

Emergenza COVID - Situazione Attuale

Aggiornato al 23 Marzo 2020

Francesco Furno[†]

Lo scopo della presente analisi e' di fornire una fotografia dello stato *attuale* dell'epidemia COVID in Italia. Non e' basata su modelli epidemiologici matematici, ma sull'analisi statistica (per lo piu' descrittiva) dei dati regionali resi disponibili dal Ministero.

1 Riassunto Punti Principali

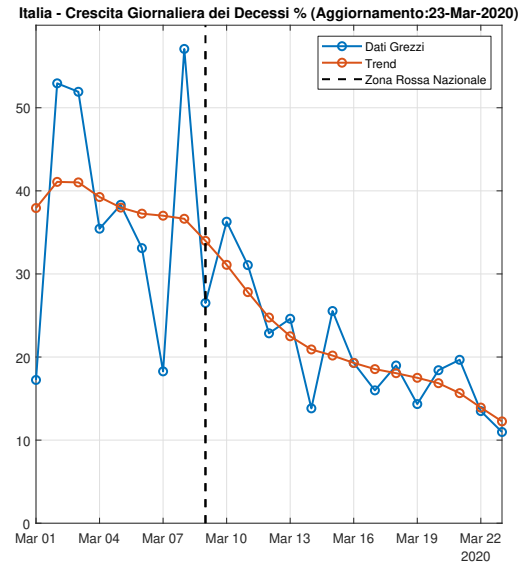
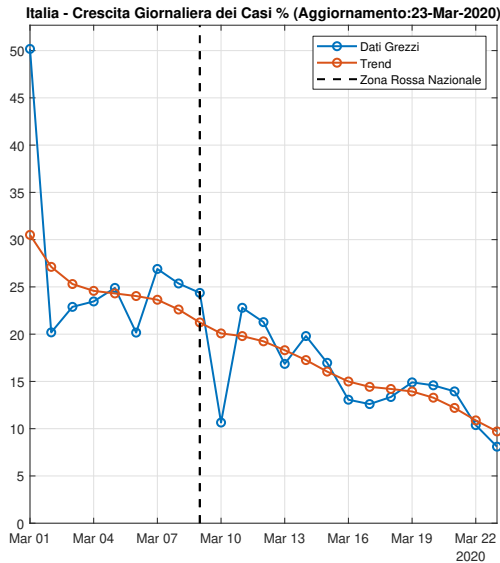
1. I dati giornalieri devono essere filtrati dalla volatilità giornaliera per poterne estrarre un **trend** piu' facilmente interpretabile.
2. Rallentamento della **crescita giornaliera dei contagi e dei decessi** in atto da giorni, efficace zona rossa nazionale. Il trend non e' pero' uniforme tra regioni.
3. **Tamponi totali** in costante aumento, ma **propensione a eseguire tamponi** in calo in tutte le regioni, misurata come tamponi totali su casi totali regione per regione. Questo riflette sia la difficolta' del sistema sanitario di stare al passo con l'epidemia sia un cambiamento di strategia. 1 test su 4 e' attualmente positivo al COVID, rispetto a 1 test ogni 12 del 1 Marzo scorso.
4. **Tassi di letalità** molto eterogenei tra regioni e particolarmente alti. Sostengo che per stimare correttamente il tasso di letalità bisogna paragonare i decessi di oggi rispetto ai casi passati, altrimenti si ottiene una immagine distorta. Stimo il tasso di letalità comparando i decessi ad oggi con i casi totali a 6 giorni fa.
5. Gli **alti tassi di letalità italiani** ("l'anomalia Italia") possono essere spiegati da bassa propensione a testare: regioni con maggiore propensione a testare registrano tassi di letalità inferiori. Le regioni con piu' bassa propensione a testare sono generalmente quelle che registrano un maggior numero di casi.
6. Possibile stimare i **casi effettivi** comparando il tasso di letalità nei dati italiani con quello 'teorico' di Cina e Ferguson et al: i casi effettivi sono tra 460 mila e i 770 mila ad oggi. Questo significa che i casi rilevati sono solamente tra il 10% e il 20% del totale effettivo. Forti differenze tra regioni.
7. Importante rapportare i casi alla popolazione a livello regionale: in Lombardia, tra il 2,5% e il 4,2% della popolazione e' probabilmente infetta, ma anche nelle Marche il tasso varia dallo 0,94% al 1,56% come risultato di una popolazione inferiore. Importante monitorare le regioni piu' piccole, anche se il numero totale dei casi e' inferiore.
8. Sulla base dei casi *ad oggi*, e' previsto il raggiungimento di 11800 decessi totali tra 6 giorni. Questo e' dovuto al fatto che i decessi nei prossimi giorni possono essere stimati combinando il tasso di letalità nei dati regionali e il numero di casi ad oggi.

[†]Dipartimento di Economia, New York University. I dati verranno aggiornati giorno per giorno mano a mano che sono resi disponibili dal Ministero. Vi prego di scusare eventuali errori e di segnarmi a francesco.furno@nyu.edu.

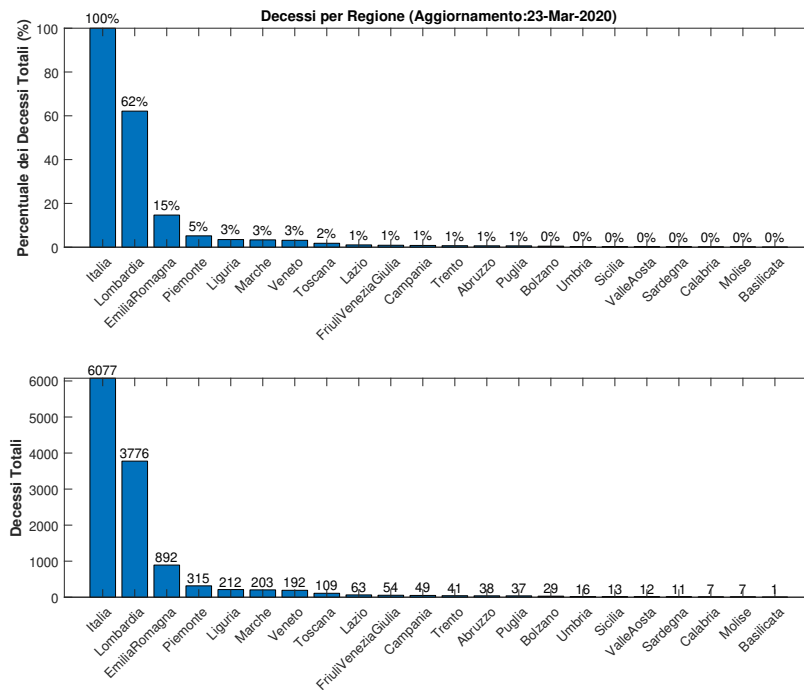
2 Statistiche Descrittive: Casi, Decessi e Tamponi

2.1 Casi e Decessi

I dati giornalieri sono volatili, e utilizzo una tecnica statistica per separare il trend.¹ Il trend di crescita percentuale sia dei casi sia dei decessi e' in miglioramento da giorni.

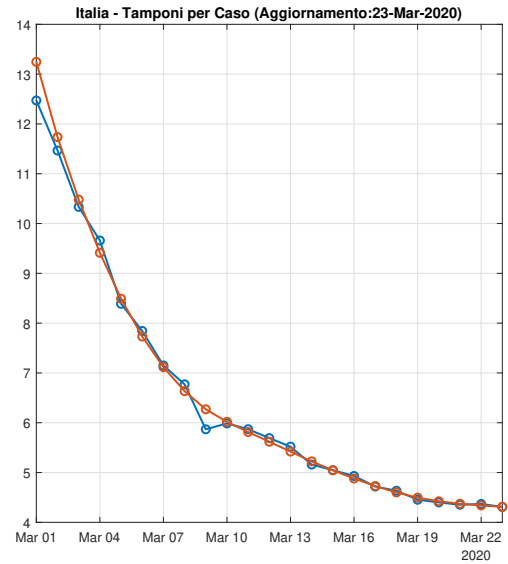
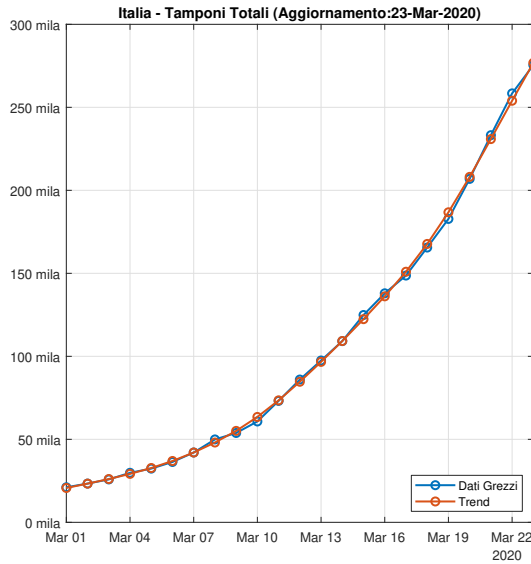


Il dato e' uniforme tra regioni (Appendice) ed e' in miglioramento dall'istituzione della zona rossa nazionale. Distribuzione dei recessi per regione:



¹Il trend e' estratto con un filtro HP, parametro di smoothing uguale a 5.

2.2 Tamponi

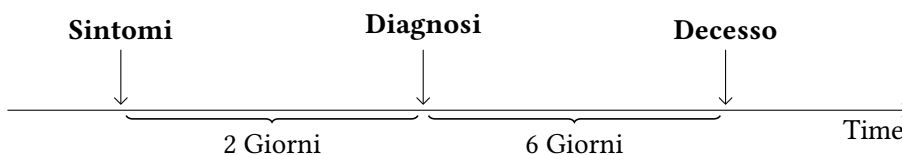


I tamponi totali eseguiti sono in costante crescita. Tuttavia, un dato preoccupante è l'evoluzione del numero medio di tamponi effettuati per ogni caso positivo. È in continua discesa in tutte le regioni (si veda Appendice) e a livello nazionale: segnale che la capacità di testare del sistema sanitario nazionale è al limite. È naturale che questo risulti in una *sottostima* del numero effettivo dei positivi.

3 Stime

3.1 Stima Tasso di Letalità

Il tasso di letalità nei dati italiani e' una variabile chiave per calcolare altre variabili importanti. La mia stima del tasso di letalità e' basata sull'idea che, in caso di decesso, intercorrono in media 8 giorni tra sintomi e decesso (fonte ISS). E' molto probabile pero' che la diagnosi si presenti con un paio di giorni di ritardo rispetto ai sintomi.

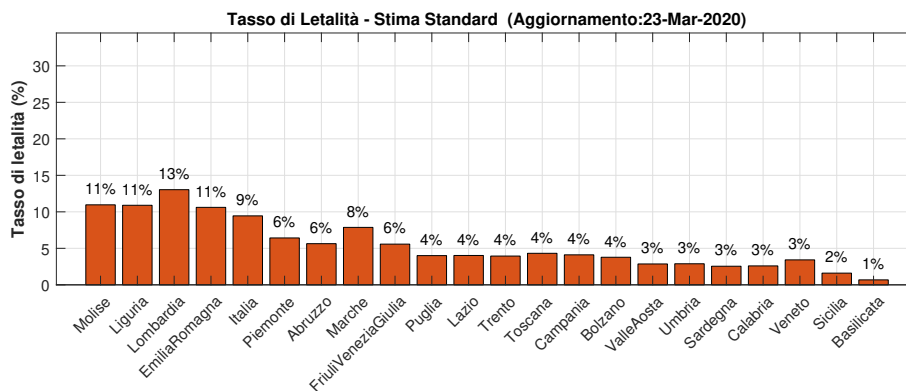
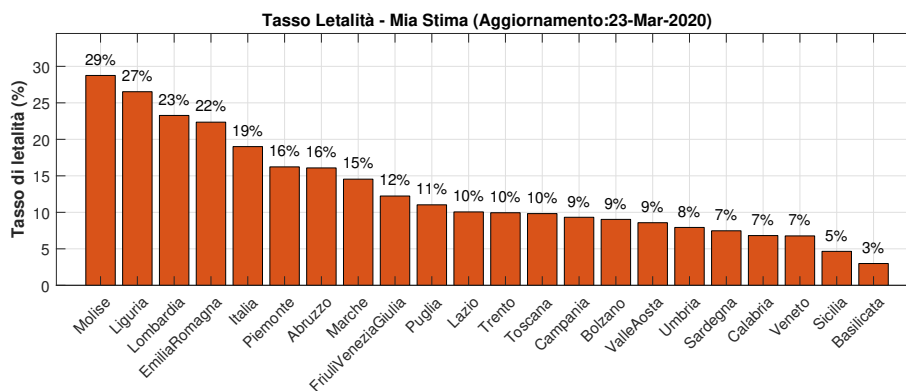


Sulla base di questa ipotesi, i decessi registrati oggi derivano in media dai casi di 6 giorni fa. Quindi:

$$\text{Letalità}_t = \frac{\text{Deceduti Totali}_t}{\text{Casi Totali}_{t-6}}$$

Il mio modo di calcolare il tasso di letalità si contrappone a una modalità utilizzata in molti mezzi di informazione. Il tasso di letalità viene definito come:

$$\text{Letalità}_t = \frac{\text{Deceduti Totali}_t}{\text{Casi Totali}_t}$$



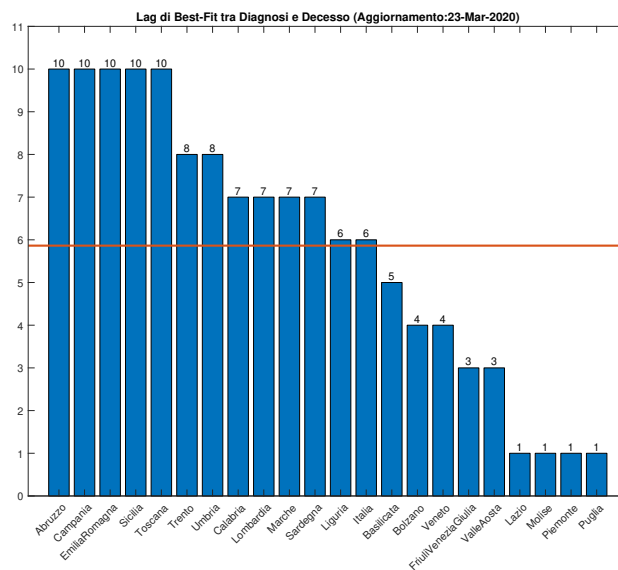
Ritengo che il modo ‘standard’ di calcolare il tasso di letalità sia ingannevole se si applica *durante* l’epidemia. Questo perché le fonti ministeriali ci dicono che intercorrono in media 8 giorni tra sintomi e decesso. Calcolare la letalità comparando i deceduti di oggi con i casi di oggi equivale ad assumere che il decesso avviene lo stesso giorno in cui avviene la diagnosi.

Scelgo di adottare 6 giorni in media tra il momento in cui viene diagnosticato il caso (e quindi venga conteggiato nei casi totali) e il momento in cui viene registrato il decesso, perché ipotizzo due giorni tra la comparsa dei sintomi e la diagnosi (un giorno per eseguire il test e un giorno per avere i risultati). Ovviamente potrei sbagliarmi, ed è per questo ricorro a una procedura statistica.

Eseguo la seguente regressione lineare per ogni valore del lag k :

$$\text{Deceduti Totali}_t = \alpha + \beta \cdot \text{Casi Totali}_{t-k} + \varepsilon$$

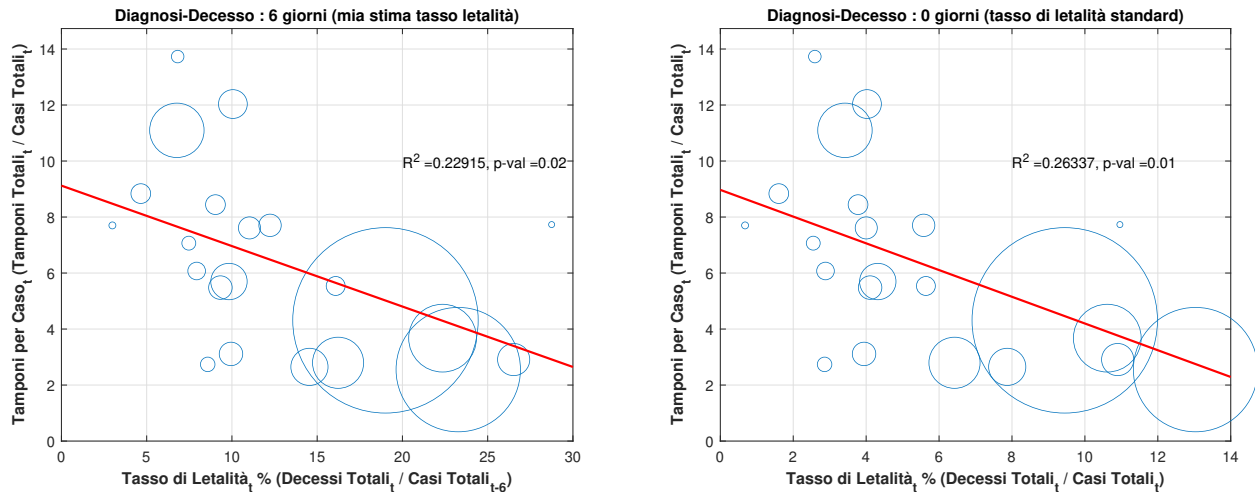
e scelgo il valore di k che massimizza la statistica R^2 , cioè l’abilità del regressore di prevedere la variabile dipendente nel campione. Ripeto l’esercizio regione per regione:



Ci sono significative differenze regionali, ma l’idea di fondo è che in media, il numero di giorni intercorsi tra diagnosi e decesso che massimizza la statistica R^2 è circa 6, confermando la mia ipotesi. Si noti che maggiore è il valore di k , maggiore è il tasso di letalità stimato, perché si comparano i deceduti totali al tempo t con il numero di casi totali al tempo $t - k$ (che è per definizione inferiore al crescere di k). Il tasso di letalità ottenuto con il metodo standard (cioè con $k = 0$) è quindi sempre il minor tasso di letalità stimabile con i dati disponibili *durante* un’epidemia.

3.2 Letalità e Tamponi: Confronto tra Regioni

Per cercare di razionalizzare gli alti tassi di letalità italiani, comparo la mia stima del tasso di letalità a livello regionale con la propensione a testare a livello regionale, misurata come ‘tamponi totali’ diviso per ‘casi totali’ regione per regione. Ogni cerchio rappresenta una regione, e la grandezza del cerchio rappresenta il numero totale di casi in quella regione. Per assicurarmi che il risultato non sia indotto dal mio metodo di stima della letalità, ripeto l’esercizio anche con la misura standard di letalità (deceduti oggi diviso per casi oggi).



La figura mostra che il tasso di letalità e' piu' elevato nelle regioni che hanno una propensione a testare minore. L'ipotesi naturale e' che quando si testa di meno, si testano casi piu' gravi e non si testano i casi meno gravi, quindi il tasso di letalità sale. Inoltre, le regioni con piu' casi (rappresentato da un cerchio piu' grande), sono quelle che testano di meno (probabilmente in maniera strategica per preservare risorse sanitarie, ma forse anche perche' faticano a stare al passo con l'epidemia).

3.3 Stima Casi Effettivi

I tassi di letalità presenti nei dati del ministero sono eccezionalmente piu' alti dei tassi di letalità suggeriti sia dalla Cina sia dal rapporto dell'Imperial College di Ferguson et al., riportati nella tabella sottostante e applicati alla struttura demografica italiana.

Fascia Età	Letalità % (Dati Cina)	Letalità % (Dati Ferguson et al)	Popolazione Italiana %
0-9	0,00	0,00	8,43
10-19	0,20	0,01	9,56
20-29	0,20	0,03	10,27
30-39	0,20	0,08	11,72
40-49	0,40	0,15	15,31
50-59	1,30	0,60	15,49
60-69	3,60	2,20	12,16
70-79	8,00	5,10	9,88
80+	14,80	9,30	7,17
Complessiva	2,62	1,57	

Data la struttura demografica italiana, i tassi di letalità 'teorici' dovrebbero essere per l'Italia: 2.62% (dati cinesi), 1.57% (dati Ferguson et al.). Ipotizzando che la vera letalità in Italia coincida con questi valori, posso stimare quindi il numero di casi totali effettivi in Italia e regione per regione che sarebbero necessari a generare il tasso di letalità teorico. Infatti:

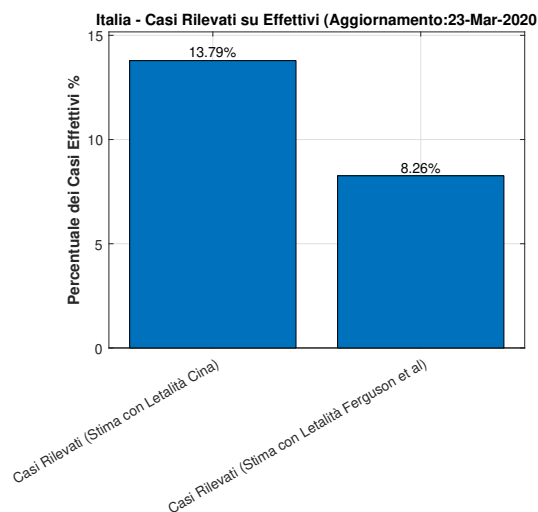
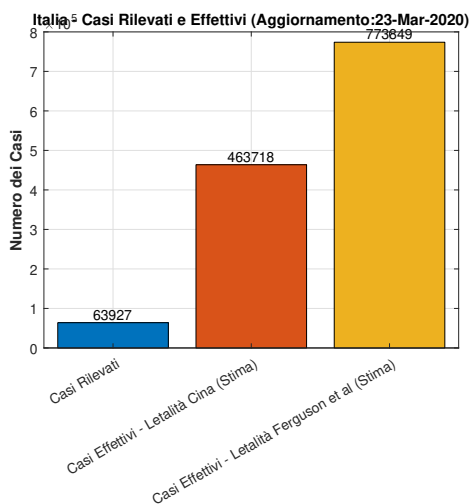
$$\text{Letalità teorica} = \frac{\text{Deceduti}}{\text{Casi Effettivi}}, \quad \text{Letalità dati italiani} = \frac{\text{Deceduti}}{\text{Casi Rilevati}}$$

Abbiamo quindi che:

$$\frac{\text{Letalità dati italiani}}{\text{Letalità teorica}} = \frac{\text{Casi Effettivi}}{\text{Casi Rilevati}}$$

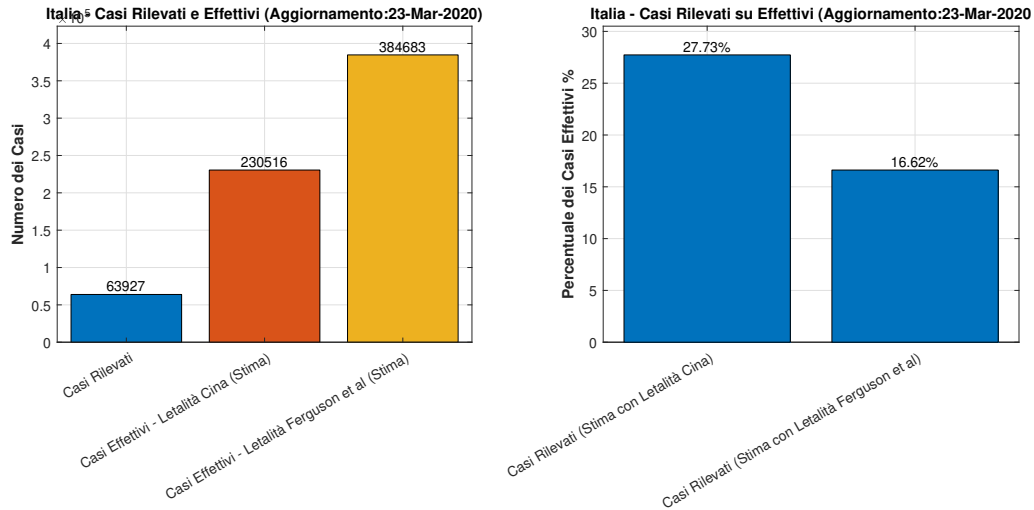
e che si possono stimare i casi effettivi come segue:

$$\text{Casi Effettivi} = \text{Casi Rilevati} \times \frac{\text{Letalità dati italiani}}{\text{Letalità teorica}}$$



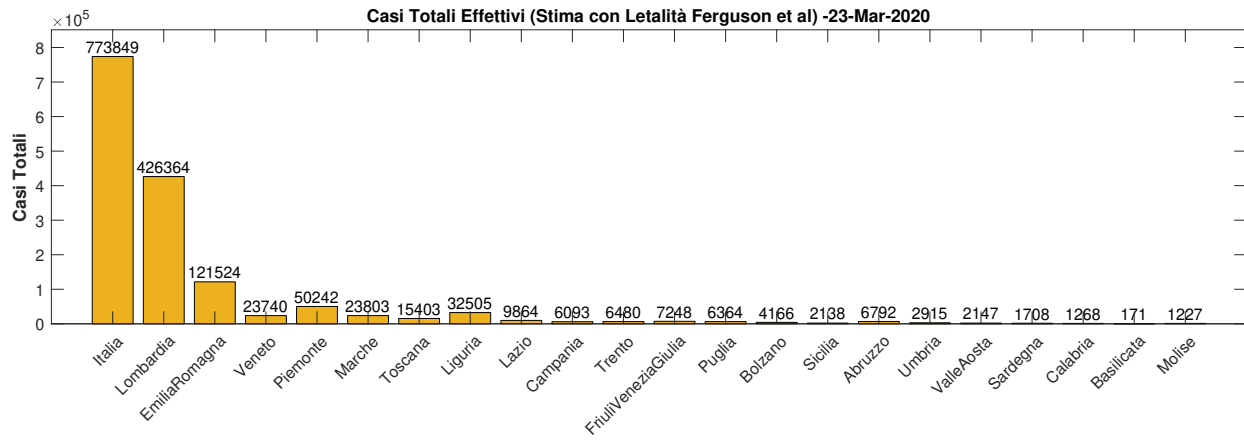
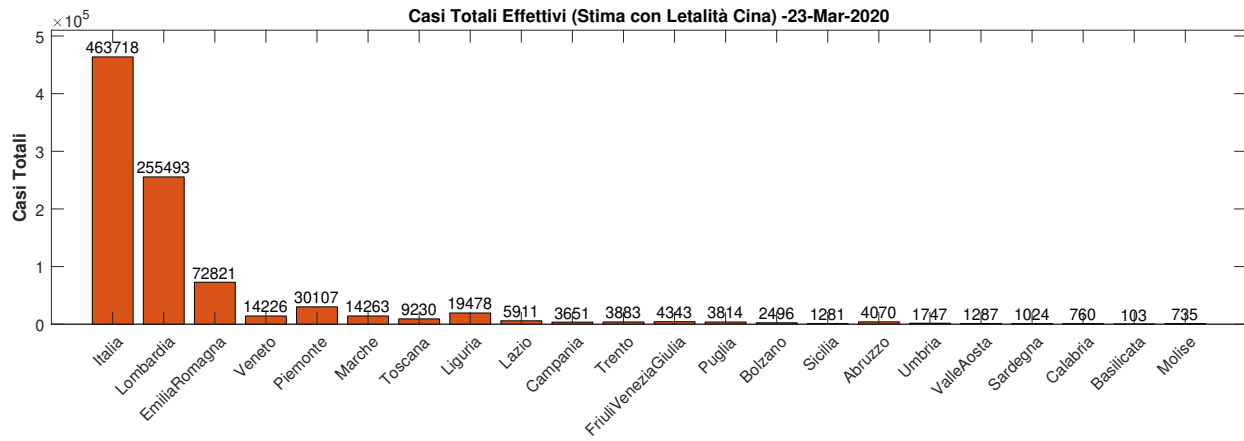
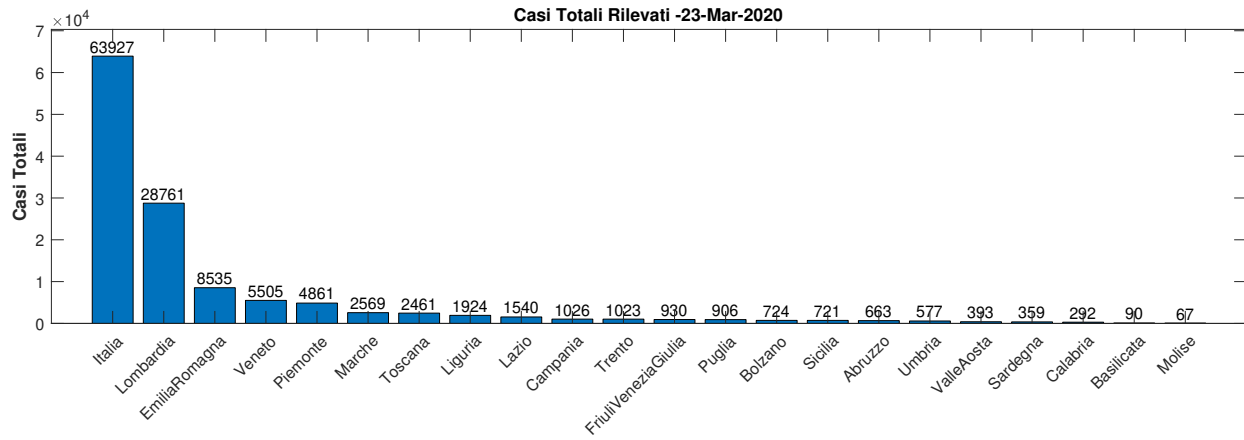
Si noti che, maggiore il tasso di letalità italiano stimato, maggiore il numero dei casi effettivi. Quindi, il numero di casi effettivi varia enormemente se si usa il tasso di letalità standard o quello da me stimato. Difendo la mia scelta di stimare il tasso di letalità comparando i decessi ad oggi con i casi di 6 giorni fa sulla base di: 1) informazioni del Ministero; 2) selezione attraverso procedura statistica.

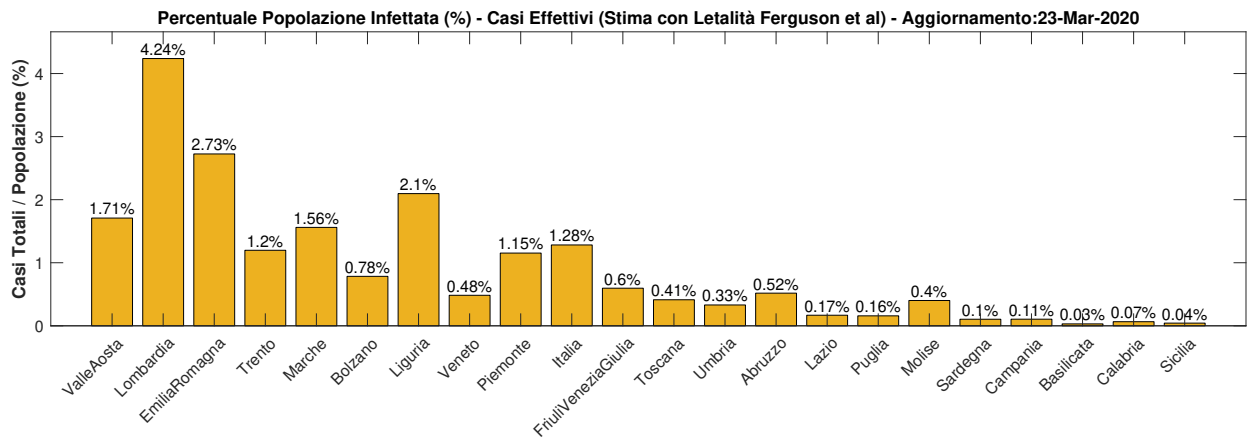
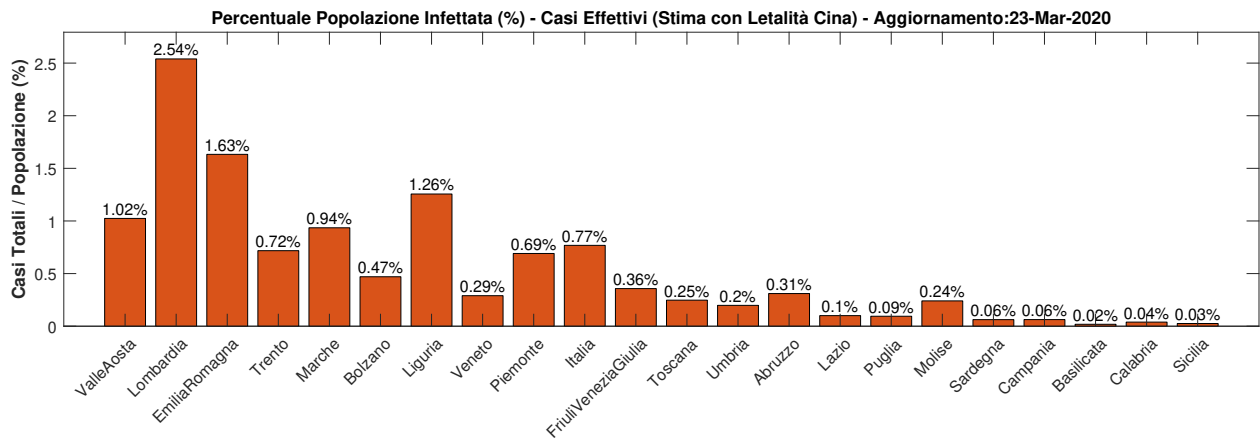
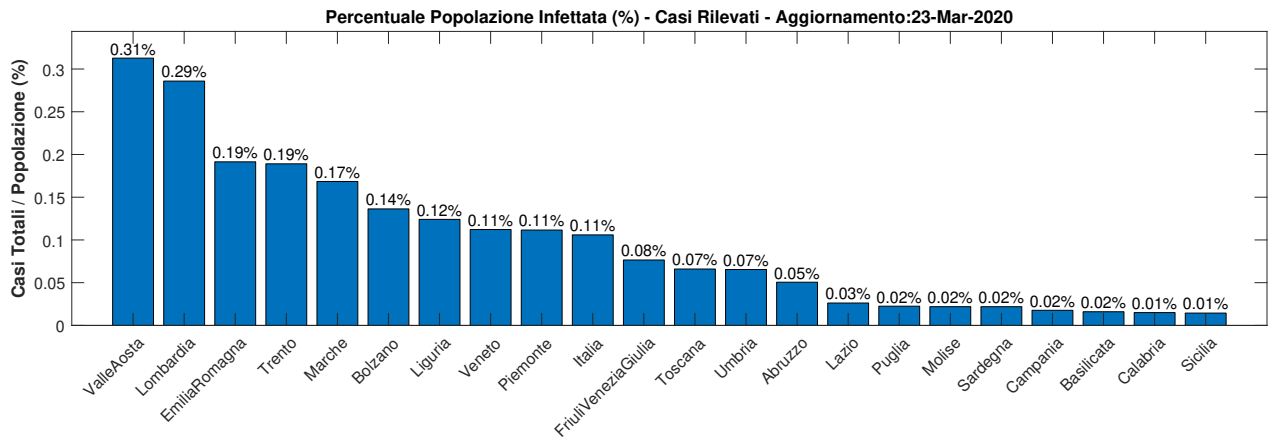
Ripeto pero' per completezza l'esercizio usando i tassi di letalità ottenuti con il metodo standard (deceduti oggi diviso per casi oggi):



Poiche' il tasso di letalità calcolato con il metodo 'standard' e' sempre il minor tasso di letalità stimabile, ne risulta che i casi effettivi stimati con esso siano il *limite inferiore* dei casi effettivi. In altre parole, il numero di casi effettivi e' sempre *superiore* alla stima riportata nella figura sopra.

Le stime regione per regione che seguono sono ottenute utilizzando la mia stima del tasso di letalità:



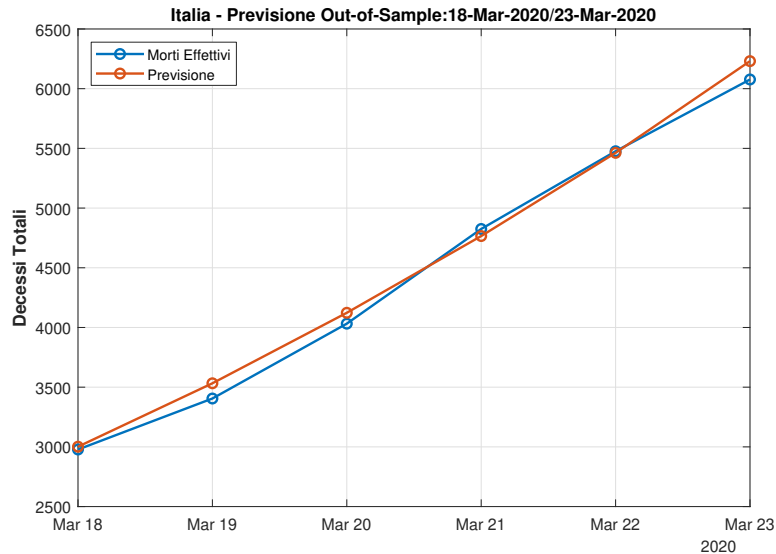


Dividere per la popolazione aiuta a capire quali sono le regioni piu' colpite, tenendo conto che alcuni regioni sono piu' grandi (in termini di popolazione residente) di altre. E' importante affiancare al numero totale dei casi, il numero totale dei casi diviso per la popolazione della regione.

3.4 Decessi

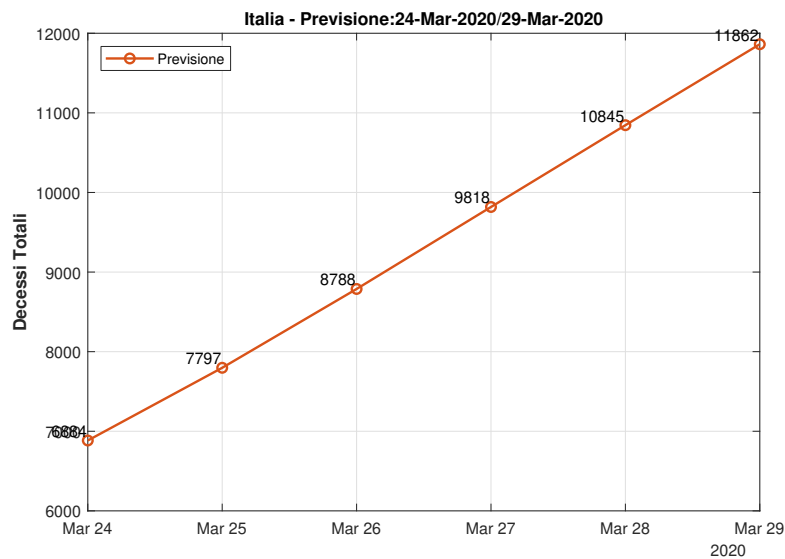
3.5 Validazione

Se la stima dei tassi di letalità è corretta, dovrei essere in grado di fornire una previsione corretta dei decessi a 6 giorni. Per prima cosa, mi limito ai dati disponibili 6 giorni fa e provo a prevedere i decessi negli ultimi 6 giorni. La figura sottostante riporta la stima aggregata, mentre le stime regionali sono in Appendice. Le previsioni sono accurate per regioni con molti casi, e meno accurate per regioni con pochi casi.



3.6 Previsione Prossimi 6 Giorni

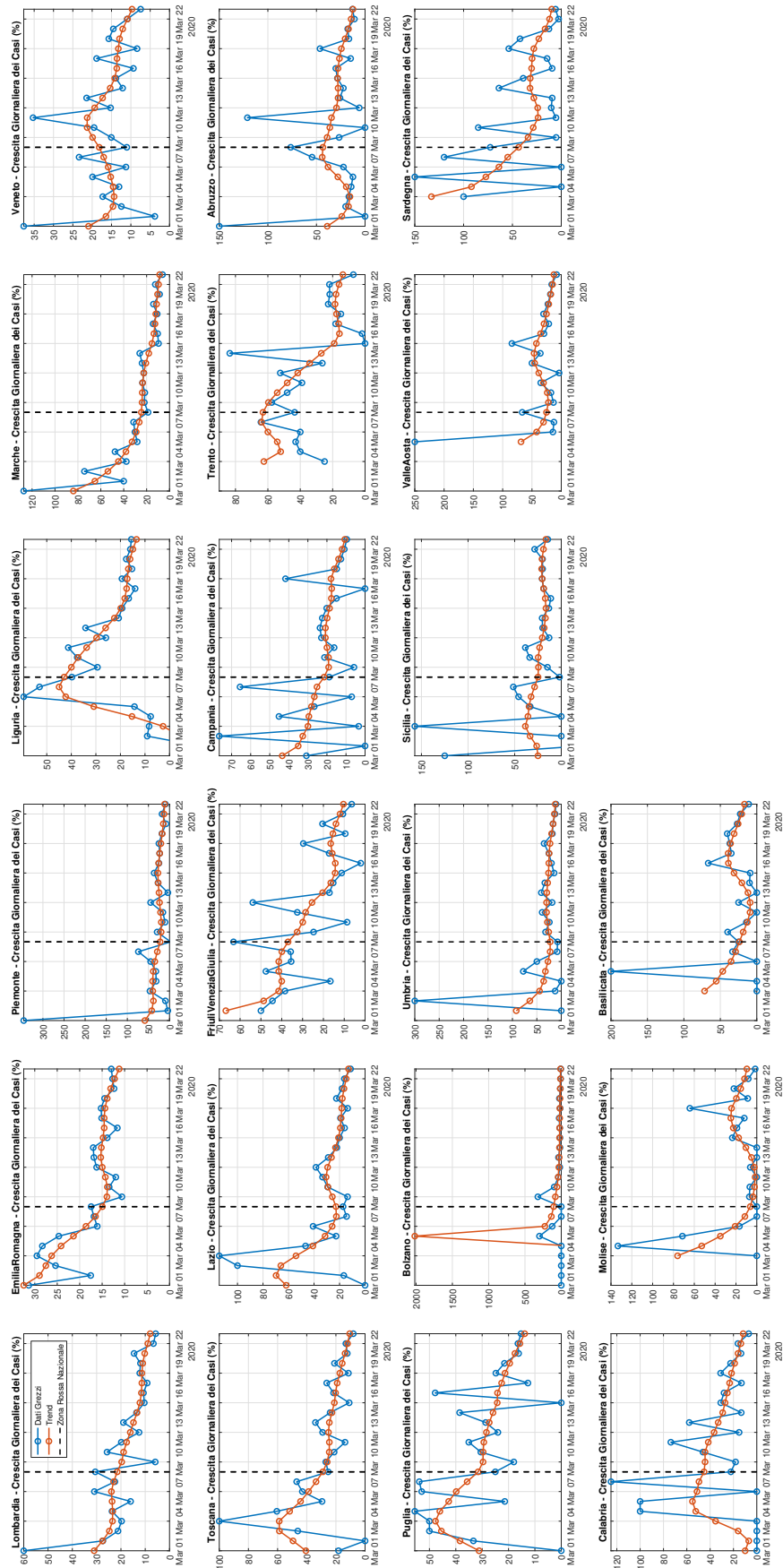
Sulla base di casi ad oggi, è possibile prevedere il numero di decessi totali per i prossimi 6 giorni:²



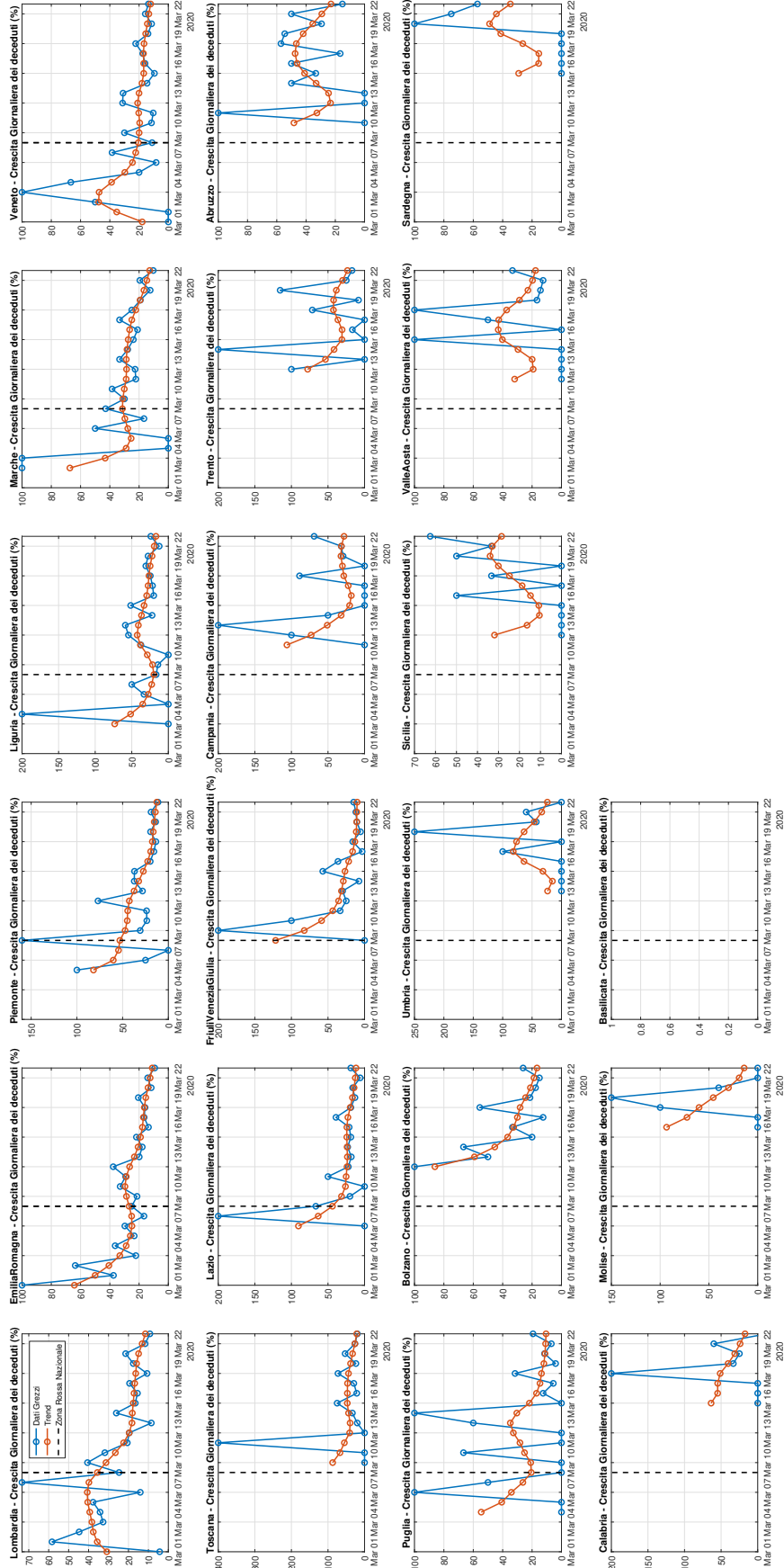
²Si noti che questa è più simile a una *nowcast* che a una *forecast*.

Appendice

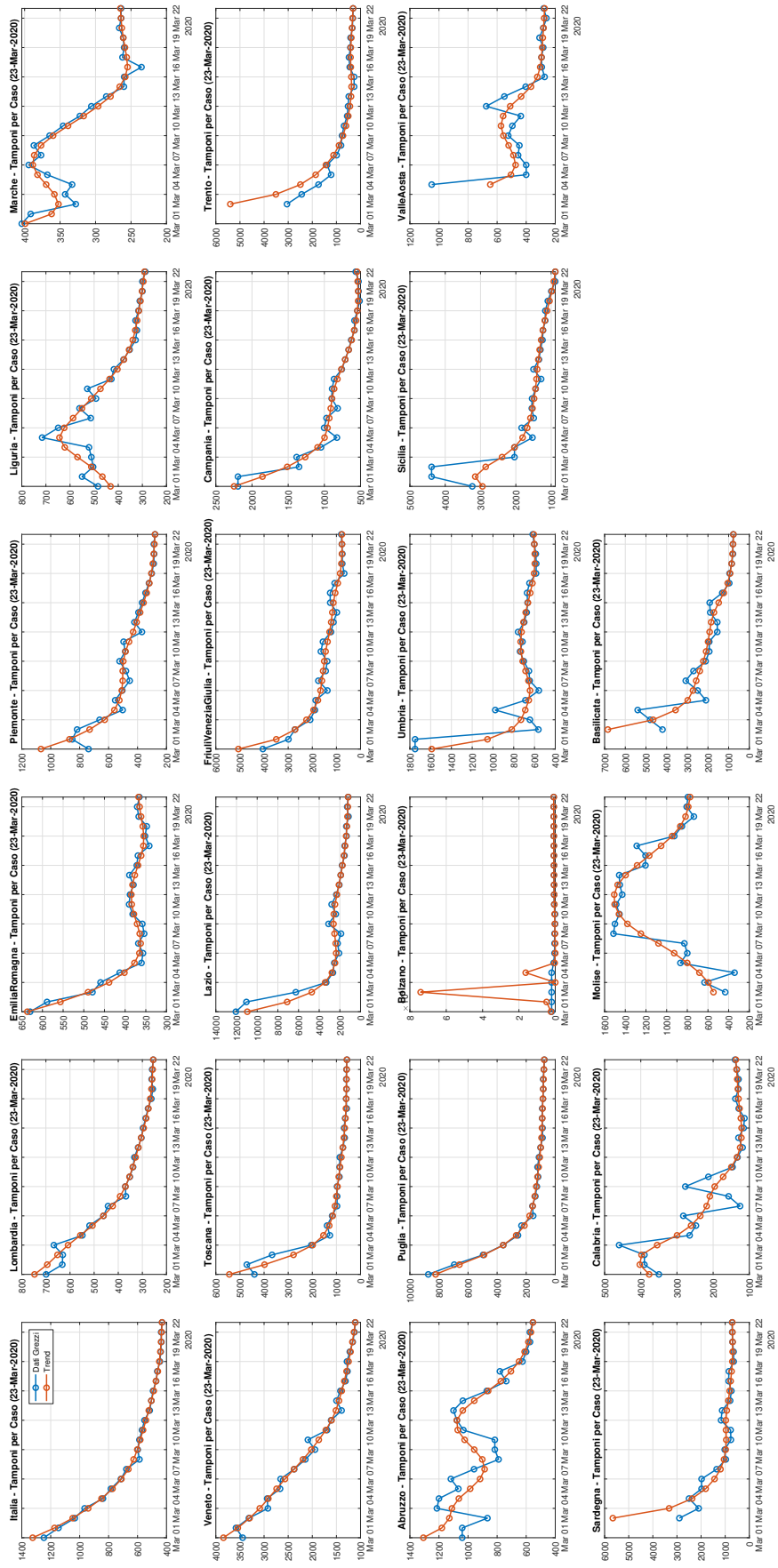
Evoluzione Casi - Regioni



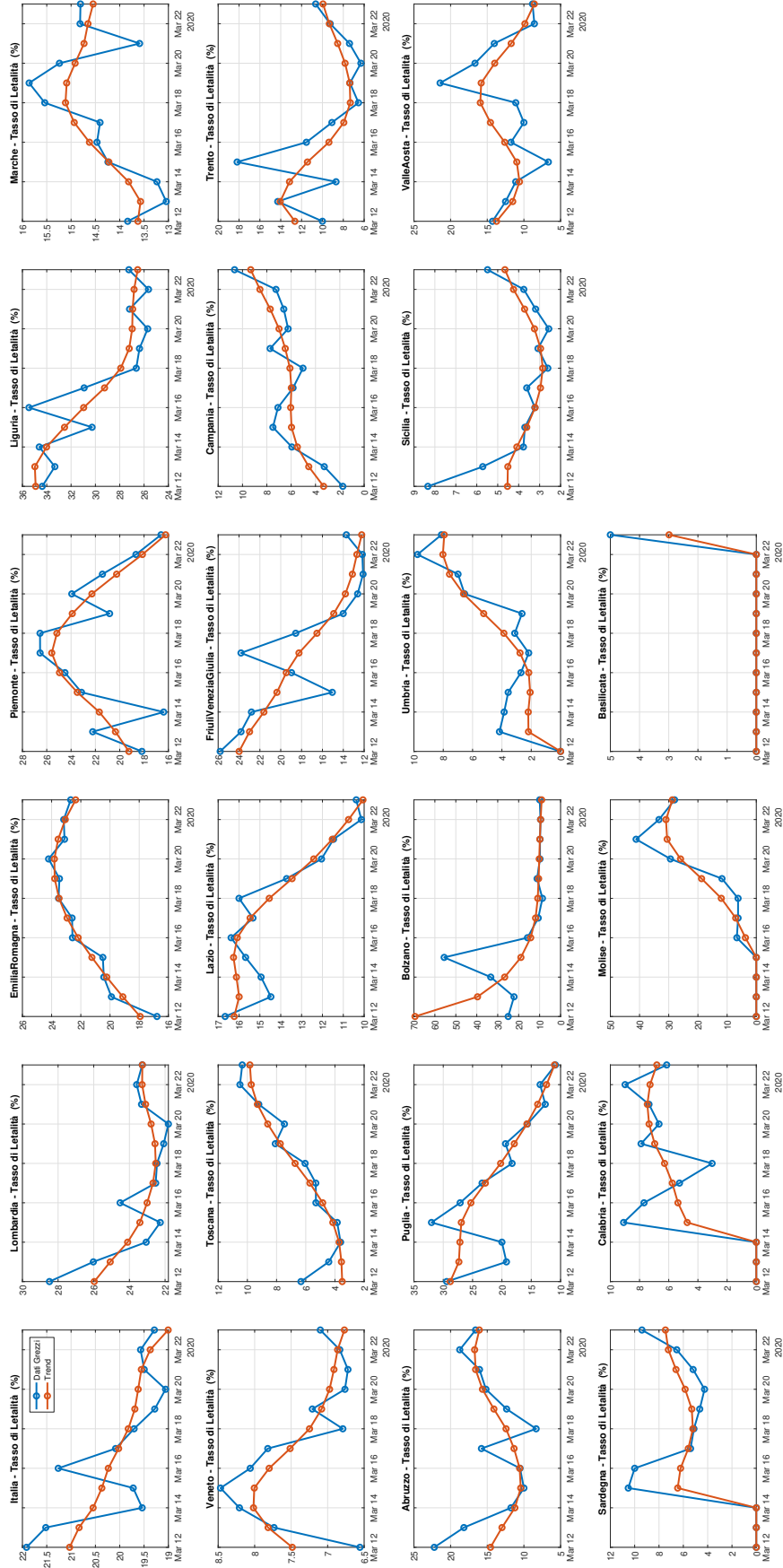
Evoluzione Deceduti - Regioni



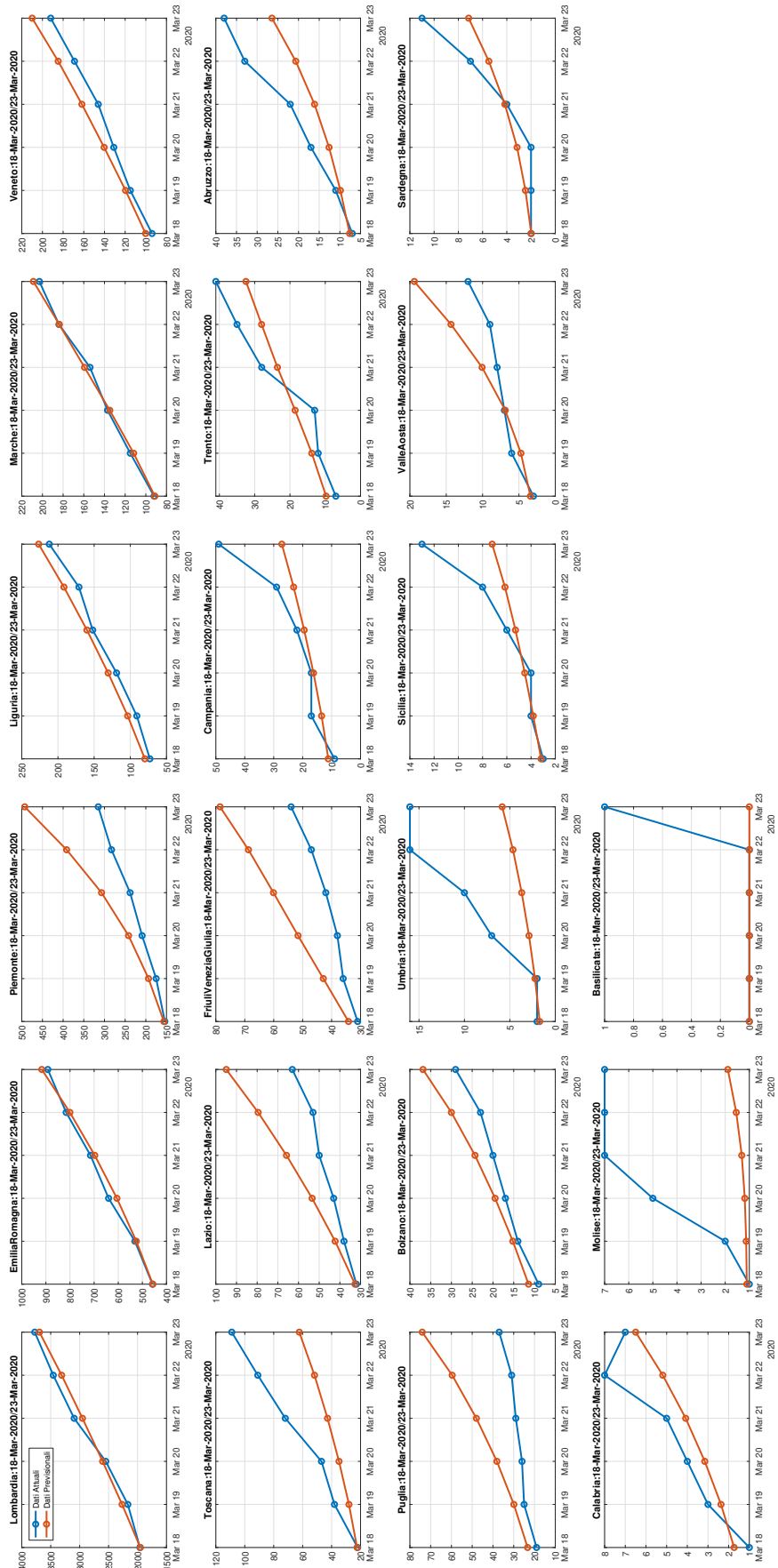
Evoluzione Tamponi per Caso - Regioni



Evoluzione Tasso di Letalità Stimato - Regioni



Previsione Decessi - Validazione - Regioni



Previsione Decessi - Prossimi 6 Giorni - Regioni

