



FOCUS *n. 10/2021*

Studio della struttura spaziale dei decessi in Puglia

INTRODUZIONE

L'obiettivo di modellare la dinamica spazio-temporale dell'attuale pandemia risponde a diverse esigenze, tra le quali la pianificazione dei servizi sanitari, l'adozione delle misure di contenimento, solo per citarne alcune.

Fra gli studi recenti aventi questo scopo, alcuni sperimentano con risultati incoraggianti i metodi e le tecniche della geostatistica. Con la finalità di mappare e visualizzare il rischio spaziale di diffusione del Covid-19, Cardoso et al. [1] e Azevedo et al. [2] utilizzano la simulazione sequenziale geostatistica a blocchi, calcolando l'incertezza spaziale associata della previsione; Ruchjana et al. [3] conducono uno studio per predire la distribuzione dei contagi in 27 distretti/città nel West Java, attraverso il Kriging Ordinario. Per applicazioni ai problemi generali di salute pubblica, Diggle e Giorgi [4] e Goovaerts [5] analizzano e sistematizzano le fasi di un'analisi geostatistica.

Come sottolineato nel lavoro di Nuvolone et al. [6] che utilizzano dati epidemiologici per derivare un indicatore ambientale sulla salute respiratoria, l'individuazione della continuità spaziale alla scala di osservazione del fenomeno è preliminare a qualsiasi altra analisi geostatistica. Potrebbe accadere, infatti, che la distribuzione geografica dei dati non sia in grado di rilevarla.

Sulla base dei dati pubblicati mensilmente dall'ISTAT riguardanti i decessi registrati nei comuni italiani dal 2011 fino a settembre 2021, lo studio proposto in questo *focus* ha l'obiettivo di indagare sulla struttura spaziale (variogramma) dei decessi avvenuti in emergenza sanitaria, nei primi sette mesi del 2021, nei comuni della Puglia, confrontando il risultato con il variogramma di due anni pre-Covid-19 (2011 e 2014). Preliminarmente, tale variabile è descritta statisticamente, confrontata temporalmente, rappresentata spazialmente attraverso mappe.

FONTE DEI DATI

Istat – *Dataset* con i decessi giornalieri in ogni singolo comune di residenza per sesso e classi di età quinquennali (aggiornamento mensile)

METODOLOGIA

La geostatistica mette a disposizione gli strumenti per l'analisi e la modellizzazione delle variazioni spaziali di una variabile partendo dai dati, la cui distribuzione sul territorio caratterizza la scala di osservazione del fenomeno. Tale approccio trova giovamento dalle conoscenze a priori sulla situazione fisica e territoriale rispetto alla quale il fenomeno in esame si manifesta o quant'altro utilizzabile nella sua interpretazione.

Si consideri $c(\mathbf{u}_\alpha)$, il numero di decessi in ciascun comune α , con $\alpha=1,\dots,N$, il cui centroide geometrico \mathbf{u}_α ha coordinate (x_α, y_α) e $n(\mathbf{u}_\alpha)$ rappresenta la dimensione della popolazione del comune α . La variabile $z(\mathbf{u}_\alpha)$ può essere espressa come:

$$z(\mathbf{u}_\alpha) = c(\mathbf{u}_\alpha) / n(\mathbf{u}_\alpha)$$

In geostatistica, per ogni fissato punto \mathbf{u}_0 si assume che $z(\mathbf{u}_0)$ sia la realizzazione di una variabile aleatoria e, al variare del vettore posizione \mathbf{u} , $Z(\mathbf{u})$ sia una funzione aleatoria. Ipotesi sui parametri della distribuzione di probabilità di $Z(\mathbf{u})$ permettono di semplificare il modello. Le *ipotesi intrinseche* sono quelle più deboli e riguardano la stazionarietà della media

$m(\mathbf{h})$ e della covarianza $C(\mathbf{h})$ della funzione aleatoria degli incrementi

$$Z(\mathbf{h}) = Z(\mathbf{u} + \mathbf{h}) - Z(\mathbf{u})$$

In tal caso, in luogo della funzione di $C(\mathbf{h})$ si introduce la funzione variogramma $\gamma(\mathbf{h})$ essendo, per variogrammi limitati, valida la seguente relazione:

$$\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h})$$

Il variogramma descrive la continuità spaziale del fenomeno che si può esplicitare in questo modo: quanto più coppie di comuni a distanza \mathbf{h} sono vicine tra loro tanto meno varia il valore della variabile in esame. In $\mathbf{h}=0$, la variazione è nulla

$$Z(\mathbf{0}) = Z(\mathbf{u}) - Z(\mathbf{u}) = 0$$

L'obiettivo dell'analisi strutturale è l'individuazione di un comportamento crescente nelle vicinanze dell'origine per il variogramma sperimentale, stimato con i dati disponibili.

N comuni producono $N*(N-1)/2$ coppie distinte di incrementi. Nell'ipotesi di fenomeno isotropico, ciascuna di esse va a finire in una classe di distanze. Tutte le classi devono contenere un numero significativo di valori, in particolare le prime, quelle più vicine a $\mathbf{h}=0$. Le classi sono multipli di un intervallo fisso di riferimento, detto *lag*. Il grafico che si costruisce sulla media delle distanze e degli incrementi quadratici della variabile per ciascuna classe è detto variogramma sperimentale.

Per il calcolo degli incrementi quadratici si è applicato il seguente adattamento del variogramma che tiene conto della popolazione, proposto da diversi autori, fra i quali Monestiez et al. [7]:

$$w(h) = \frac{n(u_\alpha)n(u_\alpha + h)}{n(u_\alpha) + n(u_\alpha + h)}$$

$$(1) \quad \gamma(h) = \frac{1}{2 \sum_{\alpha=1}^{N(h)} w(h)} \sum_{\alpha=1}^{N(h)} w(h) [c(u_\alpha) - c(u_\alpha + h)]^2$$

STATISTICHE DESCRITTIVE

L'incidenza di decessi in soggetti di età 50-74 (fig. 1) e 75-100+ (fig. 2) cresce nel 2021 in tutte le città capoluogo di provincia rispetto alle incidenze pre-Covid del 2019.

Fig. 1 - Incidenza dei decessi in soggetti di età 50-74 sul totale decessi, per comuni capoluogo. Anni 2019-2021

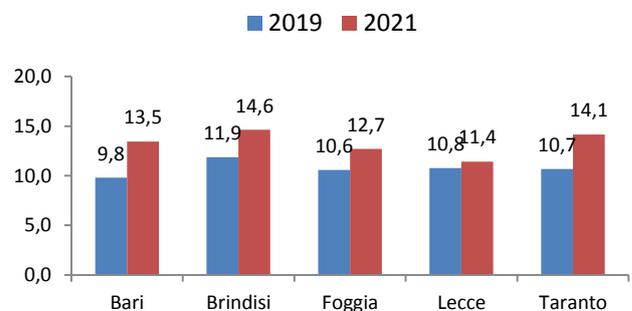
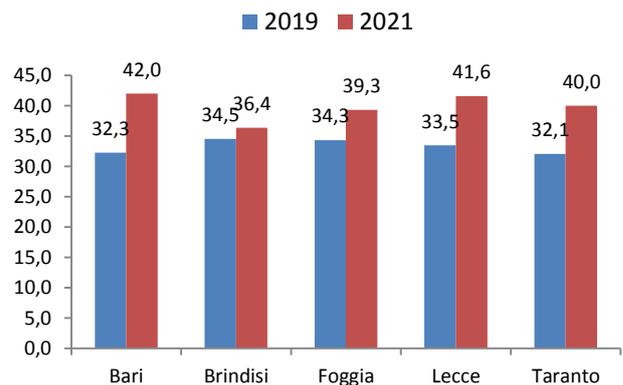


Fig. 2 - Incidenza di decessi in soggetti di età 75-100+ sul totale, per comuni capoluogo. Anni 2019-2021



L'andamento giornaliero dei decessi nel 2019 è decrescente, minore nel periodo più caldo. Per i comuni capoluogo esaminati, il modello lineare di regressione è altamente significativo ($p < 0,0001$). Lo stesso andamento e la stessa significatività del modello di regressione si rileva nel 2021 confermando la minore circolazione del virus e il minor numero di

contagi nella vita all'aperto, tipico del periodo estivo (da fig. 3 a fig. 7).

Fig. 3 - Andamento del numero di decessi per giorno (1 gen – 31 luglio) e rette di regressione. Bari. Anni 2019-2021

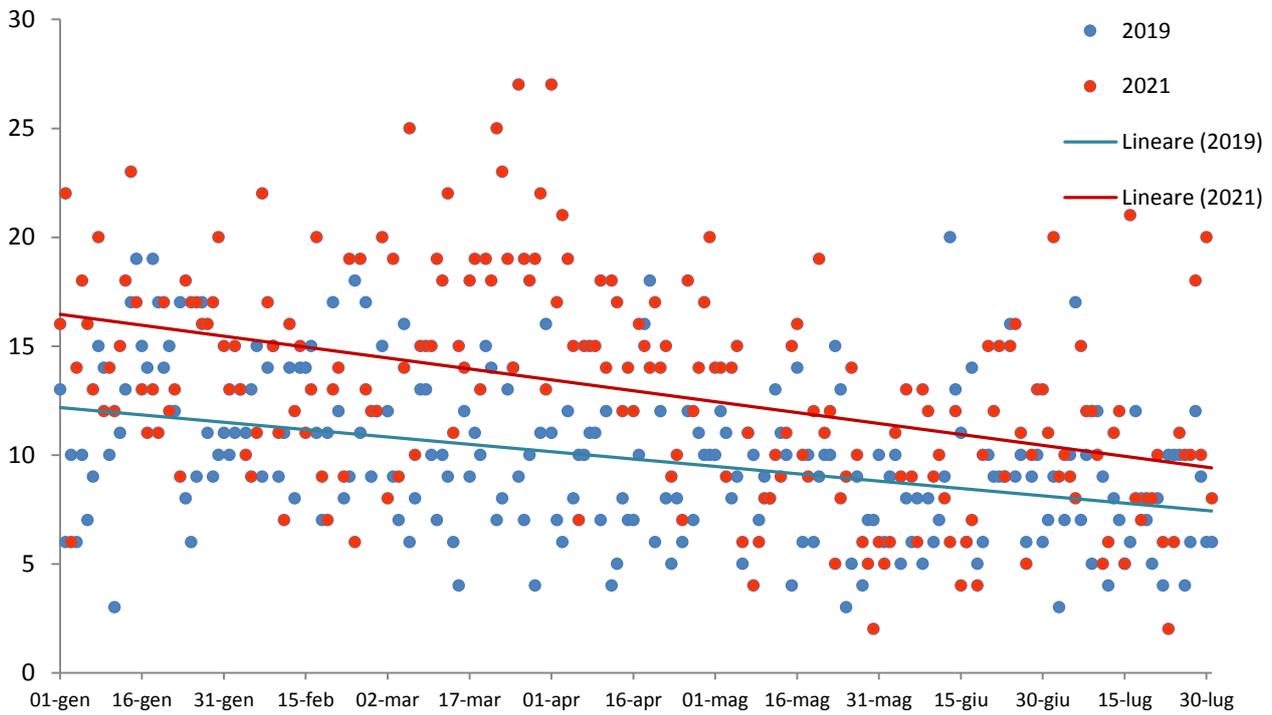


Fig. 4 - Andamento del numero di decessi per giorno (1 gen – 31 luglio) e rette di regressione. Brindisi. Anni 2019-2021

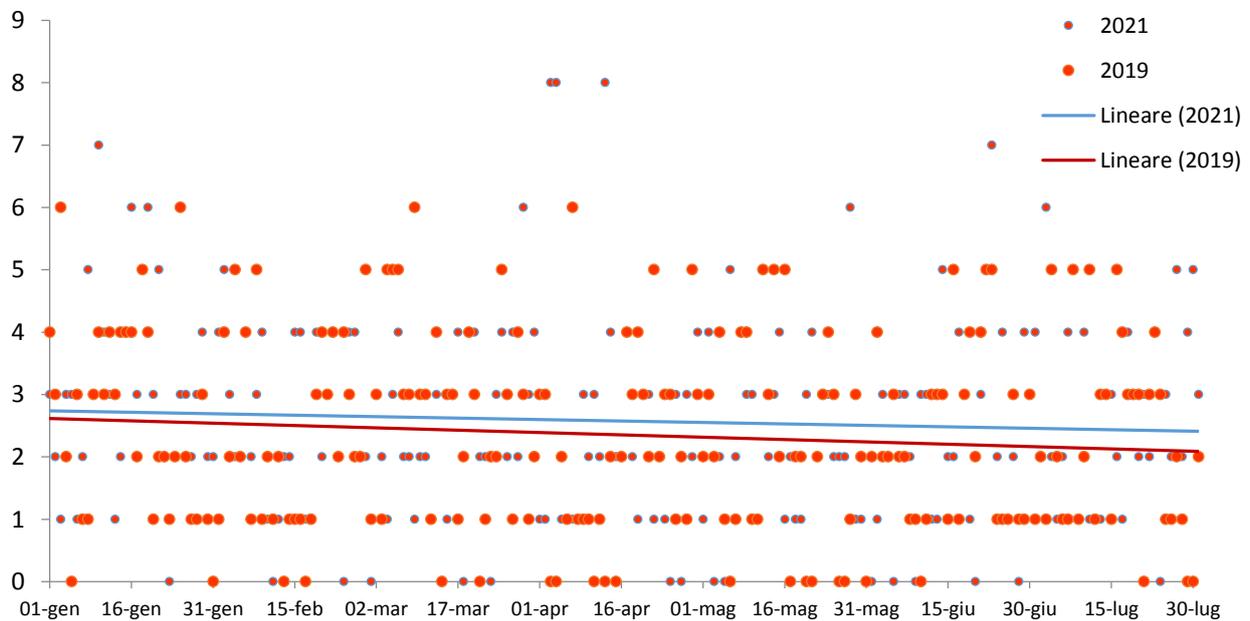


Fig. 5 - Andamento del numero di decessi per giorno (1 gen – 31 luglio) e rette di regressione. Foggia. Anni 2019-2021

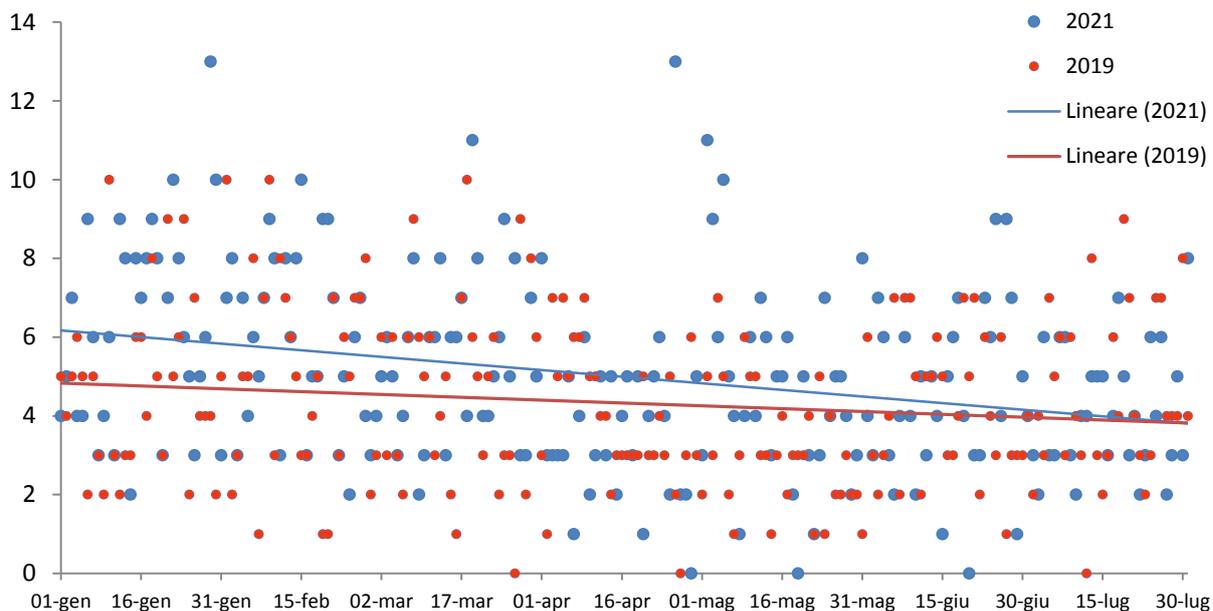


Fig. 6 - Andamento del numero di decessi per giorno (1 gen – 31 luglio) e rette di regressione. Lecce. Anni 2019-2021

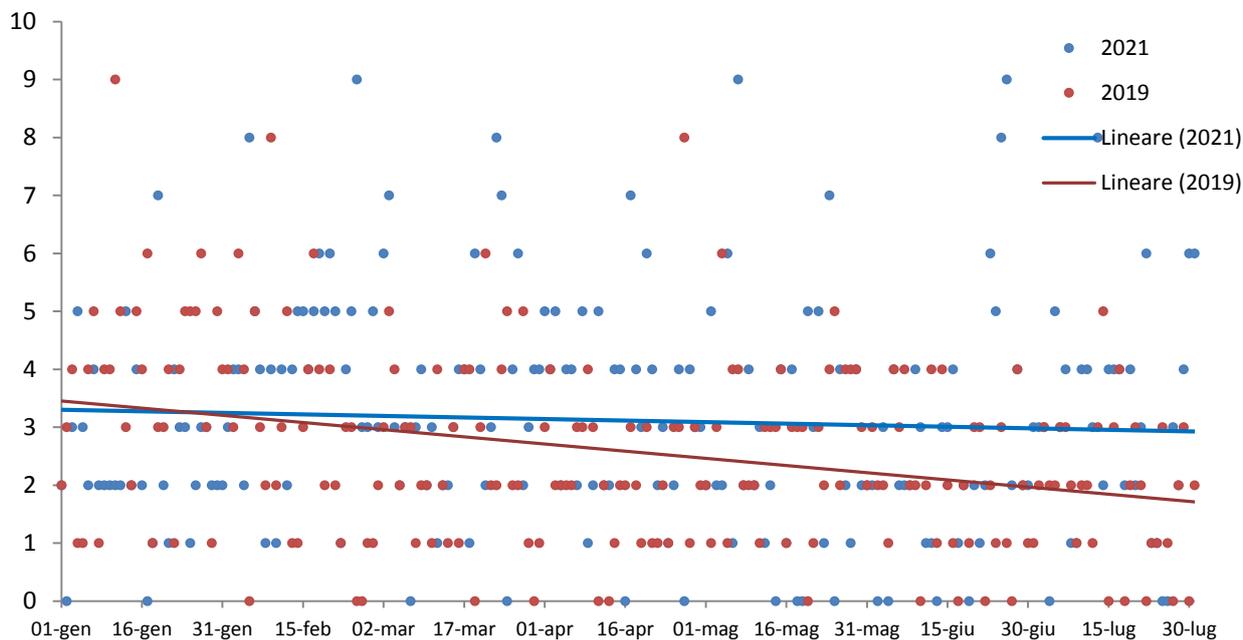
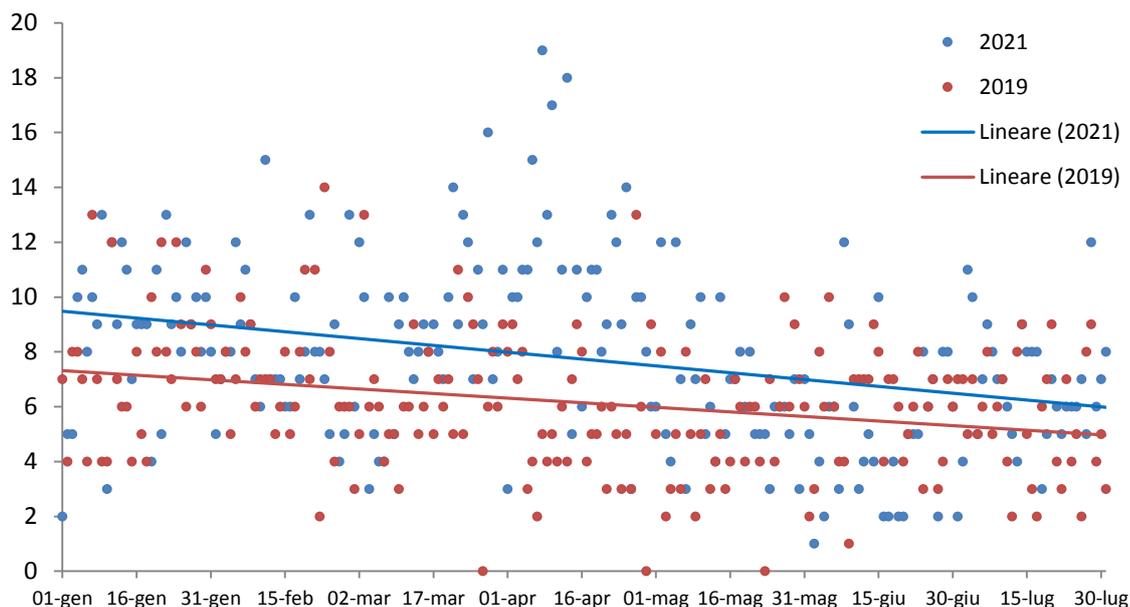


Fig. 7 - Andamento del numero di decessi per giorno (1 gen – 31 luglio) e rette di regressione. Taranto. Anni 2019-2021



Le serie storiche, ottenute dalla rappresentazione cartesiana del numero di decessi e della data di avvenimento nei primi sette mesi del 2019 (periodo pre covid) e del 2021 (periodo covid), possono essere descritte attraverso le medie mobili, calcolate rispetto a un intervallo temporale, nel caso in esame è pari a 8 giorni. Appare una certa similitudine negli andamenti delle medie mobili dei decessi di Bari e Taranto (figg. 8 e 12) viene

evidenziata, infatti, una differenza netta nei due anni tra i decessi avvenuti tra i primi di Marzo e i primi di Maggio nel primo caso, e tra metà Marzo e fine Maggio nel secondo. Aspetti di variabilità meno evidenti nel tempo, per quanto attiene i decessi, si osservano a Brindisi e a Foggia (fig. 9 e 10), a Lecce si riscontrano dei picchi nel periodo tra metà Febbraio e inizi di Marzo e tra Giugno e Luglio (fig. 11).

Fig.8 - Andamento della media mobile (m.m.) di decessi per giorno (1 gen – 31 luglio). Bari. Anni 2019-2021 (m.m.) di 8 gg.

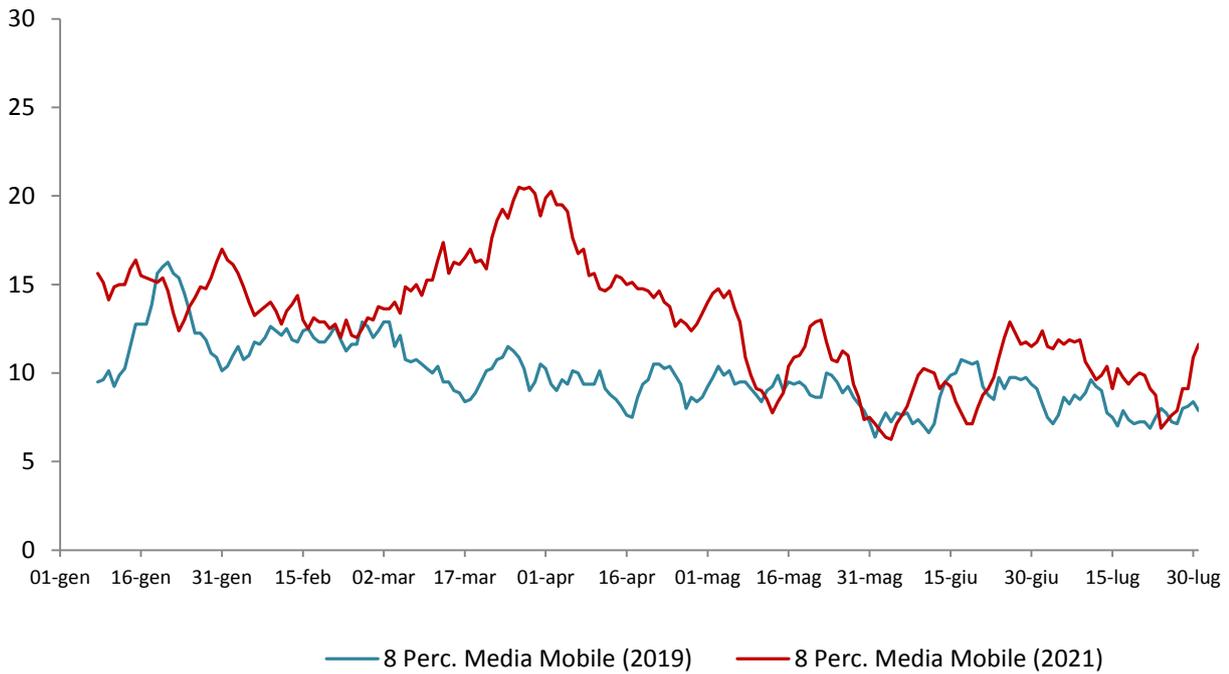


Fig. 9 - Andamento della media mobile (m.m.) di decessi per giorno (1 gen – 31 luglio). Brindisi. Anni 2019-2021 (m.m. di 8 gg.)

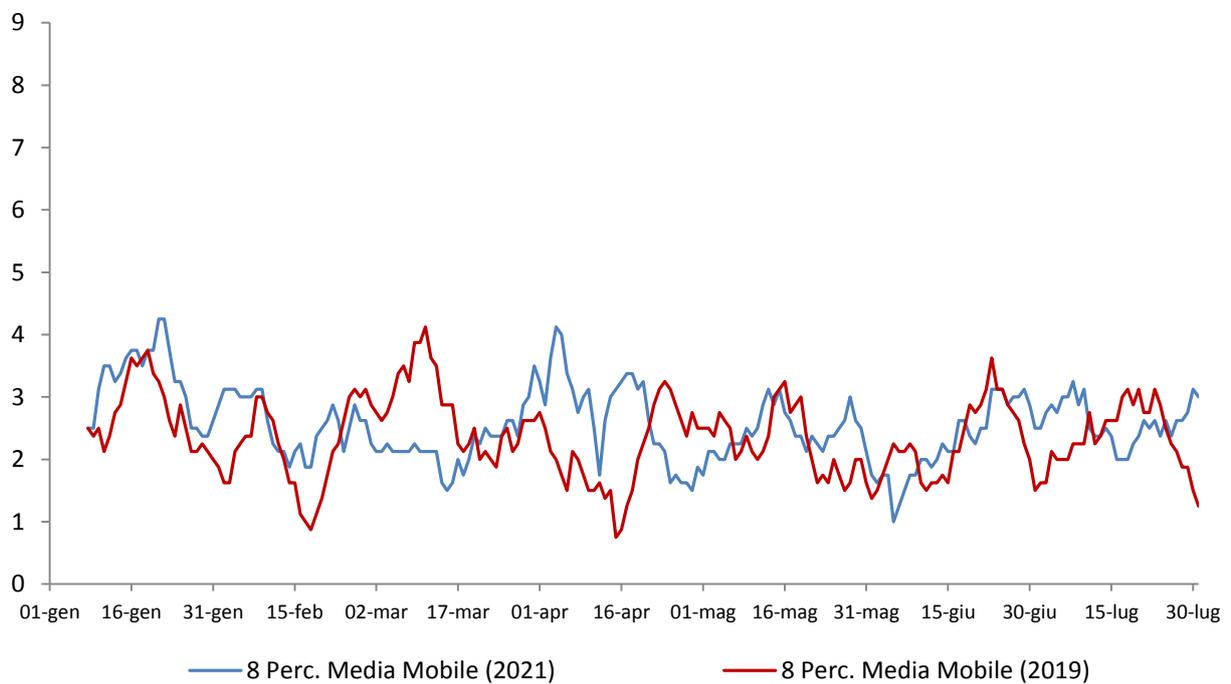


Fig. 10 - Andamento della media mobile (m.m.) di decessi per giorno (1 gen – 31 luglio). Foggia. Anni 2019-2021 (m.m. di 8 gg.)

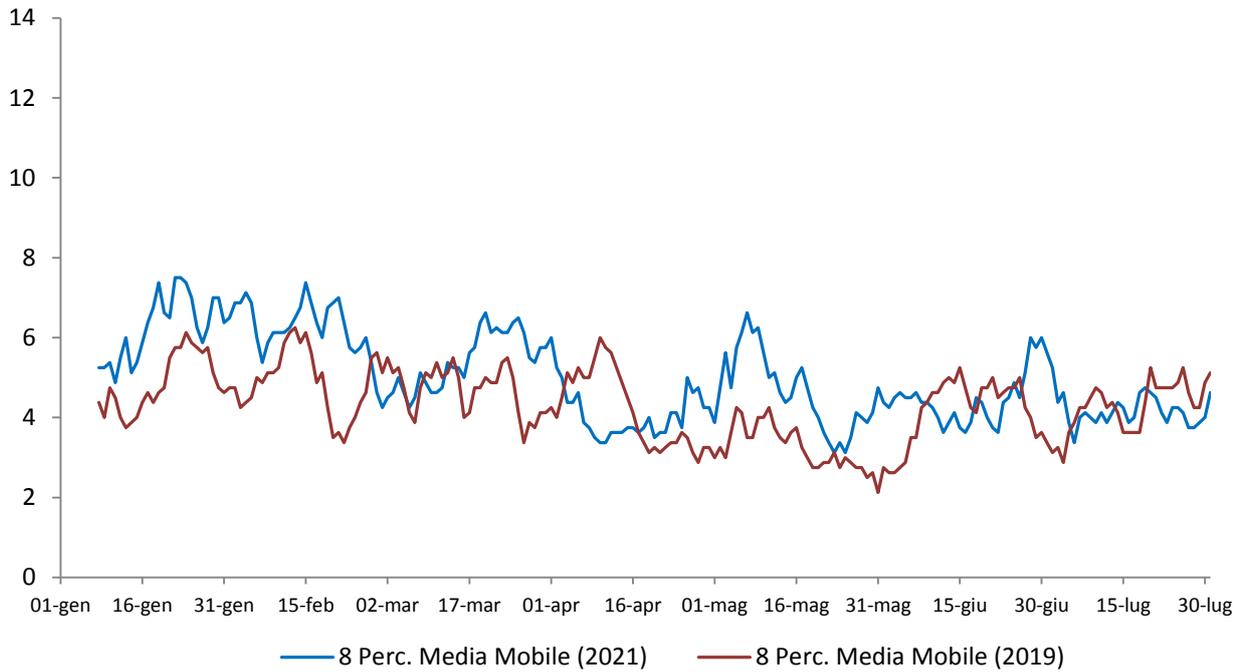


Fig. 11 - Andamento della media mobile (m.m.) di decessi per giorno (1 gen – 31 luglio). Lecce. Anni 2019-2021 (m.m. di 8 gg.)

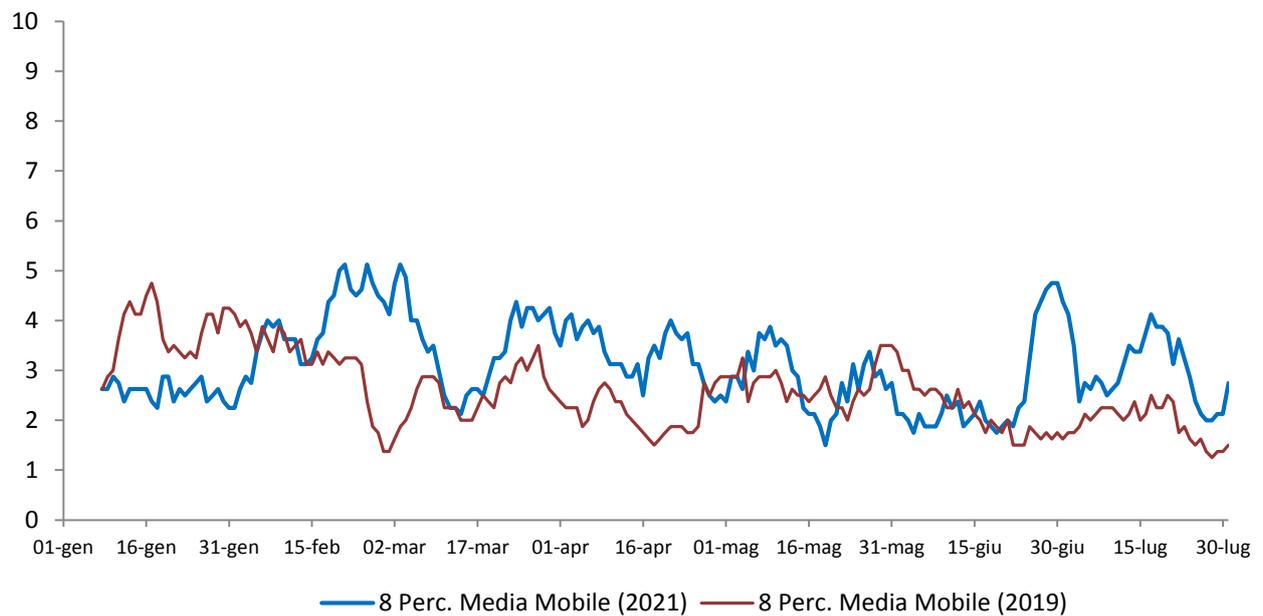
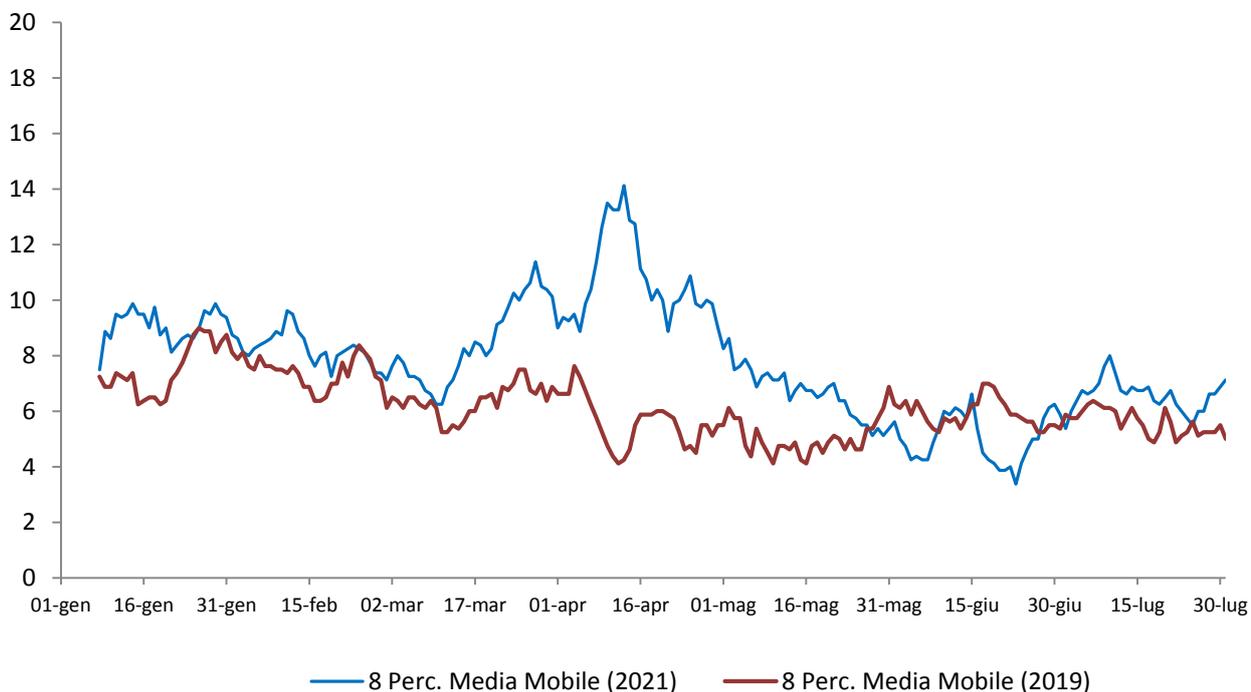


Fig. 12 - Andamento della media mobile (m.m.) di decessi per giorno (1 gen – 31 luglio). Taranto. Anni 2019-2021 (m.m. di 8 gg.)



Nella mappa del 2021 (fig. 13), a parità di classe di colorazione rispetto a quella del 2019 (fig. 14), si nota una maggiore incidenza dei decessi

sul totale della popolazione nella zona del foggiano, all'interno della provincia di Bari e di Taranto. Meno evidente nella provincia di Lecce

Fig. 13 - Mappa spaziale dei decessi sul totale della popolazione comunale. Anno 2019

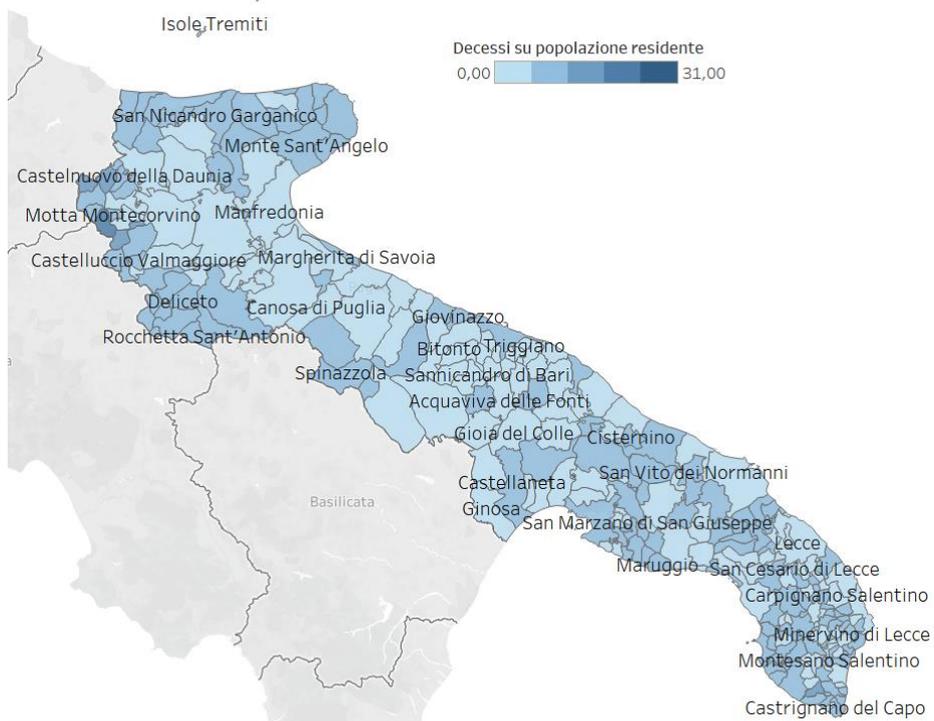
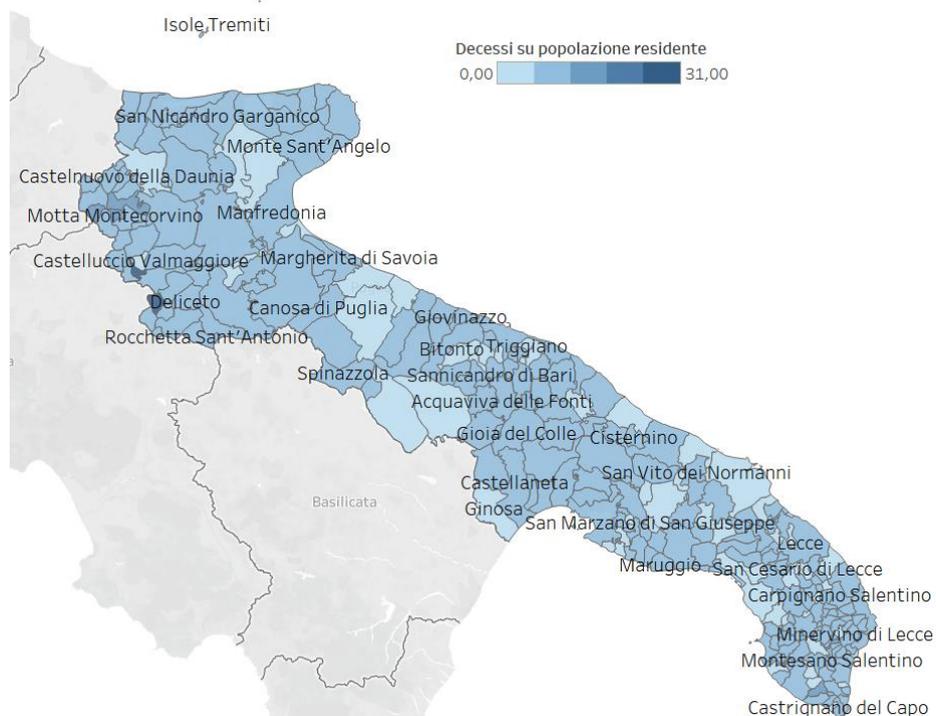


Fig. 14 - Mappa spaziale dei decessi sul totale della popolazione comunale. Anno 2021



STRUTTURA SPAZIALE DEI DECESSI

L'Istat rilascia le matrici origine-destinazione delle distanze in metri fra tutti i comuni italiani. L'elaborazione è effettuata attraverso strumenti GIS, utilizzando un grafo stradale commerciale. Gli elementi vettoriali utilizzati sono i centroidi dei comuni Italiani al 2013 e il grafo stradale MultiNet 2013 di TomTom. I centroidi dei comuni Italiani sono ottenuti partendo dalla sezione di censimento che contiene il municipio. Il processo di calcolo che consente la produzione delle matrici delle distanze è ottenuto mediante gli algoritmi di analisi di rete dei *Geographic Information System* (GIS).

Per ciascuna coppia di comuni pugliesi di cui si ottiene la distanza attraverso i dati Istat è calcolata la differenza quadratica della variabile, pesata con la popolazione dell'anno di riferimento (formula 1).

Si utilizza una classe di distanze pari a multipli di 2 km (*lag*). Tra 0 e 2 Km si trovano 14 coppie di comuni, con una distanza media di 1,7 km; fra 2 e 4 km si trovano 57 coppie che producono una distanza media di 2,5 km, e così via (tab. 1)

Tab. 1 – Statistiche del variogramma sperimentale

Numero coppie comuni	h	2011	2014	2021
14	0	0	0	0
57	1720,2	367,9	274,6	475,4
81	2528,9	476,4	417,2	734,1
110	3528,3	765,7	808,4	1102,5
122	4452,9	751,9	753,8	1158,5
188	5485,6	515,8	506,2	884,7
211	6508,2	8623,3	9135,6	11993,1
199	7453,3	12407,1	13085,8	18093,0
201	9534,7	76298,0	69465,2	141751,5

Escludendo il punto $h=0$, il variogramma sperimentale del 2021 mostra un andamento crescente fino alla quarta classe, quella compresa fra 4 e 5 km. Anche per i due anni pre covid presi in considerazione (2011 e 2014) la

monotonia crescente si manifesta tendenzialmente fino alla quarta classe: crescente fino alla terza, con una lieve variazione decrescente nella quarta (fig. 15). Dopo, in tutti e tre gli anni, diminuisce in modo spiccato l'incremento quadratico medio degli incrementi.

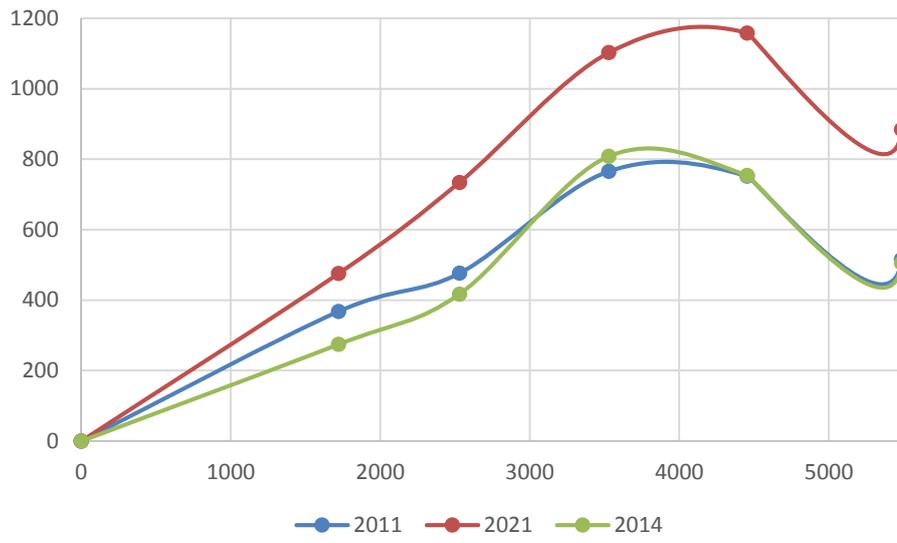
Si conclude che la continuità spaziale si manifesta fino a circa 5 km di distanza fra i comuni, nel 2014 e 2011 così come nel 2021: non varia, dunque, il massimo raggio di influenza della variabile decessi nell'anno Covid. Le misure di contenimento spaziale adottate durante la pandemia, quindi, sono risultate efficaci in quanto non si evidenzia l'alterazione della dimensione d'influenza che i fattori di ordine spaziale/ambientale hanno sulla mortalità, così come si registra negli anni pre-Covid.

E' pur vero, però, che all'interno delle misure di confinamento, conseguenti all'aumento dei contagi in alcuni periodi del 2021, il virus ha continuato a circolare a livello comunale. L'esistenza di focolai ha inciso sull'aumento della mortalità causata dal Covid alla scala di singolo comune, causando una maggiore crescita dell'incremento quadratico dei decessi rispetto ai comuni più vicini. Così si spiega il posizionamento della curva sperimentale del variogramma del 2021 sempre al disopra delle altre due.

La modellizzazione del variogramma sperimentale attraverso una forma funzionale definita positiva permetterebbe di utilizzare le tecniche di valutazione del rischio.

Questi primi risultati confermano l'utilità delle tecniche geostatistiche, incoraggiando il proseguimento delle sperimentazioni.

Fig. 15 – Andamento del variogramma sperimentale. Anni 2011-2014-2021



Bibliografia

[1] Mário Cardoso, André Cavalheiro, Alexandre Borges, Ana F. Duarte, Amílcar Soares, Maria João Pereira, Nuno Jardim Nunes, Leonardo Azevedo, Arlindo L. Oliveira. "Modeling the geospatial evolution of COVID-19 using spatio-temporal convolutional sequence-to-sequence neural networks". Anno 2020

[2] Leonardo Azevedo, Maria João Pereira, Manuel C. Ribeiro & Amílcar Soares. "Geostatistical COVID-19 infection risk maps for Portugal", International Journal of Health Geographics volume 19, Article number: 25 (2020)

[3] B N Ruchjana, A N Falah and A S Abdullah. "Application of the Ordinary Kriging method for prediction of the positive spread of Covid-19 in West Java," Journal of Physics: Conference Series, Volume 1722, Tenth International Conference and Workshop on High Dimensional Data Analysis (ICW-HDDA-X) 12-15 October 2020

[4] Peter J. Diggle and Emanuele Giorgi. "Model-Based Geostatistics for Global Public Health: Methods and Applications", FL: Chapman & Hall/CRC Press, 2019, xxvi + 247 pp., \$99.95(H), ISBN: 978-1-13-873235-3

[5] Goevaerts P. "Geostatistical Analysis of Health Data: State-of-the-Art and Perspectives", Geostatistics for Environmental Applications pp 3-22, 2020

[6] Nuvolone D., Fresco R., Maio S., Baldacci S., Angino A., Martini F., Borbotti M., Viegi G., della Maggiore R. "Application of Geostatistical Methods for Public Health Risk Mapping", Association of Geographic Information Laboratories in Europe, Paper presentato alla Conference 2021.

[7] Monestiez P, Dubroca L, Bonnin E, Durbec J, Guinet C. "Geostatistical modeling of spatial distribution of *Balaenoptera physalus* in the

Northwestern Mediterranean Sea from sparse count data and heterogenous observation efforts". Ecol Model. 2006;193:615–28, 2019

UFFICIO STATISTICO, Via Gentile 52 - 70126 Bari email: ufficio.statistico@regione.puglia.it; www.regione.puglia.it/ufficiostatistico