

Un'analisi sulla diffusione internazionale del COVID-19

basata su dati e strumenti interattivi del sito [COVID-19](#) dell'Istituto CNR-SPIN

Mario Barra¹, Emiliano Di Gennaro^{1,2}, Fabio Miletto Granozio¹

¹CNR-SPIN, Compl. Univ. di Monte S. Angelo, Via Cintia, Napoli, Italia

²Dip. di Fisica "E. Pancini", Università di Napoli "Federico II", Compl. Univ. di Monte S. Angelo, Via Cintia, Napoli, Italia

Abstract

Dopo aver duramente colpito la Cina nei primi mesi del 2020, l'epidemia di COVID-19 è esplosa in Italia e in tutte le maggiori nazioni occidentali. In questo contributo, vengono confrontati gli andamenti temporali del numero di contagi e decessi relativi a Stati Uniti, Spagna, Italia, Germania, Francia e Regno Unito. La difficoltà di confrontare i diversi livelli di evoluzione dell'epidemia nelle varie nazioni è stata affrontata ricorrendo principalmente all'introduzione di opportune traslazioni temporali oppure considerando grafici in cui la variabile tempo è resa implicita. In questo modo, è possibile evidenziare l'esistenza di una legge di scaling che, mediante l'uso di shift temporali e di fattori moltiplicativi, consente di far sovrapporre le curve con una precisione sorprendente. Anche i grafici in cui la variabile tempo non è presentata direttamente sono uno strumento efficace per confrontare dati relativi a nazioni in cui l'epidemia è a diverso stadio evolutivo. In generale, questi metodi alternativi di analisi possono fornire dettagli nuovi sull'evoluzione passata e, verosimilmente, su quella futura della pandemia a livello internazionale. Le metodologie qui riportate sono accessibili a tutti gli utenti interessati, attraverso il sito [COVID-19](#) dell'Istituto CNR-SPIN.

Introduzione

Dopo aver duramente colpito la Cina nei primi mesi del 2020, l'epidemia COVID-19 si è rapidamente diffusa in Europa a partire dalla seconda metà di febbraio. L'Italia è stato il primo paese occidentale interessato, con, attualmente (i dati sono riferiti al 23 di Aprile), circa 190.000 casi ufficialmente confermati e un numero di decessi che superano i 25.000. Dopo l'Italia, l'epidemia ha iniziato la sua diffusione in tutte le maggiori nazioni europee, con un ritardo variabile da una ad oltre due settimane. In un primo contributo pubblicato il 13 Marzo su [scienzairete](#)² e in un successivo articolo³, uno degli autori (fmg) ha evidenziato come un efficace confronto tra le epidemie di differenti nazioni richieda il ricorso ad opportune traslazioni (shift) sull'asse dei tempi. Fu osservato come, in assenza di un immediato cambio nel trend di crescita, sia gli Stati Uniti che la Spagna fossero proiettate verso uno scenario peggiore della stessa Italia, in termini di numero totale di contagi confermati. Come previsto, il numero dei casi di queste due nazioni supera oggi quello dell'Italia. Al fine di sviluppare ulteriormente questo approccio, gli autori di questo contributo hanno recentemente realizzato il sito [COVID-19](#) dell'Istituto CNR-SPIN. Il sito è dedicato sia alla situazione dell'epidemia in Italia,

con un focus sui dati delle singole regioni, che al contesto internazionale. L'obiettivo principale è quello di coniugare due approcci che solo raramente sono coesistiti finora:

- a) l'approccio tipico del mondo scientifico, teso a esplorare scelte meno ovvie nell'osservazione e rappresentazione delle variabili di interesse, alla continua ricerca di correlazioni tra queste;
- b) l'approccio puramente informatico, adottato oggi da molteplici siti web in ogni nazione, basato sulla rappresentazione di dati grezzi che si aggiornano automaticamente.

Conseguentemente, il nuovo sito web permette all'utente di visualizzare l'evoluzione della pandemia attraverso grafici "auto-aggiornanti" che vanno però oltre l'uso delle classiche curve "casi in funzione del tempo" o delle "mappe" diventate oggi onnipresenti in internet. Al visitatore vengono offerte, inoltre, varie opzioni per modificare i grafici mediante il controllo di alcuni parametri liberi. E' possibile così concentrare la propria attenzione su specifici set di nazioni oppure verificare la presenza di mutue correlazioni.

Questo articolo è dedicato in particolare al confronto della diffusione del COVID-19 nelle nazioni che presentano, al momento, il maggior numero di casi confermati: Italia (IT), Stati Uniti (US), Germania (DE), Spagna (ES), Regno Unito (UK) e Francia (FR). Nel complesso, queste sei nazioni ospitano attualmente il 65% circa dei casi rilevati in tutto il Mondo.

Confronto tra i dati basato su scaling delle curve

Come già detto, il semplice confronto tra il numero dei casi confermati (contagi) e dei decessi, con le relative rappresentazioni in funzione del tempo, consente di ottenere solo informazioni limitate se i relativi ritardi temporali tra le varie epidemie nazionali non vengono opportunamente considerati.

L'approccio perseguito nel sito [COVID-19](#) è duplice: modificare i dati tenendo in conto dei relativi shift sull'asse dei tempi oppure concentrarsi su grafici in cui il tempo non è esplicitamente rappresentato.

La Figura 1 fornisce un quadro generale, basato sui dati del 20 Aprile, della diffusione dell'epidemia COVID-19 in IT (60 milioni di abitanti), US (328 milioni di abitanti), DE (83 milioni di abitanti), ES (47 milioni di abitanti), UK (67 milioni di abitanti), e FR (67 milioni di abitanti). I dati numerici sono espressi per milioni di abitanti. Al fine di attenuare parzialmente il rumore legato sia a fluttuazioni intrinseche che a errori nella raccolta o presentazione dei dati, è stata applicata una media Gaussiana con $\sigma = 3$ giorni su punti adiacenti.

Per ogni nazione, le Figure 1a e 1b riportano il numero dei casi confermati e dei decessi per milione di abitanti. Come consuetudine, questi grafici sono in scala semi-logaritmica così garantire una visione più chiara dell'intera evoluzione osservata durante un periodo di tempo, comunque, limitato (circa 8 settimane). La fase iniziale della crescita epidemica è stata molto rapida e, in questa regione, le curve possono essere ben descritte da una sequenza di andamenti esponenziali, caratterizzati da costanti di tempo leggermente

differenti. In questo modo, i “tempi di raddoppio” (ossia il numero di giorni necessari per osservare un valore doppio dei casi confermati o dei decessi) può essere stimato facilmente^{3,4}. Come detto, l’Italia è stata la prima nazione europea in cui il COVID-19 ha iniziato la sua diffusione. In Spagna, Francia e Germania, la crescita dell’epidemia è esplosa dopo circa una settimana con, inizialmente, valori molto simili di casi confermati per milione di abitanti. Il ritardo della diffusione è stato ancora maggiore per US e UK.

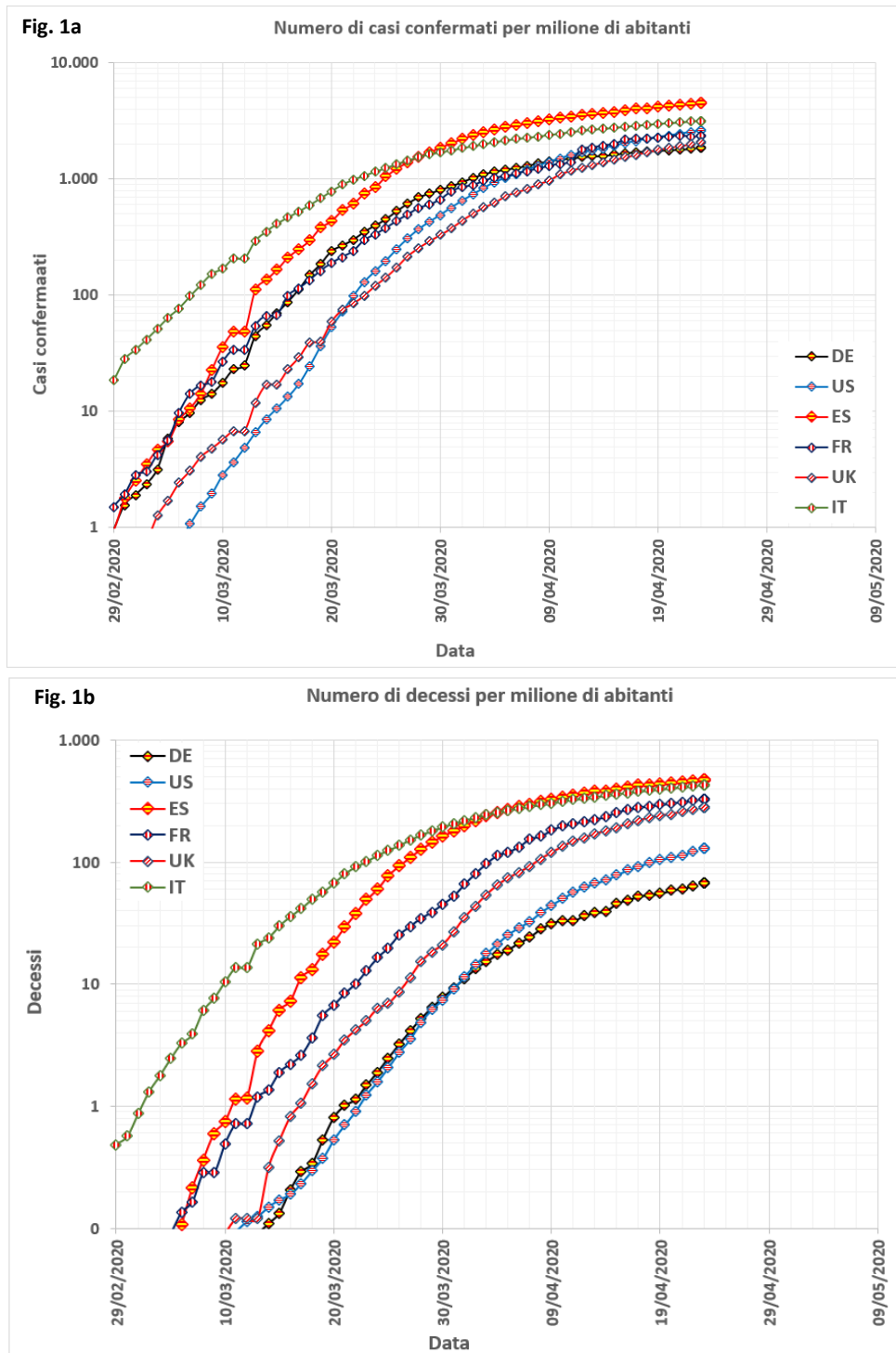


Fig.1: andamento dei (a) casi confermati e (b) decessi per milione di abitanti nelle nazioni considerate.

I dati cumulativi in funzione del tempo, riportati in Figura 1, possono essere approssimativamente modellizzati usando specifiche funzioni parametriche^{4,5} (es. sigmoide, logistica, scaling algebrico), capaci potenzialmente di descrivere sia il regime di forte crescita iniziale sia la fase finale, dove i dati seguono un comportamento crescente sempre più attenuato fino a giungere alla completa saturazione al termine dell'epidemia. Il nostro scopo qui è stato quello di evidenziare la presenza di similitudini tra le curve riferite alle diverse nazioni, che si traducono nella sovrapposizione, osservata empiricamente, delle risposte "scalate".

Utilizzando gli strumenti disponibili sul sito di CNR-SPIN, è facile riconoscere infatti che le diverse curve nazionali, relative ai casi confermati o ai decessi, mostrano una significativa sovrapposibilità qualora si proceda ad un'opportuna procedura di scaling, ottenuta per ognuna con uno shift sull'asse dei tempi e la moltiplicazione per un fattore sull'asse verticale.

Questo è stato l'approccio adottato per ottenere i grafici nelle Figure 2a,b, dove per UK e US, le nazioni che sperimentano uno stadio più iniziale dell'epidemia, le curve sono sincronizzate sul calendario effettivo. Per tutti gli altri paesi, le curve sono invece traslate avanti nel tempo di un adeguato numero di giorni.

La sovrapposizione tra le curve è stata ottenuta con semplici tentativi "manuali", senza ricorrere cioè ad alcuna procedura di ottimizzazione matematica, e può essere verificata indipendentemente, e monitorata per il futuro, da qualsiasi utente del sito [COVID-19](#) dell'Istituto CNR-SPIN.

Se si considera la grande variabilità tra nazione e nazione, sia relativamente all'evoluzione "reale" dell'epidemia sia per quanto concerne le procedure di raccolta dati, la similitudine tra le curve considerate è quasi inattesa e suggerisce la presenza di un qualche comportamento di base universale, sebbene non ancora identificato. Un accordo peggiore è ottenuto per le regioni delle curve corrispondenti agli stadi iniziali dell'epidemia. Questo risultato non è sorprendente dal momento che i dati nelle prime fasi sono caratterizzati sia da una statistica limitata che da ampie differenze tra nazione e nazione degli approcci seguiti, sia per il confinamento del virus sia per la raccolta dei dati, in una fase dove la piena "consapevolezza" dell'epidemia non era stata ancora raggiunta. Nella Tabella I, sono riassunti i fattori di scala e gli shift temporali utilizzati per ottenere i grafici in Fig.2, attribuendo convenzionalmente il fattore di scala 1 all'Italia.

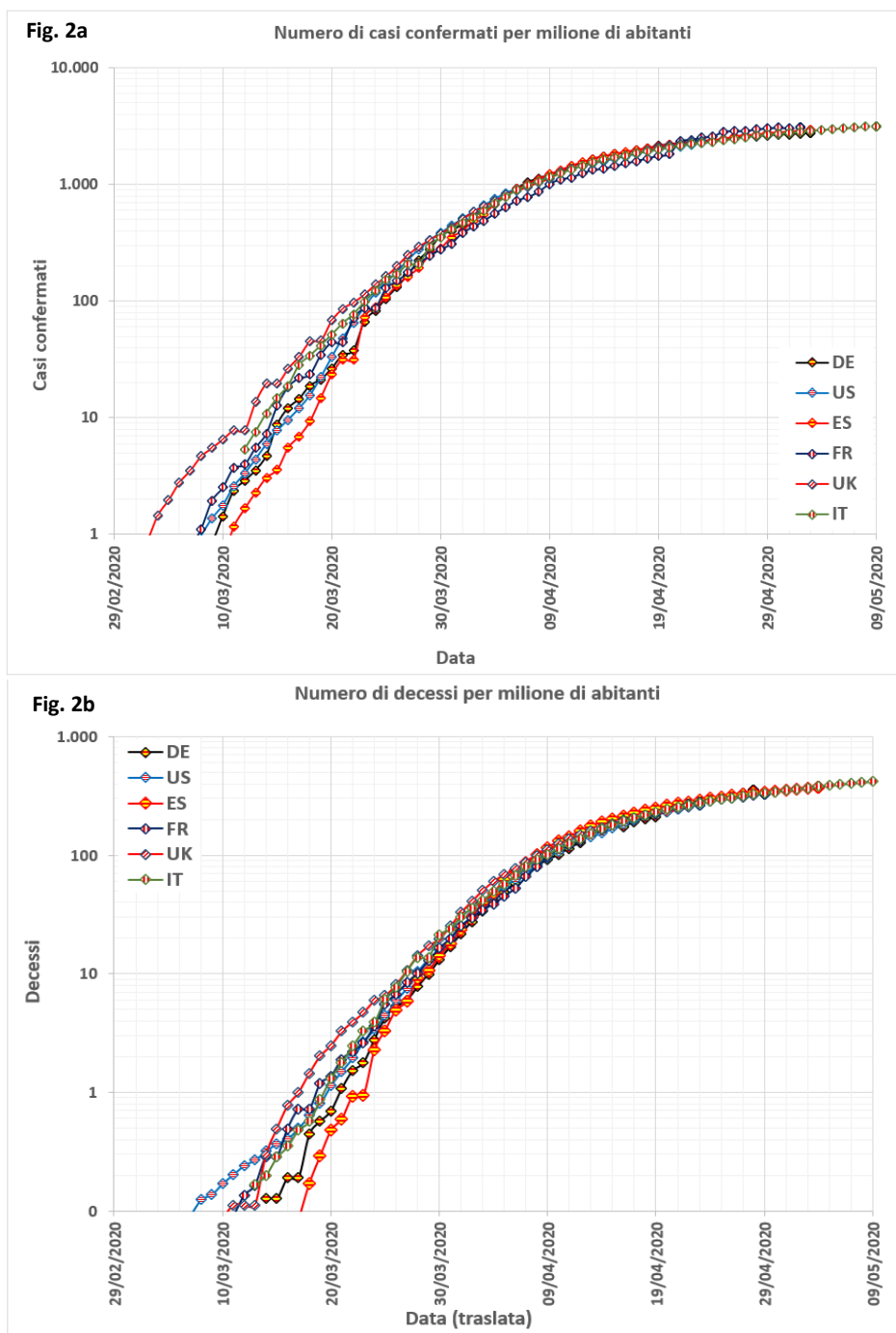


Fig.2: Sovrapposizione delle curve relative ai a) casi confermati e b) decessi per milione di abitanti.

La presenza di tempi di ritardo leggermente differenti da nazione a nazione nelle curve dei casi confermati e in quelle dei decessi (si veda, ad esempio, il caso di Germania e Francia), è a prima vista sorprendente. Dall'altro lato, mentre il valore medio del tempo di ritardo infezione-decesso è presumibilmente lo stesso dappertutto, il tempo medio infezione-diagnosi può dipendere dai differenti protocolli e dalla diversa capacità di testing di ogni nazione. Nel complesso, è ragionevole assumere quindi che il tempo medio

diagnosi-morte possa essere caratteristico di ogni nazione, e che lo shift temporale tra le curve dei casi confermati e dei decessi potenzialmente possa risultare differente.

Nazione	Fattore di scala	Shift temporali (giorni)	Fattore di scala	Shift temporali (giorni)
	Casi confermati		Decessi	
GERMANY	1.50	9	5.25	5
US	0.90	1	2.15	0
SPAIN	0.65	10	0.80	11
FRANCE	1.30	9	1.00	6
UK	1.15	0	0.95	0
ITALY	1.00	16	1.00	16

Tab. 1: Fattori di scala e shift temporali adottati per le curve in Fig.2

Possiamo ora chiederci se l'approccio fin qui descritto e i parametri riassunti in Tabella 1 abbiano un qualche potenziale predittivo. Pur astenendoci da qualsiasi previsione esplicita, vogliamo limitarci qui a commentare alcune delle cifre che potrebbero essere raggiunte, qualora i trend fin qui emersi, fossero mantenuti anche nelle prossime settimane (ossia se le curve scalate mostrate in Figura 2 continuassero a sovrapporsi anche in futuro). All'interno di questo quadro concettuale, ancora da testare, otterremo alcune indicazioni confrontando semplicemente specifiche coppie di nazioni, come negli esempi a seguire.

Casi confermati: Dato il fattore di scala (0.90) and il tempo di shift (15 giorni relativi) utilizzati per sovrapporre l'andamento degli Stati Uniti a quello dell'Italia, possiamo dedurre che, in due settimane o poco più, gli Stati Uniti avranno circa lo stesso numero di casi confermati per milione di abitanti (oltre 3000), che l'Italia ha ora, o poco più. Alla stessa data, il numero dei casi normalizzato alla popolazione del Regno Unito dovrebbe essere di circa il 20% inferiore a quello degli Stati Uniti e dell'Italia.

Decessi: Nello stesso periodo, il Regno Unito dovrebbe aver raggiunto o anche superato il numero di decessi (per milione di abitanti) che l'Italia ha ora (oltre 400), mentre questo valore dovrebbe essere circa la metà negli Stati Uniti.

Più in generale, nello scenario ipotizzato, che cioè le curve scalate conservino la loro sovrapponibilità per un tempo ancora più lungo, il rapporto tra il numero finale dei conteggi per le diverse nazioni corrisponderebbe, alla "fine dell'epidemia", agli inversi dei rapporti tra i fattori di scala in Tabella I. La Spagna (fattore di scala 0.80) pagherebbe così un costo in vite umane (normalizzate alla popolazione) di circa il 20% maggiore di quello dell'Italia; in Germania (fattore di scala 5.25) questo valore sarebbe cinque volte più basso, e così via. Il termine "fine dell'epidemia" è tenuto tra virgolette perché non è chiaro se il trend di discesa dei nuovi casi

sarà mantenuto nelle settimane a venire: ciò richiederebbe che il completo isolamento sociale (*lockdown*) rimanesse attivo ancora per molte settimane o probabilmente per mesi.

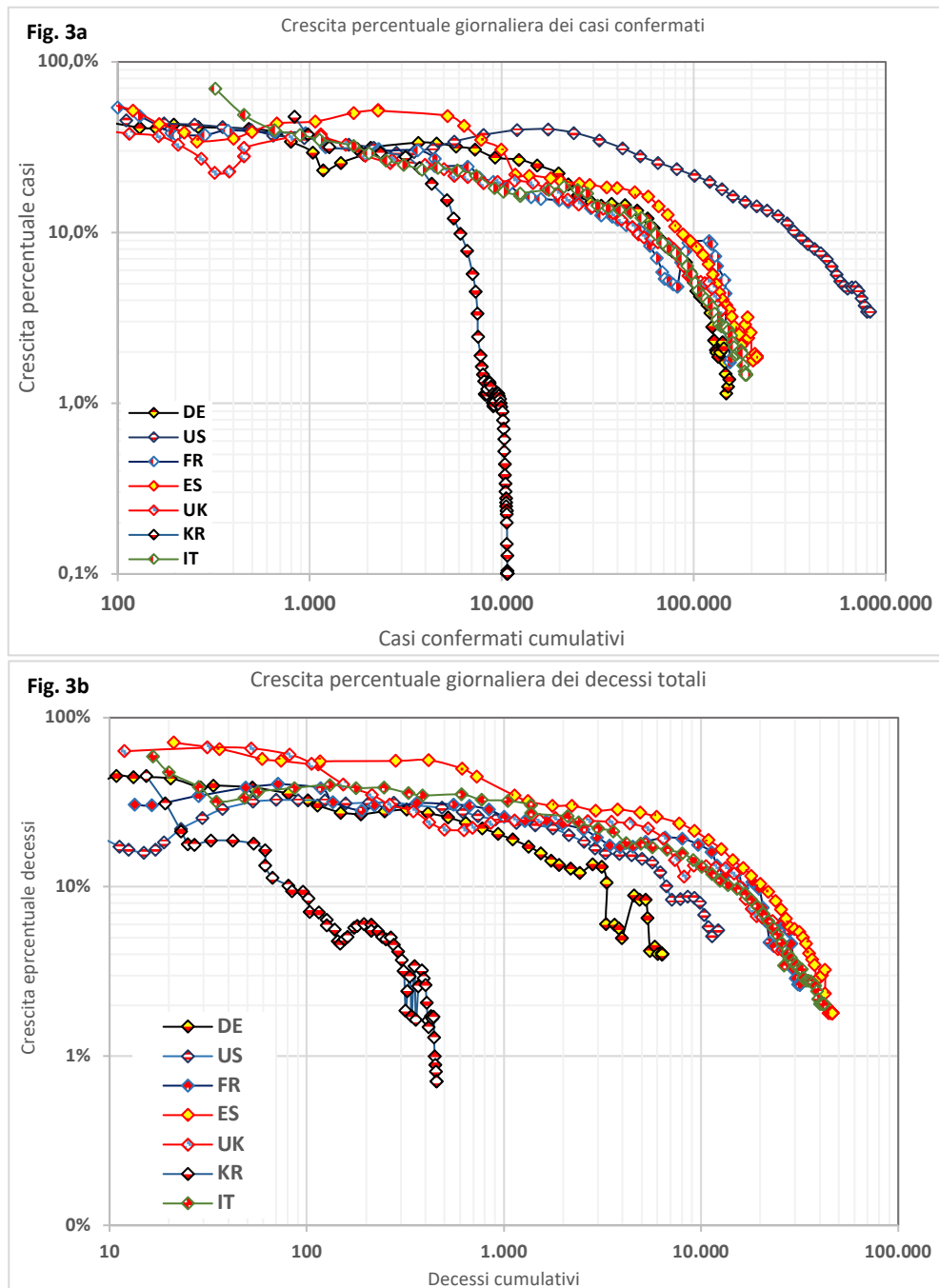


Fig.3: Crescita percentuale giornaliera dei a) casi confermati e b) decessi in funzione dei corrispondenti valori totali

Confronto tra i dati basati su grafici con la variabile tempo implicita

Un altro modo valido per confrontare l'evoluzione dell'epidemia nelle diverse nazioni, oltre alla traslazione dell'asse temporale, è quello di "rimuovere" la variabile tempo dai grafici. Questa metodologia è seguita nelle

Figure 3 a e b, dove la crescita percentuale giornaliera dei casi confermati e dei decessi è rappresentata in funzione del corrispondente numero cumulativo.

Questi grafici permettono di confrontare l'efficacia dell'azione delle diverse nazioni nel contenere la diffusione del COVID-19. Il relativo successo può essere misurato, infatti, dalla capacità di abbattere la percentuale di aumento giornaliero prima che il numero cumulativo di casi o morti diventi troppo grande. Per meglio inquadrare la situazione nelle nazioni occidentali, il grafico comprende anche il caso della Corea del Sud, spesso indicata come esempio per l'efficacia finora mostrata nel limitare la diffusione epidemica.

La Figura 3 mostra che, mentre tutte le nazioni europee presentano comportamenti simili tra loro, con una curva di variazione percentuale giornaliera scesa sotto il 10% per un numero di casi superiore di circa un ordine di grandezza a quello della Corea del Sud. Gli Stati Uniti presentano un valore di maggiore di un fattore 3-4 rispetto alle nazioni europee ed il numero di casi confermati per gli Stati Uniti sembra destinato a superare il milione.

La Figura 3b sottolinea anche che, in confronto ai paesi occidentali, il successo della Corea del Sud nel contenere finora il numero dei decessi è ancora più significativo rispetto a quanto osservato per i casi confermati. In merito, Italia, Spagna, Regno Unito e Francia mostrano comportamenti molto simili con una possibile convergenza a valori leggermente superiori ai 30.000 decessi. La Germania è scesa sotto la soglia del 10% di variazione giornaliera per un numero di vittime 4-5 volte più piccolo di quello degli altri paesi europei. Infine, gli Stati Uniti sono scesi sotto la soglia del 10% dell'aumento giornaliero per un numero complessivo di decessi circa 2-3 volte maggiore delle nazioni europee. Bisogna sottolineare nuovamente che qualsiasi commento predittivo, relativamente agli Stati Uniti, è reso maggiormente incerto dalla grande disomogeneità tra stato e stato.

Più in generale, qualsiasi evoluzione futura dipenderà fortemente dalle politiche adottate dalle diverse nazioni, questione ben al di là del potere predittivo della scienza. Politiche unitarie potrebbero spingere le curve a conservare le notevoli sovrapposizioni attuali, mentre decisioni differenti potrebbero portare le stesse a divergere. L'interruzione apparentemente prossima delle politiche di "lockdown" avrà effetti ad oggi difficilmente prevedibili. Una considerazione di ulteriore prudenza meritano gli Stati Uniti, i cui dati rappresentano la somma di un insieme (i vari Stati) fortemente disomogeneo sia in termini di stadio evolutivo che di politiche adottate, e potrebbero essere influenzate significativamente dalla crescita di focolai oggi ancora allo stato iniziale.

Conclusioni

Gli andamenti riferiti all'evoluzione dell'epidemia nelle diverse nazioni si differenziano per (a) le rilevanti diversità nelle politiche adottate, nell'organizzazione dei vari stati, nelle abitudini sociali e nelle condizioni

climatiche e b) per la diversa accuratezza con cui i casi di contagio confermati rappresentano i numeri reali dell'epidemia. Questo tipo di considerazione potrebbe scoraggiare qualsiasi tentativo di estrarre informazioni attraverso il confronto tra i vari andamenti nazionali. Tuttavia, quando le curve sono rappresentate in maniera appropriata, è possibile identificare similitudini significative. In questo articolo, abbiamo confrontato l'evoluzione del COVID-19 nelle sei nazioni (Stati Uniti, Spagna, Italia, Germania, Francia e Regno Unito) che al momento presentano il numero più elevato di casi di contagio confermati e di decessi. La difficoltà di analizzare nazioni con diversi livelli di evoluzione dell'epidemia è stata affrontata ricorrendo principalmente all'introduzione di opportune traslazioni temporali oppure ricorrendo a grafici in cui la variabile tempo è resa implicita. In questo modo, è stata messa in evidenza una relazione di *scaling* capace di portare tutte le curve a sovrapporsi con un'accuratezza sorprendente. Nel caso che tale relazione e i corrispondenti parametri finora estratti rimangano validi anche per il futuro (un'ipotesi da monitorare e verificare nelle prossime settimane), le nostre osservazioni potrebbero tradursi in una sorprendente capacità predittiva. Un altro modo efficace per confrontare l'evoluzione dell'epidemia nelle diverse nazioni è quello di rimuovere la variabile tempo dai grafici rendendola implicita. In particolare, se gli incrementi percentuali giornalieri vengono rappresentati in funzione del numero di decessi, risulta evidente che Italia, Spagna, Francia e Regno Unito seguono una tendenza significativamente simile che produrrà, probabilmente, valori finali confrontabili. Questa analisi fornisce dunque nuovi spunti riguardo l'evoluzione passata e, possibilmente, su quella futura della pandemia nelle diverse nazioni. La sezione internazionale del sito [COVID-19](#) dell'Istituto CNR-SPIN consente a qualsiasi utente di monitorare quotidianamente gli andamenti analizzati in questo articolo, valutarne altri e formulare nuove predizioni da testare nel tempo.

Referenze

1. CSSEGISandData/COVID-19. *GitHub* <https://github.com/CSSEGISandData/COVID-19>.
2. Miletto Granozio, F. L'evoluzione dell'epidemia in Europa. *Scienza in rete* <https://www.scienzainrete.it/articolo/levoluzione-dellepidemia-europa/fabio-miletto-granozio/2020-03-13> (2020).
3. Miletto Granozio, F. Comparative analysis of the diffusion of Covid-19 infection in different countries. *ArXiv200308661 Phys. Q-Bio* (2020).
4. Maier, B. F. & Brockmann, D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science* eabb4557 (2020) doi:10.1126/science.abb4557.
5. Sebastiani, G. Analisi dei dati epidemiologici del coronavirus in Italia: nota metodologica. *Scienza in rete* <https://www.scienzainrete.it/articolo/analisi-dei-dati-epidemiologici-del-coronavirus-italia-nota-metodologica/giovanni> (2020).