

Le performance sanitarie regionali durante l'emergenza Covid-19: una comparazione tra serie storiche

eyesreg.it/2020/le-performance-sanitarie-regionali-durante-lemergenza-covid-19-una-comparazione-tra-serie-storiche/

di: *Lucio Palazzo e Riccardo Ievoli*

EyesReg, Vol. 10, N. 5, Settembre 2020

Introduzione

A cinque mesi dalla sua identificazione (Luglio 2020) è possibile supporre con un certo grado di fiducia che la prima “ondata” della pandemia di COVID-19 in Italia sia quantomeno nelle sue fasi conclusive. Tale considerazione non è irrealistica sia sulla base dei dati relativi alla stima dell'indice dinamico nazionale di contagiosità, che ormai si attesta su livelli considerati accettabili (così come indicato dai report settimanali del Ministero della Salute), sia osservando l'andamento del numero dei decessi giornalieri unito alla drastica riduzione del numero di pazienti in terapia intensiva (meno di 100) e negli altri reparti preposti (circa 1.000).

Nonostante ciò, il nostro paese è certamente un “osservato speciale” ed è stato oggetto di numerosi studi scientifici a livello epidemiologico (Boccia, Ricciardi e Ioannidis, 2020 tra gli altri). Ad oggi si osservano dinamiche di squilibrio territoriale non necessariamente guidate dalle note disparità di tipo economico ed infrastrutturale (come sottolineato di recente anche da Musolino e Rizzi, 2020), e nemmeno riconducibili solamente ai flussi di propagazione del virus stesso.

Obiettivi e metodologia del lavoro

Da tali riflessioni scaturisce la seguente domanda di ricerca: al di là del dibattito pubblico generato dall'emergenza sanitaria e al netto delle diverse caratteristiche delle strutture sanitarie (e modelli di *governance*) esistono davvero dei pattern regionali sostanziali nei modelli evolutivi nella diffusione della pandemia, tenendo conto delle differenze di popolazione residente?

Pertanto lo scopo del presente lavoro è quello di indagare come e in che modo si siano manifestati gli squilibri sanitari a livello territoriale, andando ad identificare caratteristiche comuni in termini di incidenza del virus. Tale analisi viene effettuata utilizzando le serie storiche relative ai ricoveri in terapia intensiva e negli altri reparti ospedalieri per le 19 regioni italiane e le due province autonome di Trento e Bolzano. Tali serie rappresentano, a livello regionale, una diretta espressione non solo della diffusione del virus, ma anche dell'impatto della malattia stessa sul sistema sanitario.

La possibilità di utilizzare modelli per serie storiche durante la fase acuta della pandemia è già stata esplorata da autori come Agosto e Giudici (2020), per le dinamiche del contagio, e da Farcomeni et al. (2020), per la previsione a breve termine dei ricoveri nelle

terapie intensive. E autori quali James e Menzies (2020) e Rojas, Rojas e Valenzuela (2020) hanno effettuato comparazioni di serie storiche multivariate concernenti i casi positivi e i decessi per sottoinsiemi di unità territoriali.

Dopo una prima analisi descrittiva riguardante l'impatto della pandemia in Italia a livello territoriale, per perseguire lo scopo del lavoro si è deciso di effettuare una comparazione basata su indici di distanza tra le serie storiche regionali, allo scopo di corroborare quanto osservato nelle analisi empiriche preliminari. La similarità tra le serie degli ospedalizzati e delle terapie intensive di ciascuna regione è stata misurata attraverso metodi già noti in letteratura in ambiti quali l'analisi delle serie storiche finanziarie o l'analisi delle frequenze per i riconoscimenti vocali.

Per effettuare tale confronto si utilizza la metrica denominata *dynamic time warping* (DTW), introdotta da Berndt e Clifford (1994) per la misurazione della distanza tra due serie storiche attraverso il confronto della forma e dell'andamento dei dati rispetto al tempo. Questo metodo appartiene al cosiddetto gruppo delle distanze "*shaped-based*", costruite sulla "forma" delle serie. Come *benchmark* verranno riportati anche i risultati ottenuti dall'adozione di una diversa metrica, quella *di Fourier* (si vedano, tra gli altri, Agrawal, Faloutsos, e Swami, 1993), che, appartenente alla classe delle distanze "*feature-based*", permette di confrontare la similarità tra due sequenze temporali dopo averle convertite in una combinazione di elementi strutturali delle serie (trend, ciclo, ecc.).

Entrambi gli approcci assicurano una flessibilità tale da poter essere applicati agli andamenti regionali preservandone la metrica; la differenza principale è che il primo metodo dipende fortemente dalla posizione e dal numero dei nodi caratteristici (quali ad esempio punti di massimo/minimo e flessi) mentre nel secondo caso la trasformazione di Fourier tende a confrontare le componenti armoniche attraverso le serie "filtrate". Inoltre, allo scopo di raggruppare unità territoriali simili, viene utilizzata la classificazione gerarchica basata sulla distanza di Ward (per maggiori dettagli si veda la *review* di Murtagh e Contreras, 2012) considerando un numero di *cluster* pari a quattro.

Le serie storiche regionali a confronto

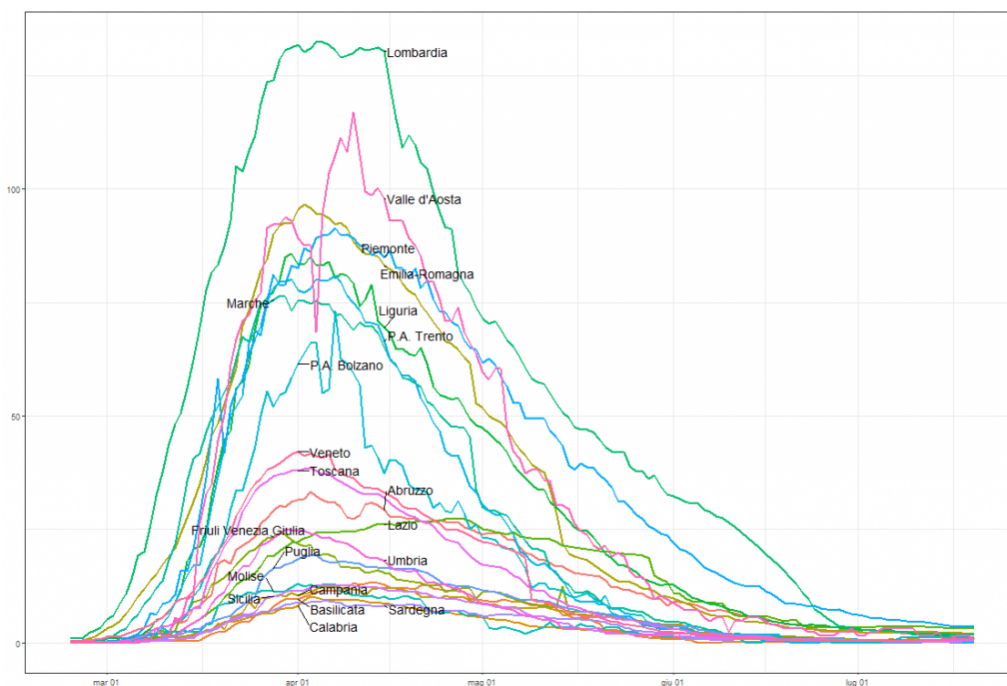
I dati delle serie storiche regionali giornaliere relative a 1) il numero di pazienti ricoverati nei reparti di terapia intensiva e 2) il numero di pazienti ricoverati negli altri reparti sono stati ottenuti attraverso il sito ufficiale della protezione civile e successivamente normalizzati in base alla popolazione di ciascuna unità territoriale. Il campione presenta una dimensionalità pari a 134 giorni che vanno dal 24/02 al 6/07 del 2020 e va tenuto in conto che l'andamento delle curve è dipeso anche da fattori esogeni, quali l'attuazione delle misure restrittive introdotte dal Governo (Sebastiani, Massa e Riboli, 2020). Attraverso le figure 1 e 2 è possibile osservare l'andamento dell'epidemia per quanto riguarda le terapie intensive e i pazienti ospedalizzati nelle regioni italiane, normalizzati rispetto al numero di abitanti e moltiplicati per 100 mila. Nonostante il picco endemico appaia in una fascia temporale simile per tutte le unità territoriali, si nota una differenza sostanziale in termini di variabilità e di impatto del fenomeno. Tra i territori con la più elevata variabilità emergono in particolare la Lombardia, l'Emilia Romagna, il Piemonte, la Valle d'Aosta e la Liguria; mentre tra le regioni con minore variabilità, si possono citare

Campania, Sardegna, Sicilia e Calabria. Sorprendentemente, la Lombardia è superata dalla Valle d'Aosta e in parte dalla Provincia autonoma di Trento se si considera l'incidenza delle terapie intensive.

Figura 1: Tasso giornaliero di terapie intensive per regione per 100.000 abitanti nel periodo 24/02/2020 – 06/07/2020

Fonte: elaborazione degli autori su dati della Protezione civile

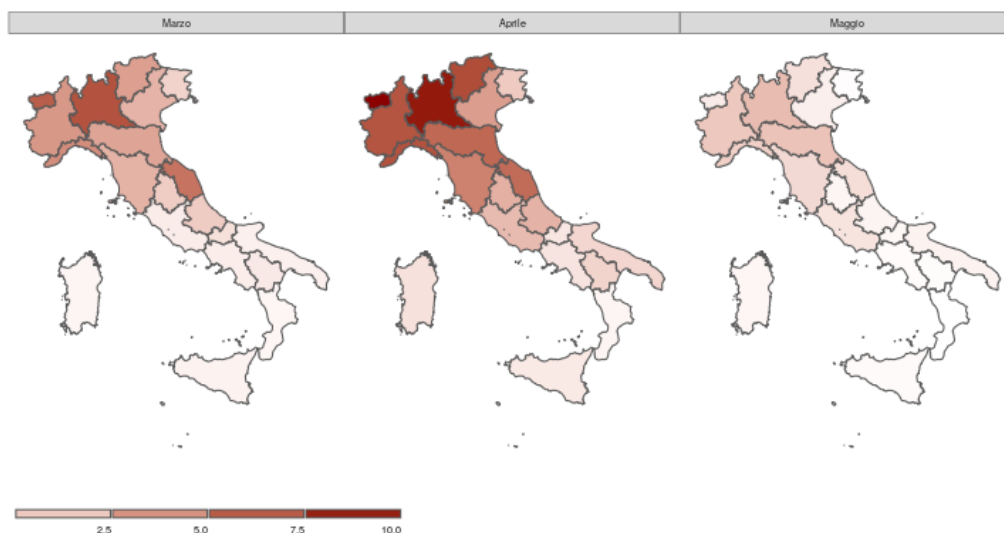
Figura 2: Tasso giornaliero di ospedalizzati per regione per 100.000 abitanti nel periodo 24/02/2020 – 06/07/2020.



Fonte: elaborazione degli autori su dati della Protezione civile

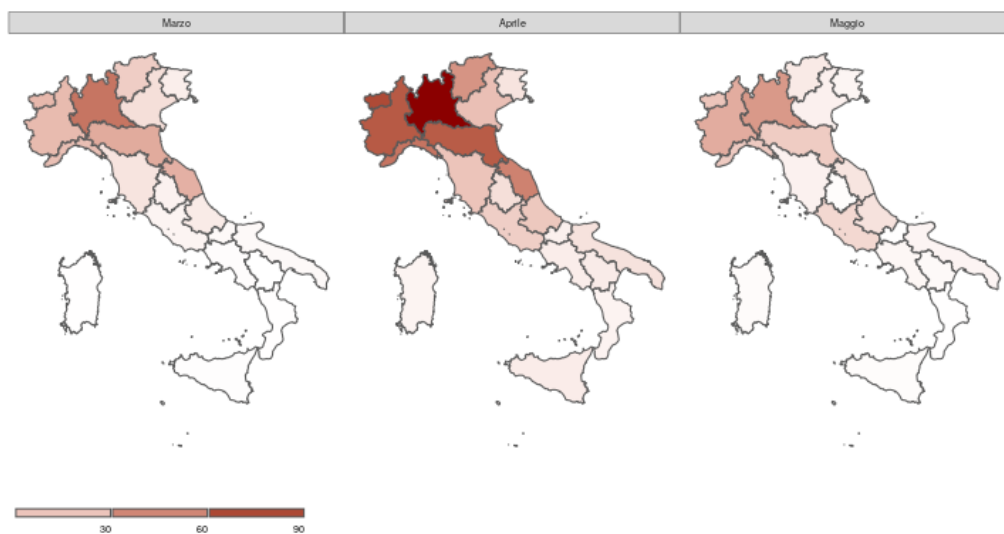
Allo scopo di evidenziare maggiormente le disparità territoriali, si presentano le mappe territoriali delle medie mensili del numero di terapie intensive e ospedalizzati associate ai mesi più critici della pandemia. Ne emerge la sofferenza di tutte le regioni del Nord e, parzialmente, di quelle del Centro nel mese di aprile per quanto concerne le terapie intensive (Fig. 3), laddove nel caso degli ospedalizzati una maggiore incidenza è visibile in Lombardia, nelle regioni del Nord-ovest, in Emilia Romagna, nelle Marche e nel Trentino-Alto Adige (Fig. 4). In questo caso si nota la minore incidenza di ospedalizzati (sul totale degli abitanti) in Veneto e Toscana rispetto al suddetto gruppo di regioni.

Figura 3: Mappa regionale del tasso medio mensile di terapie intensive per 100.000 abitanti.



Fonte: elaborazione degli autori su dati della Protezione civile

Figura 4: Mappa regionale del tasso medio mensile di ospedalizzati totali per 100.000 abitanti.



Fonte: elaborazione degli autori su dati della Protezione civile

I profili regionali: uno sguardo oltre la prossimità territoriale

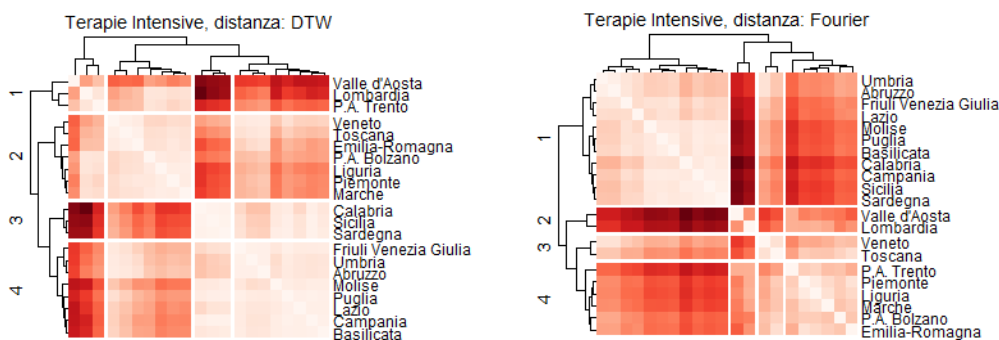
Nelle figure che seguono (Figg. 5 e 6), le distanze tra le serie storiche calcolate nei termini di DTW e Fourier sono rappresentate tramite mappe di calore (*heatmap*). In questa rappresentazione, laddove il colore diventa più scuro significa che vi è un aumento della distanza tra le serie.

Per quanto riguarda le terapie intensive, utilizzando la DTW emergono quattro gruppi (Fig. 5):

- il primo include Lombardia, Valle d'Aosta e Provincia autonoma di Trento;
- il secondo è composto da Veneto, Toscana, Emilia Romagna, Provincia autonoma di Bolzano, Piemonte e Marche;
- un terzo gruppo contiene la Calabria e le due Isole;
- l'ultimo gruppo include tutte le rimanenti regioni del Centro-sud e, sorprendentemente, il Friuli-Venezia Giulia.

Utilizzando invece la distanza di Fourier, si denota una forte distinzione tra due “micro” cluster, formati soltanto da due unità ciascuno (Valle d'Aosta – Lombardia e Veneto – Toscana), e tutte le altre regioni. Ancora una volta il Friuli-Venezia Giulia viene associato alle regioni del Centro-sud, che questa volta comprendono anche le Isole.

Figura 5: Matrice delle distanze DTW e Fourier delle serie storiche per le terapie intensive nel periodo 24/02/2020 – 06/07/2020.



Fonte: elaborazione degli autori su dati della Protezione civile

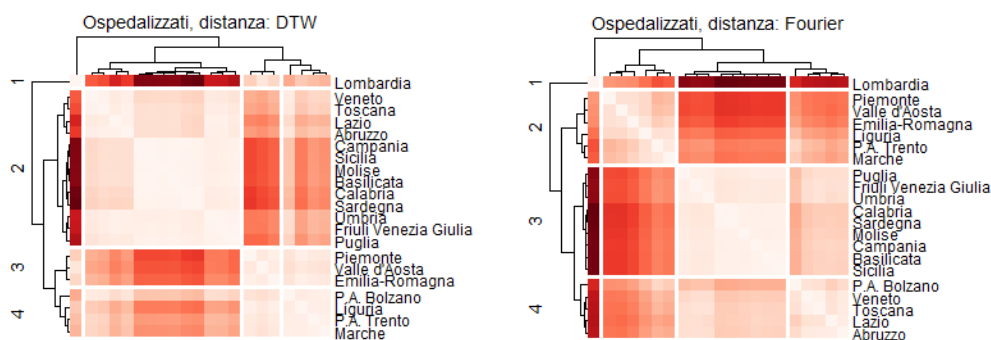
I gruppi così ottenuti risultano più omogenei utilizzando la DTW piuttosto che la distanza di Fourier e, in particolare nel primo caso, si evidenzia una maggiore caratterizzazione tra i gruppi, specialmente tra quelli del Sud rispetto alle regioni più colpite.

Andando ad analizzare invece il caso degli ospedalizzati, la DTW e la classificazione gerarchica hanno permesso di identificare i seguenti gruppi (Fig. 6):

- un primo cluster che include solo la Lombardia;
- un secondo cluster formato da Piemonte, Valle d'Aosta ed Emilia Romagna;
- un terzo cluster che include Bolzano, Trento e le Marche;

ed un ultimo cluster comprendente tutte le regioni del Centro-sud unite a Veneto e Friuli-Venezia Giulia.

Figura 6: Matrice delle distanze DTW e Fourier delle serie storiche per gli ospedalizzati nel periodo 24/02/2020 – 06/07/2020.



Fonte: elaborazione degli autori su dati della Protezione civile

Utilizzando invece la distanza di Fourier, la Lombardia si conferma unico elemento del proprio *cluster* e gli altri gruppi risultano più omogenei in termini numerici rispetto al DTW; in particolare citiamo il gruppo composto da Valle d'Aosta, Piemonte, Marche, Liguria, Provincia di Trento e Emilia Romagna. Ancora una volta si nota come il Friuli-Venezia Giulia venga ricompreso all'interno del gruppo delle regioni del Mezzogiorno. Anche in questo caso, per entrambi i metodi adottati, si evidenzia una buona caratterizzazione tra i gruppi, mentre il cluster della Lombardia, seppur discostandosi in maniera evidente dalle regioni del Centro-sud, non risulta troppo distanziato dalle regioni confinanti.

Riflessioni conclusive

Le analisi condotte in questo lavoro evidenziano affinità e divergenze tra le regioni italiane in relazione all'incidenza sanitaria della prima ondata della pandemia COVID-19. Tali unità di indagine risultano particolarmente rilevanti per il ruolo centrale che rivestono in qualità di livelli territoriali responsabili dell'organizzazione delle strutture e dei servizi sanitari, e delle misure di contrasto nel periodo delle emergenze sanitarie.

In base ai risultati ottenuti attraverso un confronto tra le distanze DTW delle serie storiche relative a ospedalizzati e ricoverati in terapia intensiva, si può notare un'Italia caratterizzata da più cluster regionali, non tutti identificabili unicamente dalla prossimità geografica e dove la pandemia ha impattato in maniera piuttosto differente. All'interno delle regioni del Nord, ossia laddove vi è stata una maggiore incidenza del fenomeno endemico, sia il Veneto che soprattutto il Friuli-Venezia Giulia hanno presentato un andamento più in linea con quello delle regioni del Mezzogiorno piuttosto che con quello delle regioni maggiormente colpite dalla pandemia. Dall'analisi emerge inoltre che Valle d'Aosta, Lombardia e la Provincia di Trento risultano le regioni maggiormente colpite se si considerano in particolar modo il tasso di ricoverati in terapia intensiva, seguite da Piemonte, Liguria, Emilia Romagna e Marche.

L'osservazione di tali cluster potrebbe rappresentare un punto di partenza per il miglioramento della *governance* sanitaria ed essere utile nel monitoraggio dei potenziali squilibri dovuti ad ulteriori ondate, purtroppo non del tutto escludibili in una prospettiva scientifica (Xu e Li, 2020).

Lucio Palazzo, Università di Napoli
Riccardo Ievoli, Università di Ferrara

Riferimenti bibliografici

Agosto, A., Giudici, P. (2020). "A Poisson Autoregressive Model to Understand Covid-19 Contagion Dynamics." *Available at SSRN 3551626*.

Agrawal, R., Faloutsos, C., Swami, A. (1993). "Efficient Similarity Search in Sequence Databases." In *International Conference on Foundations of Data Organization and Algorithms*, 69–84. Springer, Berlin Heidelberg.

Berndt, D.J., Clifford, J. (1994). "Using Dynamic Time Warping to Find Patterns in Time Series." In *KDD Workshop*, 10:359–70. 16. Seattle, WA, Stati Uniti:

Boccia, S., Ricciardi W., Ioannidis, J.PA. (2020). "What Other Countries Can Learn from Italy During the Covid-19 Pandemic." *JAMA Internal Medicine*, 180 (7): 927–928.

Farcomeni, A., Maruotti, A., Divino, F., Jona Lasinio G., Lovison, G. (2020). "An Ensemble Approach to Short-Term Forecast of Covid-19 Intensive Care Occupancy in Italian Regions." *arXiv Preprint arXiv:2005.11975*.

James, N., Menzies, M. (2020). "Cluster-Based Dual Evolution for Multivariate Time Series: Analyzing Covid-19." *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 30 (6): 061108.

Murtagh, F., Contreras, P. (2012). "Algorithms for Hierarchical Clustering: An Overview." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2 (1): 86–97.

Musolino, D., Rizzi, P. 2020. "Covid-19 E Territorio: Un'analisi a Scala Provinciale." *EyesReg*, 10 (3).

Rojas, I., Rojas, F., Valenzuela, O. (2020). "Estimation of Covid-19 Dynamics in the Different States of the United States Using Time-Series Clustering." *medRxiv*. Doi: <https://doi.org/10.1101/2020.06.29.20142364>

Sebastiani, G., Massa, M., Riboli, E. (2020). "Covid-19 Epidemic in Italy: Evolution, Projections and Impact of Government Measures." *European Journal of Epidemiology*, 35 (4): 341. Xu, S., Li, Y. (2020). "Beware of the Second Wave of Covid-19." *The Lancet*, 395 (10233): 1321–2.