei risultati ottenuti dalla CLA specialmente quando si voglia considerare una partizione sub-ottimale. Quest'ultima è la principale ragione geografica per cui, da ora in poi, si concentrerà l'attenzione su questa seconda tipologia di analisi.

Spostando ora l'attenzione sulla seconda scelta, ossia quella riguardante la strategia seguita dall'algoritmo di aggregazione, la letteratura propone diversi algoritmi da quello del legame singolo (o del vicino più prossimo) a quello del legame completo (o del vicino più lontano). Il primo di questi presenta il ben noto *effetto a catena*, che implica la tendenza di questa strategia a creare pochi gruppi iniziali a cui aggregare successivamente tutte le altre unità territoriali, rischiando sovente di compromettere la coesione interna dei gruppi così ottenuti; è però anche vero che si tratta dell'algoritmo che meno risente dell'eventuale correlazione esistente tra gli indicatori della batteria. Il legame completo invece non presenta l'effetto a catena, ma individua una serie di gruppi omogenei bilanciati dal punto di vista della loro numerosità, dimostrandosi un buon algoritmo soprattutto quando ci si trova a studiare fenomeni che si manifestano in modo netto, idealmente dicotomico, sul territorio senza presentare aree di sovrapposizione. Più ci si allontanerà da questo comportamento dicotomico, sempre più si formerà uno sbilanciamento nella numerosità dei gruppi. Inoltre, fatto più grave, a differenza del precedente, questo algoritmo risente della correlazione tra gli indicatori. Per cercare di ovviare a questi difetti, sono stati proposti diversi algoritmi alternativi, come quelli del legame medio e mediano e il metodo del centroide che forniscono risultati intermedi a quelli ottenuti dalle strategie "estreme" appena discusse, per i quali si rimanda direttamente alle monografie già segnalate oppure a Gordon (1987).

Rispetto a queste osservazioni teoriche, nella gran parte dei casi di studio che si riferiscono al Friuli Venezia Giulia sono state evidenziate tre tipologie di problematiche presenti simultaneamente: a) nessuna particolare distribuzione degli indicatori della batteria; b) presenza di correlazioni anche marcate tra gli stessi; c) presenza di aree di sovrapposizione dove il fenomeno indicato si comporta territorialmente in modo non netto. In questi casi è stato utilizzato, con una certa soddisfazione, l'algoritmo di Ward (1963) che, basandosi sulla minimizzazione dell'incremento della devianza totale ad ogni passo dell'algoritmo, non solo non risente dell'effetto a catena, garantendo quindi per costruzione la coesione interna dei gruppi, ma a differenza della strategia del legame completo produce gruppi equilibrati anche in presenza di aree di sovrapposizione (Marra 1990, pp. 171-175). Il "tallone d'Achille" nell'applicazione di questo algoritmo è legato al fatto che, a differenza del legame singolo, risulta fortemente sensibile alla correlazione tra indicatori.

Un modo per risolvere all'origine questo problema è quello di utilizzare *l'Analisi delle Componenti Principali* (ACP) come passo preliminare alla CLA. Infatti, le Componenti Principali non ruotate non sono altro che indicatori sintetici, ottenuti come combinazioni lineari degli indicatori della batteria, che per costruzione risultano non correlati (Cerioli, Zani, 2007). Questo non solo risolve il problema ma permette anche di legittimare la scelta della distanza euclidea, ossia la terza decisione che il ricercatore deve prendere, perché ora le dimensioni dello spazio matematico degli indicatori sintetici ottenuti sono ortogonali per costruzione. Se questa scelta viene legittimata in seno all'ACP, utilizzando il quadrato di tale distanza risulta molto più facile scegliere l'appropriato livello di taglio del dendrogramma, individuando quindi sia il numero ottimale sia la struttura dei gruppi che non solo garantisce la coesione interna, ma anche la separazione esterna.

Un simile modo di procedere viene confermato da un altro risultato teorico, ossia invece di utilizzare la distanza euclidea sulle componenti principali è sempre possibile utilizzare direttamente la distanza di Mahalanobis sugli indicatori della batteria poiché quest'ultima, per definizione, permette di tenere conto direttamente della correlazione tra gli stessi. Sadocchi (1987) dimostra formalmente che quest'ultima distanza applicata agli indicatori standardizzati, passaggio sempre necessario per via delle diverse unità di misura presenti nella batteria, porta agli stessi risultati del quadrato della distanza euclidea applicata alle componenti principali ottenute senza rotazione (ossia garantendo per costruzione la loro ortogonalità). In

altre parole, l'autore dimostra che seguire il procedimento prima consigliato equivale a tenere in considerazione la correlazione esistente tra gli indicatori della batteria, non a eliminarla in modo artificioso.

L'ultimo aspetto che qui si vuole discutere, forse il più immediato dal punto di vista delle caratteristiche dei risultati attesi, è quello di un'eventuale introduzione, "esplicita" secondo la terminologia di Duque et al., (2007), di vincoli territoriali all'interno dell'algoritmo di aggregazione. Questo è il tipico caso in cui viene richiesto dalla normativa di costruire gruppi composti da unità territoriali geograficamente contigue, nel senso di confinanti, tra di loro: il problema è tecnicamente risolvibile attraverso l'introduzione di un vincolo di contiguità nell'algoritmo di aggregazione che permette il raggruppamento di due unità territoriali, in questo caso i comuni, se e soltanto se questi confinano (Gordon, 1996; 1999). Purtroppo è ben noto che un ragionamento simile sconvolge a monte l'idea dell'analisi di tipo gerarchico poiché, per effetto del vincolo, si andranno ad aggregare comuni contigui meno omogenei di quelli non lo sono, in altre parole non viene più garantita la coesione interna dei gruppi, problema noto in letteratura come inversione delle distanze (Murtagh, 1985). Tale inversione non permette nemmeno di costruire il dendrogramma, quindi anche la possibilità di determinare il numero di gruppi viene meno. In altre parole, l'introduzione del vincolo riporta alle condizioni dell'analisi di tipo non gerarchico, ossia alla necessità di conoscere a priori il numero di gruppi. Inoltre, ci si deve sempre ricordare che i gruppi così ottenuti non sono più coesi al loro interno.

Proprio per ovviare a questi inconvenienti, sono fiorite in letteratura diverse proposte che oramai vanno di pari passo con la produzione del relativo software di classificazione automatica. Una recente rassegna di queste si trova in Chavent et al. (2017), che aggiorna quelle già citate, dove viene anche proposto un algoritmo Word-like vincolato che permette di separare l'omogeneità socio-economica degli indicatori dagli aspetti geografici attraverso l'opportuna selezione di un parametro, ma fissando a priori il numero di gruppi. Anche il ben noto software GeoDa di Luc Anselin, che ha permesso per anni un'agevole costruzione di alcune delle matrici geografiche più note sulla base delle geometrie GIS, propone un algoritmo basato sul metodo della k-medie di tipo vincolato, ma sempre per un numero di gruppi prefissato a priori. Per superare questo problema, e quindi per avvicinarsi alla caratteristica principale delle classificazioni gerarchiche, lo stesso software implementa la proposta di Duque et al. (2012) relativa al cosiddetto Max-p Regions Problem per il quale s'impone a priori che il raggruppamento delle unità territoriali porti a un numero massimo sconosciuto di regioni omogenee che però soddisfano un valore soglia minimo rispetto a un indicatore predefinito, come ad esempio la popolazione residente. Non si tratta di un algoritmo *prêt-à-utiliser* poiché la manualistica del programma (Anselin, 2017) mostra come, per arrivare alla soluzione definitiva, sia necessario portare a termine un'analisi di sensibilità che, come per molte procedure iterative di stima, si ottiene incrementando il numero di iterazioni dell'algoritmo oppure cambiando la soluzione iniziale.

Venendo invece alla legittimità di inserire un simile vincolo, non è possibile dare una risposta univoca, ma dipende dal problema scientifico studiato. Ad esempio, in campo ecologico la contiguità spaziale sembra essere imprescindibile per lo studio dei sistemi tanto che Legendre (1987) considera "artefatti" i risultati ottenuti senza l'introduzione del vincolo. Se invece spostiamo l'attenzione al problema geografico qui trattato, è possibile affermare che negli anni Settanta la posizione era ben espressa da Fischer (1980) che, richiamando la teoria generale dei sistemi, richiede la contiguità fisica delle unità spaziali costituenti la regione omogenea, ma le già citate parole del Dematteis (1989), alla fine degli anni Ottanta, vanno nel senso di ridurre l'importanza della contiguità fisica.

Volendo sintetizzare dal punto di vista prettamente operativo, si consiglia di utilizzare un simile vincolo quando vi sia una precisa necessità di ricerca o che vi sia un'esplicita richiesta del committente della ricerca che, informato delle problematiche, preferisca comunque privilegiare gli aspetti di gestione operativa magari sbarcandosi anche la decisione *ex ante* di quanti gruppi individuare attraverso la procedura di zonizzazione o di quale requisito minimo soddisfare in termini di indicatori.

A questo punto sembra necessario chiedersi se sia possibile ottenere una ripartizione territoriale che, almeno in via tendenziale, restituisca gruppi abbastanza contigui senza dover per forza inserire esplicitamente il vincolo di contiguità nella strategia seguita dall'algoritmo di aggregazione. In altre parole, come si può "aiutare" un algoritmo non vincolato a restituire gruppi i più possibili contigui al loro interno senza però esigere un ferreo rispetto di tale caratteristica geografica.

### 3. Le direzioni privilegiate del fenomeno socio-economico

Ricapitolando quanto fin qui discusso, il problema di zonizzazione trattato riguarda la possibilità di costruire gruppi di unità territoriali costruiti non solo nel pieno rispetto della coesione interna ma anche della separazione esterna, che siano ben bilanciati dal punto di vista della numerosità dei gruppi e costituiti da unità territoriali il più possibile contigue tra di loro. Questo perché l'analisi teorica appena condotta ha reso evidente come il pieno rispetto della coesione interna sia sostanzialmente incompatibile con la forzatura sulla contiguità territoriale delle unità coinvolte nella procedura di ricerca dei gruppi omogenei.

Senza ricorrere a metodi statistici più complessi e ancora in via di raffinazione, quali quelli già citati o ai metodi di classificazione con sovrapposizione e quelli sfuocati (Zani, 1993), l'esperienza empirica ha reso evidente come con l'algoritmo di Ward, in taluni casi e per alti livelli di aggregazione, sia comunque possibile ottenere gruppi sufficientemente contigui senza dover pagare il prezzo dell'introduzione di un vincolo. Questo poiché, in base alle caratteristiche del caso sotto esame, spesso è possibile introdurre nella batteria un indicatore che abbia la funzione di fornire la direzione privilegiata del fenomeno studiato, ovviamente nell'ipotesi che questo non sia per sua natura multidirezionale.

Per quanto riguarda il FVG, nel caso di fenomeni socio-economici, l'altitudine è sicuramente un indicatore legittimo poiché questo risulta sostanzialmente legato ad ogni attività antropica presente sul suo territorio, aspetto facilmente riscontrabile dal punto di vista empirico attraverso un'analisi delle correlazioni bivariate di Bravais-Pearson di tutti gli indicatori presenti nella batteria<sup>12</sup>. Questa semplice osservazione, valida anche per molte altre realtà territoriali, permette di fornire una soluzione empirica al problema qui posto poiché, per via della particolare configurazione morfologica, nel FVG l'altitudine cresce seguendo prevalentemente una direzione sud-nord (tranne che per alcuni comuni posti a sud-est, in particolare quelli del Carso). Questo permette di incamerare nell'analisi, in modo implicito per usare la terminologia di Duque et al., (2007), un principio di contiguità debole – che comporta comunque una maggiore coesione geografica dei gruppi – ma che non implicando vincoli non spezza la coesione socio-economica interna dei gruppi.

Chiaramente, l'altitudine non è l'unico indicatore di direzione privilegiata di un fenomeno che si può utilizzare. Nell'esperienza legata alla manovra sulla riduzione dei prezzi alla pompa dei carburanti regionali, rispetto a quelli sloveni, si è presentata la possibilità di utilizzare un altro indicatore con una direzione spaziale ben precisa, questa volta est-ovest, ossia la distanza tra ciascun comune e il più vicino punto di rifornimento di carburanti oltreconfine praticato dagli automobilisti residenti in FVG.

Per mostrare, con un'unica immagine cartografica, il funzionamento di entrambi gli indicatori di direzione si è proceduto a un'analisi dei cluster gerarchica, utilizzando il già discusso algoritmo di Ward, esclusivamente sulla distanza con il punto di rifornimento di carburanti oltreconfine, dichiarata dalle singole amministrazioni comunali<sup>13</sup>, e sull'altitudine comunale<sup>14</sup>. Il risultato ottenuto per cinque gruppi è riportato nella parte superiore della Figura 1: risulta graficamente evidente che, mentre la distanza dal punto di rifornimento oltreconfine ha prodotto la separazione rispetto alla direzione est-ovest, l'altitudine comunale ha prodotto la separazione dei gruppi rispetto alla direzione sud-nord.

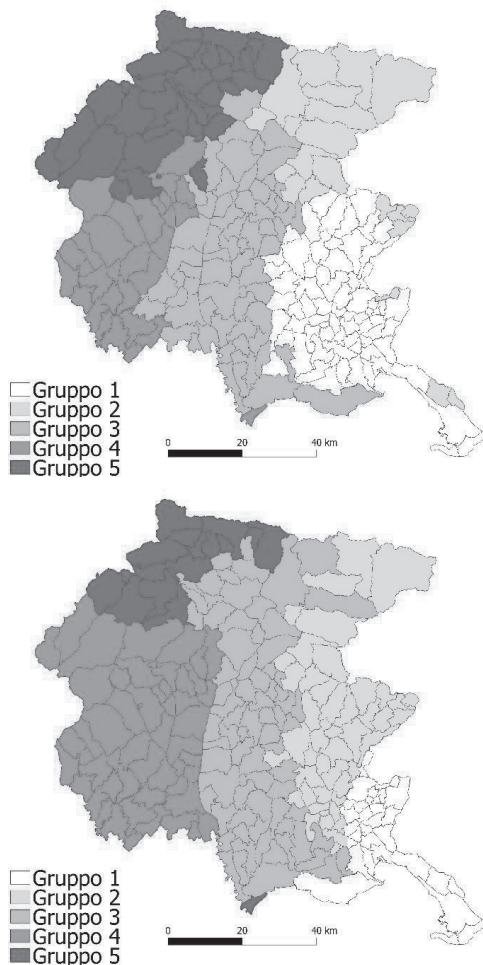
---

<sup>12</sup> La controprova di questo fatto si può ottenere calcolando le correlazioni parziali al netto degli effetti dell'altitudine stessa.

<sup>13</sup> Si tratta di una "distanza ufficiale" dichiarata nel 1997 che teneva in conto della possibilità di utilizzare valichi di prima e seconda categoria dove, per attraversare questi ultimi, non bastava un semplice documento di identità per attraversarli, ma era necessario anche uno specifico lasciapassare o *propusnica*.

<sup>14</sup> Per coerenza temporale con i dati sulla "distanza ufficiale", per altitudine si intende l'altezza sul livello medio del mare del Comune rilevata convenzionalmente in corrispondenza del Municipio in occasione del Censimento Generale della Popolazione del 1991.

Figura 1. Zonizzazioni del FVG tramite analisi dei cluster a cinque gruppi: a sinistra sulla distanza dal punto di rifornimento di carburanti oltreconfine e sull'altitudine; a destra sulla distanza dal punto di rifornimento di carburanti oltreconfine e sul tasso di evasione provinciale



Fonte: elaborazioni proprie su dati Reg. Aut. FVG, ISTAT e SWG (1997)

Il risultato appena dedotto graficamente viene confermato dalle medie per gruppo, di entrambe le variabili utilizzate, proposte in Tabella 1. Considerando solo le prime due colonne, si osserva subito come i gruppi 2 e 5, composti da comuni di montagna, presentano un'altitudine media mediamente più elevata dei gruppi posti a valle, mentre i gruppi 1 e 2 sono quelli che presentano una distanza media dal punto di rifornimenti oltreconfine, ossia in Slovenia, più bassa. Anche il coefficiente di correlazione complessivo risulta coerente, poiché pur essendo piuttosto basso in valore assoluto risulta positivo poiché nella parte nord-occidentale, così come emerge anche dall'osservazione dell'altitudine del gruppo 5, l'orografia del FVG presenta diversi rilievi di nota come il Monte Coglians che, con i suoi 2.780 metri, risulta essere la vetta regionale più elevata.

Tabella 1. *Altitudine media, distanza media e tasso di evasione mediano per gruppo*

Gruppi	Fig. 1 (sup.) - Distanza e altitudine		Fig. 1 (inf.) - Distanza e tassi di evasione		
	Altitudine media [m]	Distanza media [km]	Altitudine media [m]	Distanza media [km]	Tasso di evasione
1	61,0	16,2	67,9	9,5	36,8%
2	406,2	19,4	189,6	18,2	51,6%
3	117,8	45,1	168,9	39,8	51,6%
4	141,1	76,1	187,0	73,2	25,4%
5	632,4	69,4	681,7	67,4	51,6%

Fonte: elaborazioni proprie su dati Reg. Aut. FVG, ISTAT e SWG (1997)

Si tratta di un'analisi condotta unicamente per mettere in evidenza come i gruppi che si sono formati sono "naturalmente" contigui senza bisogno di alcun vincolo. È chiaro che tale soluzione non risulta perfetta, come quella che si sarebbe ottenuta con tale vincolo, così come risulta chiaro dalla presenza di alcuni comuni separati dal proprio gruppo di riferimento che, in realtà, sono per gran parte isolati a causa del diverso trattamento previsto dalla normativa regionale per i comuni sprovvisti di una stazione di rifornimento sul proprio territorio (Zaccomer, 2008). In questo caso però, ai soli fini gestionali, è possibile far transitare i pochi comuni "separati" nel gruppo contiguo che lo circonda, ossia si può agire successivamente alla formazione dei gruppi violando la coesione interna solo per questi pochi comuni. Si tratta quindi di un buon compromesso tra la necessaria correttezza scientifica e le esigenze di chi ha commissionato lo studio che dovrà gestire poi i gruppi territoriali individuati dalla procedura di zonizzazione.

Prima di procedere alla descrizione di un caso più complesso, ma sempre in linea con quello appena proposto, si vuole citare un'altra possibilità che si crea quando un indicatore è conosciuto solo per una partizione territoriale meno fitta, ad esempio per quella provinciale (ma potrebbe essere quella dei Sistemi Locali del Lavoro o quella delle Unioni Territoriali Intercomunali). Questo accade soprattutto quando le informazioni sono frutto di un campionamento la cui numerosità non è tale da fornire stime comunali attendibili. Le vie normalmente seguite in questo caso sono sostanzialmente due, ossia l'esclusione dalla batteria di questo indicatore oppure la stima del valore comunale utilizzando un qualche metodo di riparto, come quello sintetico proporzionale piuttosto utilizzato nella pratica corrente. Quest'ultima soluzione viene utilizzata quando l'informazione portata dall'indicatore provinciale è ritenuta imprescindibile per lo studio del fenomeno e quindi, in mancanza di una sua misura *proxy*, non può essere scartata. In via generale, il metodo di riparto, per quanto complesso sia, si basa sempre e comunque sull'informazione contenuta in altre variabili conosciute a livello più fine (tipicamente la popolazione o la superficie comunale) che non solo porta a un valore teorico, che in alcuni casi può essere ben diverso da quello reale, ma che, nel momento in cui l'indicatore "ripartito" viene inserito nell'ACP, l'intera operazione di stima rischia di perdere d'importanza soprattutto quando gli altri indicatori della batteria utilizzano anch'essi, per la loro normalizzazione, la medesima variabile impiegata per l'operazione di ripartizione. Inserendo invece l'indicatore direttamente a livello superiore, si evita questo problema e si tende a favorire l'unione di comuni che appartengono alla stessa provincia, quindi prevalentemente limitrofi.

Sempre nel caso dei prezzi dei carburanti, disponendo della stima dei tassi di evasione provinciali, determinati su base campionaria, e combinandoli con l'indicatore di direzione relativo alla distanza dal punto di riferimento oltreconfine si sono ottenuti i risultati proposti nella parte inferiore della Figura 1. Da questa si può osservare che la zonizzazione, a parità di gruppi ottenuti, è anch'essa sostanzialmente rispettosa della contiguità territoriale e rende ben evidente l'effetto di separazione del fiume Tagliamento<sup>15</sup> che separa, a sud, la provincia di Udine e quella di Pordenone (aspetto assolutamente non presente nella carta di sinistra della stessa Figura 1) dove, alla foce dello stesso corso d'acqua, rimane isolato dal suo gruppo di riferimento il solo comune di Lignano Sabbiadoro.

In sostanza, l'aver utilizzato l'informazione conosciuta solo a livello territoriale superiore non fa altro che spingere la classificazione a mantenere uniti i comuni della stessa provincia, così come accade per quelle di Pordenone (gruppo 4) e Gorizia e Trieste, ma l'effetto della distanza comunale è tale da spezzare la vasta provincia udinese in tre gruppi (2, 3 e 5). Ancora la Tabella 1 mette ben in evidenza come

<sup>15</sup> Il suo effetto è presto spiegato: gli automobilisti del pordenonese, potendo attraversare l'imponente corso d'acqua solo in pochi punti del suo cammino, vedono incrementare la propria distanza fino al punto di rifornimento oltreconfine. Tale aumento viene immediatamente intercettato dall'algoritmo di classificazione.

l'altitudine media dei gruppi 2 e 3 non sia troppo dissimile, ma quella del gruppo 5 rimane decisamente più elevata. Un'ultima osservazione: potrà sembrare anomalo che il tasso di evasione del primo gruppo, più vicino al confine sloveno, risulta inferiore a quello dei comuni dell'udinese. Questo risultato si spiega per il fatto che tali province erano, a suo tempo, già assoggettate a un diverso regime agevolato per i carburanti, ossia i residenti disponevano già di un contingente mensile di carburanti a un prezzo inferiore a quello oltreconfine. Solo terminato il contingente, gli automobilisti di questo gruppo indirizzavano il loro rifornimento oltreconfine.

Ovviamente quest'ultimo metodo di procedere è quello più "forzato", utilizzabile solo quando non vi è il medesimo dettaglio di conoscenza territoriale per tutti gli indicatori della batteria e non si voglia ricorrere a metodi di stima più complessi, come quello per piccole aree<sup>16</sup>, mentre quello praticabile nel caso dell'esistenza di una direzione privilegiata del fenomeno è sicuramente più elegante e formalmente ineccepibile qualora sulla batteria di indicatori venga applicata l'ACP.

Rimane un ultimo aspetto di non poco conto da chiarire: aumentando la dimensione della batteria, un unico indicatore di direzione del fenomeno può essere ancora utile? La risposta è sì, soprattutto quando molti degli indicatori ne sono fortemente influenzati e il criterio di massima separazione esterna porti a un numero di gruppi piuttosto limitato. Per verificare quanto appena detto, si propone di seguito lo studio basato su una variegata batteria di indicatori sociali<sup>17</sup>, relativa al periodo di benchmark statistico 2010-11, in cui è stata anche associata la variabile altimetrica media comunale, desunta dai dati dell'ISTAT, derivati da elaborazioni GIS sui dati del DEM dell'ISPRA (20×20 metri)<sup>18</sup>, per un totale di 15 indicatori. Come si osserva a sinistra della Figura 2, il criterio del massimo salto nel dendrogramma porta all'individuazione di due sostanziosi gruppi sufficientemente contigui senza la necessità di introdurre alcun vincolo. Dal punto di vista spaziale, la stessa figura mostra come la massima differenza si ha tra i comuni di montagna "socialmente più disagiati" e il resto del FVG<sup>19</sup>.

Come si può notare il gruppo dei comuni montani è diviso in due poiché la parte centrale della Carnia, con il capoluogo Tolmezzo e i suoi comuni limitrofi, non può essere classificata al pari degli altri comuni di montagna avendo caratteristiche sociali più simili a quelle riscontrate nell'area collinare e pianeggiante posta a valle. Naturalmente, nel momento in cui si desidera aumentare il numero dei gruppi, andando quindi a violare l'ipotesi di massima differenza tra i gruppi, l'aver inserito un solo indicatore di direzione del fenomeno può risultare insufficiente. Tagliando il dendrogramma a livello subottimale, in modo da ottenere tre gruppi, emergono nel gruppo 3 le diversità sociali delle maggiori realtà urbane del FVG (Pordenone, Udine, Gorizia, Monfalcone e Trieste) e la nota realtà balneare di Lignano Sabbiadoro, che spesso costituisce una realtà molto particolare all'interno dell'intero panorama regionale, così come si può facilmente verificare a destra della Figura 2.

Per quest'ultimo caso, la Tabella 2 indica il comportamento fortemente differenziato per singolo gruppo. Il gruppo 2 è sempre costituito dai comuni montani della Fig. 2, come si capisce immediatamente dall'altitudine media<sup>20</sup>, mentre gli altri gruppi sono costituiti dai comuni più a valle. Guardando a questi

---

<sup>16</sup> Si tratta di un coacervo di tecniche sviluppate con l'intento di superare i problemi di affidabilità statistica laddove i metodi classici falliscono. Si tratta soprattutto di un problema legato a una numerosità campionaria troppo bassa quindi non sorprende che lo stesso concetto di "piccola area" sia in realtà più un concetto statistico che geografico. L'ISTAT, anche in previsione del passaggio dal classico censimento decennale a quello "continuo", basato soprattutto su archivi amministrativi e indagini campionarie, ha recentemente sviluppato questa tematica in particolare per i dati delle rilevazioni sulla popolazione (Borrelli et al., 2012).

<sup>17</sup> Gli indicatori selezionati concernono: gli aspetti demografici (dalla densità abitativa, natalità, anziani per un bambino, vecchiaia alla presenza di stranieri); la disoccupazione; l'istruzione (dalla presenza di scuole dell'infanzia e primarie al possesso titoli universitari e terziari non universitari); la spesa sanitaria; la capacità di ospitare (espressa in termini di posti letto disponibili); la partecipazione socio-politica (dai volontari di istituzioni no-profit alla partecipazione al referendum del 2011 per il quesito sull'acqua); infine, la mobilità in generale (considerando la popolazione residente che si sposta giornalmente, per motivi sia di lavoro sia di studio, e il parco circolante ACI).

<sup>18</sup> Cfr. <https://www.istat.it/it/archivio/156224>, ultima visita: settembre 2019.

<sup>19</sup> Per completezza si segnala che a un risultato simile si perviene introducendo, accanto ai 15 indicatori già citati, altri 8 indicatori di carattere esclusivamente economico quali quelli relativi alla superficie agricola utilizzata, agli occupati in attività agricole, alle aziende con attività connesse, al carico zootecnico, agli addetti in attività industriali, alla densità industriale, agli addetti in attività terziarie e, infine, agli sportelli bancari. In questo caso vengono individuati i comuni di montagna "più disagiati" dal punto di vista socio-economico (Zaccomer, 2018a).

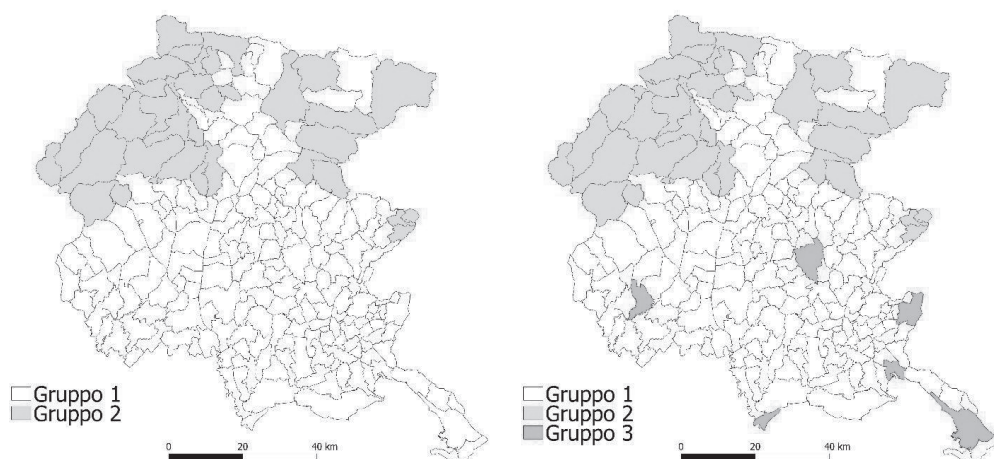
<sup>20</sup> Come si osserva, il valore medio dell'altitudine del gruppo 2 è maggiore di quelli riportati nella Tabella 1 per i gruppi 5. Questo non dipende solo dalla diversa composizione del gruppo, ma soprattutto dal cambio del metodo di rilevazione dell'altitudine di un



indicatori si può già far emergere il noto problema dello spopolamento montano – il confronto con la densità abitativa urbana è “drammatico” – e dell’elevata presenza di anziani (over 64 anni) per singolo bambino di età inferiore a sei anni, aspetto che poi si correla direttamente con il limitato possesso di titoli di studio post diploma. Infine, anche la presenza di stranieri è fortemente diversificata poiché, se questa è molto elevata nei centri urbani per la presenza di opportunità lavorative (in particolar modo a Pordenone e Monfalcone), nei comuni montani questa è legata alla presenza di case appartenenti a persone provenienti dall’estero (non solo stranieri come gli austriaci), ma anche di persone di origini friulane che, nel corso della loro vita, hanno preso una cittadinanza diversa e poi sono tornate in FVG.

Dal punto di vista scientifico questi risultati sono plausibili, poiché corrispondenti alla realtà dei fatti misurata attraverso la batteria di indicatori selezionata, ma dal punto di vista organizzativo solo un committente informato potrà valutare, sulla base dei risultati ottenuti, se sia necessario ripetere la zonizzazione ricorrendo all’introduzione di un vincolo di contiguità, forzando quindi l’inserimento del tolmezzino nell’area montana, oppure più semplicemente, come già suggerito per il caso dei carburanti, riclassificare “d’ufficio” i suoi comuni.

Figura 2. Zonizzazioni del FVG tramite analisi dei cluster su 14 indicatori sociali e sull’altitudine media: a sinistra per due gruppi (soluzione ottimale); a destra per tre gruppi (soluzione subottimale)



Fonte: elaborazioni proprie su dati Reg. Aut. FVG, ISTAT, MIUR e ACI (2010-11)

Tabella 2 Alcuni indicatori sociali per l’analisi a tre gruppi (soluzione subottimale)

Gruppi	Altitudine media [m]	Densità abitativa [residenti/km <sup>2</sup> ]	Numero anziani per bambino	Numero stranieri per 1000 residenti	Titoli post diploma (residenti > 19 anni)
1	221,4	179,4	4,7	57,2	7,3%
2	1075,2	14,4	10,4	10,4	4,3%
3	78,4	1342,1	5,1	117,2	14,1%

Fonte: elaborazioni proprie su dati Reg. Aut. FVG, ISTAT, MIUR e ACI (2010-11)

comune, prima unicamente misurata presso un punto di riferimento convenzionale quale il Municipio, attualmente attraverso una media di diverse misurazioni fatte su tutto l’estensione del suo territorio. È abbastanza palese come, per la gran parte dei comuni montani, il secondo tipo di misurazione restituisca un dato ben più elevato.

#### 4. Elementi di sintesi e conclusione

Il presente lavoro si è concentrato su alcune riflessioni di fondo relative alla costruzione di regioni omogenee, in senso statistico ossia rispetto all'informazione contenuta in una batteria di indicatori. Lo strumento più utilizzato in letteratura per questo tipo di operazioni è la *Cluster Analysis* che, prese alcune decisioni di partenza, viene considerato un valido strumento in quanto porta a una procedura di classificazione automatica trasparente e ripetibile, che garantisce quindi i requisiti di scientificità dell'operazione di zonizzazione. L'aspetto forse più delicato in questa procedura è quello iniziale, ossia la selezione degli indicatori, qualora non avvenga in conformità a una teoria, o a una prassi scientifica, va delegata ad una commissione di esperti nominata per l'occasione in modo da non lasciare a un singolo ricercatore un compito così delicato. Anche la decisione sul tipo di classificazione (gerarchica o non gerarchica), e gli aspetti tecnici che essa implica, porta spesso alla formazione di gruppi alquanto diversi. In tal caso è sempre consigliabile procedere a diverse analisi per testare la robustezza dei risultati ottenuti.

Entrando nell'ambito più strettamente empirico, è stato osservato come, nel caso del Friuli Venezia Giulia, l'utilizzo dell'algoritmo di Ward sulle componenti principali non ruotate ha permesso di ottenere buone partizioni del suo territorio basate su gruppi omogenei meglio bilanciati dal punto di vista dimensionale.

L'implementazione di un vincolo di contiguità nell'algoritmo di *clustering*, che implica l'introduzione nell'analisi di aspetti non direttamente considerati dalla batteria degli indicatori, deve sempre essere decisa in funzione delle esigenze di ricerca<sup>21</sup>, o in accordo con il committente dello studio, alla luce delle difficoltà che essa comporta. Qualora non vi sia la necessità di introdurre un simile vincolo, è stato dimostrato per via empirica come l'esistenza di direzioni privilegiate del fenomeno socio-economico può, sotto alcune condizioni, essere sfruttata per ottenere gruppi sufficientemente contigui. L'impiego dell'altitudine è una soluzione praticabile soprattutto qualora la configurazione morfologica della regione sotto osservazione risulti particolare come quella del FVG, ossia presenti ampie aree con altitudini fortemente differenziate lungo una specifica direzione geografica.

Le zonizzazioni ottenute sono solo il punto di partenza che, recuperando altre informazioni anche di tipo qualitativo, comprese quelle ottenute da indagini campionarie *ad hoc*, costituiscono un solido *parterre* per un'analisi più completa del fenomeno socio-economico sotto osservazione. Un'analisi che solo così può essere considerata a pieno titolo "territoriale" secondo la definizione qui accettata.

#### Riferimenti bibliografici

- ALVARO G. (2011), Il PIL: oltre il PIL. E oltre il PIL che c'è? *Moneta e Credito*, 64, 229-258.
- ANDERBERG M.R. (1973), *Cluster analysis for applications*, Academic Press, New York.
- ANSELIN L. (2017), *Cluster Analysis. Spatially Constrained Clustering Methods*, reperibile all'URL [http://geodacenter.github.io/workbook/8\\_spatial\\_clusters/lab8.html](http://geodacenter.github.io/workbook/8_spatial_clusters/lab8.html), ultima visita: settembre 2019.
- ARBIA G. (1989), *Spatial data configuration in statistical analysis of regional economic and related problems*, Advanced studies in theoretical and applied econometrics, 14, Kluwer Academic Publisher, Dordrecht.
- ARBIA G., PETRARCA F. (2011), *Effects of MAUP on spatial econometric models*, Letters in Spatial and Resource Sciences, 4, pp. 173-185.
- BACHI R. (1999), *New methods of geostatistical analysis and graphical presentation. Distributions of populations over territories*, Kluwer Academic/Plenum Publishers, New York.
- BONAVERO P. (2005), *L'approccio transcalare come prospettiva di analisi. Il contributo della geografia alla ricerca economica e sociale*, Quaderni dell'Istituto di Studi su Popolazione e Territorio, Serie metodologica, 3, ISU Università Cattolica, Milano.
- BONFATTI F. (a cura di) (1988), *Elaborazione automatica dei dati geografici. Strumenti per la realizzazione dei sistemi informativi territoriali*, Masson, Milano.

---

<sup>21</sup> Se, per esempio, si vuole classificare in modo diverso i comuni montani da quelli di pianura, le città dal loro hinterland, le aree rurali da quelli industriali, allora l'utilizzo di un vincolo di contiguità può risultare controproducente poiché, forzando il raggruppamento di comuni vicini, ne compromette di fatto la loro somiglianza.



- BORRELLI F., CARBONETTI G., DE FELICI L., SOLARI F. (2012), *Metodologie di stima per piccole aree applicabili a variabili di censimento*, ISTAT Working Papers, 3, Roma.
- BORRUSO (2011), *Geografie di rete. Infrastrutture, regioni, città*, Pàtron editore, Bologna.
- CERIOLO A., ZANI S. (2007), *Analisi dei dati e data mining per le decisioni aziendali*, Giuffrè, Milano.
- CHAVENT M., KUENTZ-SIMONET V., LABENNE A., SARACCO J. (2017), *ClustGeo: an R package for hierarchical clustering with spatial constraints*, reperibile all'URL <https://arxiv.org/abs/1707.03897>, ultima visita: settembre 2019.
- CONTI PUORGER A. (2018), *Analisi geografica e metodi statistici. Una introduzione*, Pàtron, Bologna.
- DEMATTEIS G. (1989), Regioni geografiche, articolazione territoriale degli interessi e regioni istituzionali, *Stato e mercato*, 27(3), 445-467.
- DUQUE J.C., ANSELIN L., REY S.J. (2012), The max-p regions problem, *Journal of Regional Science*, 52 (3), 397-419.
- DUQUE J.C., RAMOS R., SURIÑACH J. (2007), Supervised regionalization methods: a survey, *International Regional Science Review*, 30 (3), 195-220.
- FISCHER M.M. (1980), Regional taxonomy. A comparison of some hierarchic and non-hierarchic strategies, *Regional Sciences and Urban Economics*, 10, 503-537.
- GONANO G. (1994), *Il bilancio energetico regionale: alcune riflessioni metodologiche*, Quaderni del Dipartimento di Economia, Società e Territorio, 7, Università degli Studi di Udine.
- GONANO G., ZACCOMER G.P. (1994), *La costruzione degli indicatori energetici: fonti e metodi statistici*, Quaderni del Dipartimento di Economia, Società e Territorio, 4, Università degli Studi di Udine.
- GORDON A.D. (1987), A review of hierarchical classification, *Journal of the Royal Statistical Society, Series A*, 150 (2), 119-137.
- GORDON A.D. (1996), A survey of constrained classification, *Computational Statistics & Data Analysis*, 21 (1), 17-29.
- GORDON A.D. (1999), *Classification*, Monographs on Statistics & Applied Probability, 82, Chapman & Hall/CRC, Boca Roton (US), 1999.
- GRANDINETTI R. (a cura di) (1999), *Il seggiolaio e l'economia globale. La transizione evolutiva del distretto friulano della sedia attraverso i risultati di una indagine sul campo*, Cedam, Padova.
- GRIFFITH D.A. (1988), *Advanced spatial statistics*, Kluwer, Dordrecht.
- HAYNES K.E., PARAJULI J. (2015), Shift-share and its new extensions, in KARLSSON C., ANDERSSON M., NORMAN T. (eds.), *Handbook of research methods and applications in economic geography*, Edward Elgar Pub., Cheltenham, pp. 83-117.
- ISTAT (1992), *Atlante statistico italiano. Analisi geostatistica*, secondo volume, Roma.
- ISTAT (1997), *I sistemi locali del lavoro 1991*, Argomenti, 10, Roma.
- ISTAT-IRPET (1989), *I mercati locali del lavoro in Italia*, Franco Angeli, Milano.
- KENDALL M.G., YULE G.U. (1950), *An introduction to the theory of statistics*, Griffith, London.
- LEGENDRE P. (1987), Constrained Clustering, in LEGENDRE P., LEGENDRE L. (eds.), *Developments in numerical ecology*, NATO ASI Series, 14, Springer-Verlag, Berlin, pp. 289-307.
- MACQUEEN J. (1967), Some methods for classification and analysis of multivariate observations, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1: Statistics, University of California Press, pp. 281-297.
- MARRA E. (1990), Metodi per la riorganizzazione territoriale e cluster analysis, in GUALA C., MARRA E. (eds.) *Indicatori sociali e territorio*, Sagep, Genova.
- MURTAGH F. (1985), A survey of algorithms for contiguity-constrained. Clustering and Related Problems, *The Computer Journal*, 28 (1), 82-88.
- NAZARA S. E HEWINGS G.J.D. (2004), Spatial structure and taxonomy of decomposition in shift-share analysis, *Growth & Change*, 35, 476-490.
- NYSTUEN J.D. E DACEY F. (1961), A graph theory interpretation of nodal regions, *Papers and proceedings of the Regional Science Association*, 7, 29-42.
- OPENSHAW S. (1984), *The Modifiable Areal Unit Problem*, Concepts and Techniques in Modern Geography, 38, (Norwich (UK)).

- OPENSHAW S., Taylor P.J. (1979), A million or so of correlation coefficients: three experiments on the modifiable areal unit problem, in WRIGLEY N. (ed.) *Statistical methods in the spatial sciences*, Pion, London, 127-144.
- RIPLEY B.D. (1981), *Spatial Statistics*, J. Wiley and sons, New York.
- ROBINSON A.H. (1956), The necessity of weighting values in correlation of areal data, *Annals of the Association of American Geographers*, 46 (2), 233-236.
- ROMAGNOLI L. (2002), *Metodi statistici elementari per la geografia. Strumenti per l'analisi dei dati, il confronto quantitativo, lo studio e l'interpretazione delle relazioni spaziali con l'ausilio del calcolatore elettronico*, Pàtron, Bologna.
- SADOCCHI S. (1987), *Manuale di analisi statistica multivariata per le scienze sociali*, Angeli, Milano.
- STRASSOLDO M., MATTIOLI E., SCHIFINI S. (1996), *Teoria dei numeri indici dei prezzi e degli indicatori economici, finanziari e sociali*, Cedam, Padova.
- SWG (1997), *Il consumo di benzina in Friuli Venezia Giulia*, rapporto interno commissionato dalla Reg. Aut. Friuli Venezia Giulia, 25 febbraio 1997, Trieste.
- TOBLER W.R. (1970), A computer movie simulation urban growth in the Detroit region, *Economic Geography*, 46, 234-240.
- TOBLER W.R. (1989), Frame independent spatial analysis, in GOODCHILD M., GOPAL S. (eds.), *The accuracy of spatial databases*, Taylor & Francis, London, 76-79.
- UNWIN D. (1986), *Analisi spaziale. Un'introduzione geocartografica*, F. Angeli, Milano.
- UPTON G.J.G, FINGLETON B. (1985), *Spatial data analysis by example. Point pattern and quantitative data*, vol. I, Wiley and sons, New York.
- UPTON G.J.G, FINGLETON B. (1989), *Spatial data analysis by example. Categorical and directional data*, vol. II, Wiley and sons, New York.
- WARD J.H. (1963), Hierarchical grouping to optimize an objective function, *Journal of American Statistical Association*, 58, 236-244.
- ZACCOMER G.P. (2006), Shift-Share Analysis with Spatial Structure: an Application to Italian Industrial Districts, *Transition Studies Review*, 42, 213-227.
- ZACCOMER G.P. (2008), *Economia, Statistica e Territorio. Informazione e metodologia statistica per la conoscenza dell'economia del Friuli Venezia Giulia*, Forum, Udine.
- ZACCOMER G.P. (2011), *Carburanti, statistiche e prezzi. Esperienze di ricerca legate alla manovra di riduzione dei prezzi delle benzine e del gasolio per autotrazione in Friuli Venezia Giulia*, Forum, Udine.
- ZACCOMER G.P. (2012), La manovra di riduzione dei prezzi dei carburanti in Friuli Venezia Giulia: un quadro di sintesi dal 1997 al 2012, *Rivista di Economia e Statistica del Territorio*, 2, 34-59.
- ZACCOMER G.P. (2018a), *L'analisi territoriale socio-economica in ambito paesaggistico. Gli indicatori compositi per la zonizzazione territoriale del Friuli Venezia Giulia*, Forum, Udine.
- ZACCOMER G.P. (2018b), La costruzione di una misura caratteristica della matrice dei pesi spaziali: un confronto empirico sulle geometrie del Friuli Venezia Giulia, in DE TONI F. (a cura di), *Scritti in onore di Marzio Strassoldo*, Forum, Udine, 139-149.
- ZACCOMER G.P., GRASSETTI 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, 56-72.
- ZACCOMER G.P., GRASSETTI L. (2016), L'impiego di algoritmi AMOEBA per lo studio delle variazioni temporali di un fenomeno economico: prime evidenze generalizzabili da un caso di studio, *Bollettino dell'Associazione Italiana di Cartografia*, 158, 131-144.
- ZACCOMER G.P., MASON P. (2008), A definition of neighborhoods based on local labor systems: a regional application on employment data, *Statistica & Applicazioni*, 6 (2), 215-236.
- ZACCOMER G.P., 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, 329-356.
- ZANI S. (1993), a cura di, *Metodi statistici per le analisi territoriali*, F. Angeli, Milano.
- ZANI S., NAPOLITANO P. (1992), Problemi di rilevazione e di presentazione dei dati spaziali, in ISTAT, SIS, *Avanzamenti metodologici e statistiche ufficiali: atti delle prime giornate di studio Roma, 13-14 dicembre 1990*, Roma.