

Primo premio Distretto 2050

## Assessing COVID-19 healthcare system costs and vaccine effects: a dynamic system approach

**Simone Milanese**

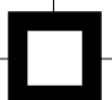
Corso di Laurea Magistrale in Matematica,

Università di Pavia

Relatore della tesi: Giuseppe De Nicolao

Il progetto di tesi si è concentrato sulla valutazione tramite sistemi dinamici dei costi sanitari e degli effetti delle vaccinazioni legati alla pandemia COVID-19. In particolare, ci si è concentrati sullo studio dei decessi, dell'occupazione delle unità di terapia intensiva e dell'occupazione degli ospedali durante la seconda ondata in Italia, indagando le loro relazioni causali con la serie temporale di nuovi soggetti positivi. Secondariamente, è stata studiata l'evoluzione della letalità in Italia e Regno Unito, al fine di avere una valutazione sia dell'aggressività della variante Alfa, sia del ruolo salvavita dei vaccini. Infine, è stato ricostruito il profilo temporale dei nuovi positivi durante la prima ondata: tale problema è stato motivato dal grave underreporting verificatosi a causa dell'impreparazione sanitaria e burocratica nel gestire un improvviso focolaio epidemico. Diversi risultati ottenuti all'interno del progetto hanno permesso di interpretare e prevedere alcuni aspetti del fenomeno pandemico: l'attività svolta ha fornito uno strumento che contribuisce a una visione consapevole dei costi umani e sanitari attesi in un determinato periodo e dell'effetto delle vaccinazioni.

L'innovatività dell'approccio risiede nella natura stessa del modello causale proposto e approfondito nel seguito: essendo i costi umani e sanitari determinati dal numero dei nuovi positivi, ne segue che dal punto di vista di un decisore politico è necessario ma anche sufficiente concentrare gli sforzi sul contenimento dei casi. La conoscenza di quanti nuovi infetti ci sono oggi fornisce, infatti, informazioni sul numero di giorni-letto d'ospedale di cui i



soggetti in questione avranno necessità, così come una stima istantanea del numero di morti che si verificheranno nel periodo immediatamente successivo.

Sono state utilizzate metodologie tipiche della 'Identificazione di sistemi' ed il principale strumento di calcolo è stata la piattaforma di calcolo Matlab, integrata dal System Identification Toolbox. L'analisi è stata condotta attraverso lo sviluppo di un modello matematico alle differenze finite tempo-invarianti di tipo SISO (single input, single output), chiamato d'ora in avanti modello di riferimento, o, più brevemente, modello. Sono stati identificati tre modelli, ciascuno dei quali ha come ingresso il numero giornaliero di nuovi soggetti positivi e ha come uscita, a seconda dei casi, il numero di ospedalizzazioni, numero di letti occupati in terapia intensiva o numero di decessi.

Il modello descrive ciascuna delle uscite come una convoluzione discreta tra l'ingresso (il numero dei nuovi positivi) e un'esponenziale negativa

$g(t) = \beta\alpha^{-t}, t \geq 0$  Segue in modo naturale che il modello è univocamente determinato dalla conoscenza dei parametri  $\alpha$  e  $\beta$ , i quali vengono calibrati a partire dall'analisi statistica delle serie storiche dei dati di ingresso e uscita.

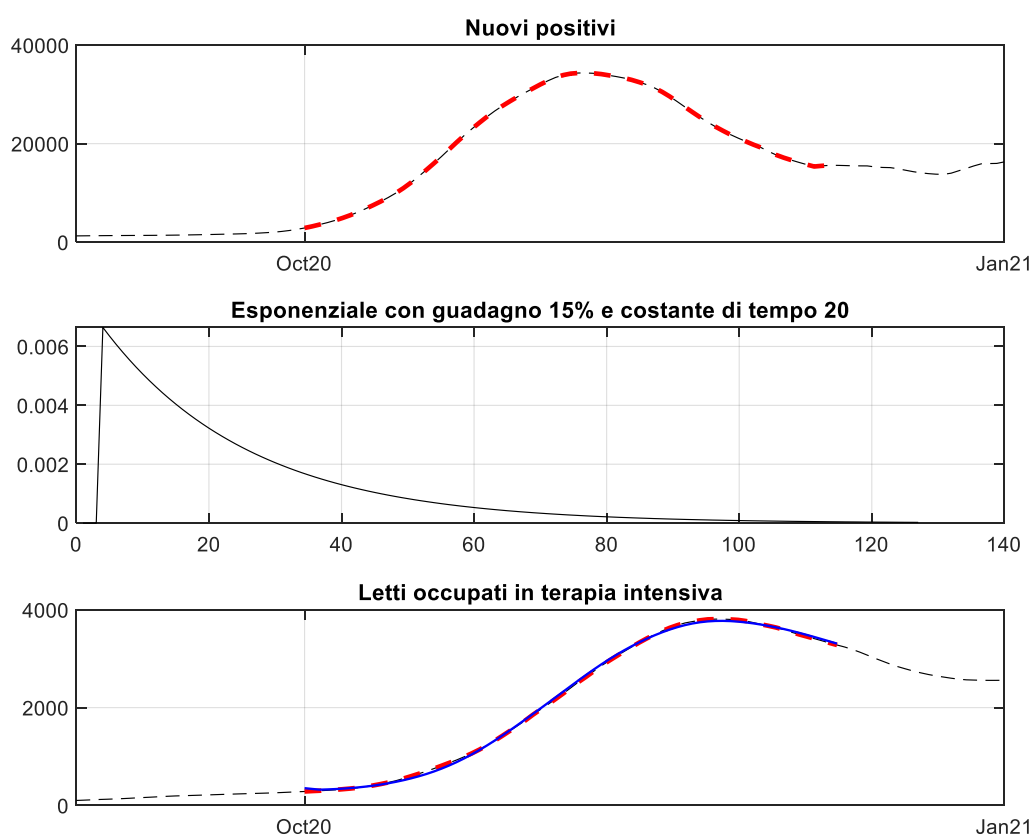
Due quantità fondamentali risultano associate al modello: il guadagno  $\mu$  e la costante di tempo  $\tau$  (Figura 1). La conoscenza di queste due quantità è equivalente alla conoscenza dei parametri  $\alpha$  e  $\beta$  ma risulta di interpretazione più immediata. Supponiamo di registrare oggi un certo numero  $x$  di nuovi positivi e di chiederci quanti di questi entreranno in terapia intensiva: ebbene, il modello permette di prevedere che, mediamente, tale numero sarà pari al prodotto  $\mu x$ . Il guadagno è pertanto una stima della frazione di  $x$  che, nel periodo successivo, entrerà in terapia intensiva. La costante di tempo permette invece di calcolare quanto tempo trascorre in media tra la registrazione del contagio e l'ingresso in terapia intensiva. In conclusione, un guadagno  $\mu = 0.15$  e una costante di tempo  $\tau = 20$ , implicano che, se oggi osserviamo un numero di nuovi positivi pari a 25'000, circa 3750 di essi finiranno in terapia intensiva entro un numero di giorni che, per ciascuno, è mediamente pari a 20.

Le forme esplicite di  $\mu$  e  $\tau$  sono ottenute facilmente a partire dai parametri  $\alpha$  e  $\beta$  il guadagno coincide con  $\beta/(1 - \alpha)$ ; la costante di tempo  $\tau$  è invece ottenuta come  $(-1)/(\ln(\alpha))$ .

Supponiamo quindi di essere un decisore politico in procinto di prendere una decisione sulle nuove misure restrittive. È indispensabile un lockdown? Quale sarà l'impatto sull'economia, sull'istruzione e sul benessere psicologico dei cittadini? Se si evitassero misure drastiche, fino a che punto e in quale misura il sistema sanitario riuscirebbe a reggere? O, ancora, come si può ottenere una valutazione quantitativa affidabile degli effetti dei vaccini sui decessi e i costi sanitari? La complessità di queste domande è significativa e necessita di uno strumento oggettivo che fornisca al decisore un'adeguata e istantanea analisi dei rischi e dei benefici.

Una prima risposta potrebbe essere quella di farsi guidare dai modelli epidemiologici, come ad esempio il modello SIR o le sue varianti più complesse. Tale modello descrive l'evoluzione della pandemia in termini di soggetti

suscettibili, infetti e rimossi, e ha come punto di partenza la conoscenza dei parametri pandemici, come ad esempio il reproduction number, meglio noto come  $R_t$ . Nonostante il modello SIR possa essere migliorato (introducendo nuove variabili oppure ampliandolo ad un modello stocastico) non si può prescindere dalle stime dei parametri pandemici. Questo fatto è alla base dell'incertezza fisiologica che accompagna le previsioni di questi modelli: per predire il numero dei casi attivi bisogna essere in grado di predire il comportamento dell' $R_t$ , il quale dipende da fattori di diversa natura (non solo biologici, ma anche comportamentali, ad esempio) che non possono essere facilmente modellizzati.



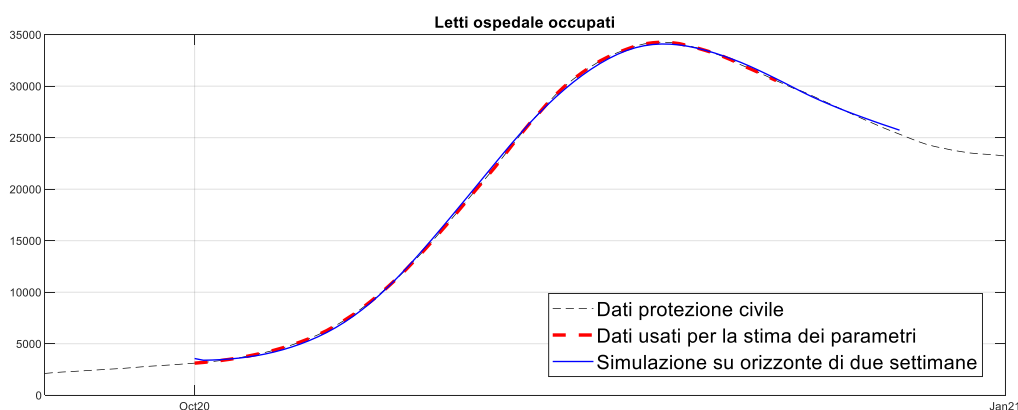
**Fig. 1**

**Stima di guadagno e costante di tempo del modello associato alla terapia intensiva.**

*Nel pannello in alto è visualizzato il profilo dei nuovi positivi dal primo di settembre 2020 al primo di gennaio 2021; viene evidenziato in rosso il periodo considerato per la stima dei parametri. Nel secondo pannello è rappresentata la funzione esponenziale, la cui convoluzione discreta con i nuovi positivi restituisce la curva blu, disegnata nel terzo pannello. Come si può osservare, la curva riprodotta dal modello è aderente alla curva dei dati reali nel periodo di interesse, tratteggiata in rosso.*

Una conferma dei limiti dei modelli epidemiologici in campo decisionale si riscontra nel fatto che tali modelli sono usati dai tecnici per produrre scenari differenti sulla base di diverse configurazioni dei parametri pandemici. In altri termini, data la difficoltà della stima dell' $R_t$ , si propongono diversi valori plausibili del parametro  $e$ , e per ciascuno di essi, si crea una casistica ad hoc. Un decisore che voglia usare i modelli epidemiologici dovrà quindi tenere in conto diversi scenari, preparare una strategia per ciascuno di essi e saper comunicare in modo efficace la ragione delle sue decisioni ad un pubblico generalista e, in una qualche misura, inesperto.

Il modello di riferimento studiato nella tesi, invece, è uno strumento numerico complementare ai modelli epidemiologici, il quale, pur limitatamente alla catena causale contagi-effetti sanitari, garantisce però un livello significativo di oggettività e di affidabilità concentrato in un unico scenario (Figura 2)



**Fig. 2**

***Affidabilità della previsione su un orizzonte di due settimane per i letti occupati in ospedale.***

*La curva grigia tratteggiata in figura corrisponde ai dati della protezione civile sulle ospedalizzazioni da metà settembre 2020 al primo di gennaio 2021. In particolare, sono stati evidenziati in rosso i dati tra il primo di ottobre 2020 e il sei dicembre 2020, utilizzati per le stime dei parametri  $\alpha$  e  $\beta$ . La curva blu, invece, corrisponde alla previsione ottenuta a partire dal modello. La prima parte (tra primo di ottobre e sei dicembre) consiste nel risultato di una simulazione e mostra un'aderenza significativa ai dati rossi su cui è stata stimata. La seconda parte (dal sei dicembre al venti dicembre) è invece di pura previsione: un tecnico che, arrivato al sei dicembre, si interroga sulla futura evoluzione delle ospedalizzazioni può fare riferimento al modello proposto. I giorni successivi, come mostrato dal grafico, ne confermeranno l'affidabilità.*

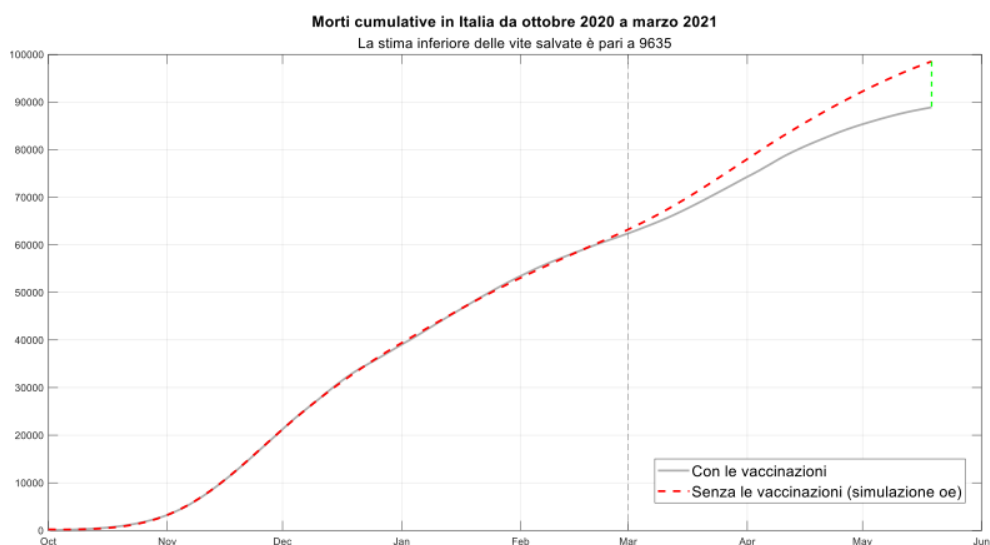
L'efficacia dell'analisi proposta risiede nel fatto che il punto di partenza non è la stima dei parametri pandemici, bensì coincide con i dati registrati dalla protezione civile su nuovi positivi, ospedalizzazioni, occupazione delle terapie intensive e morti. Il buon funzionamento delle previsioni è verificabile anche visivamente in Figura 2: i modelli proposti sono in grado di prevedere con accuratezza l'evoluzione nelle due settimane future dei costi umani e sanitari. Un decisore

politico può quindi vagliare eventuali restrizioni conscio, ad esempio, della percentuale di saturazione delle terapie intensive, le quali costituiscono una componente rilevante del peso che grava sul sistema sanitario.

Se accettassimo l'ipotesi che il virus si ripresenterà stagionalmente per diversi anni, l'utilità del modello proposto è evidente e risulterà fondamentale, a patto di aggiornare i parametri del modello in funzione delle varianti del virus e del grado di immunità acquisita dalla popolazione.

Un secondo risultato fondamentale ottenuto in tesi è la possibilità di stimare le vite salvate dai vaccini nei mesi successivi alla loro introduzione. In particolare, quella che viene ottenuta è una stima per difetto: le vaccinazioni hanno effetti su più livelli e la nostra stima si occupa di descrivere il contributo diretto. Per effetto diretto intendiamo l'abbassamento della letalità, nel senso che i soggetti vaccinati non si infettano o, se infettati, la probabilità di morire diminuisce. Un secondo effetto, più difficile da quantificare con precisione, è indiretto: anche se i vaccini non danno una copertura completa dal virus, riducono la probabilità di contagio e quindi in una popolazione parzialmente vaccinata, la circolazione del virus è ostacolata e di conseguenza si assiste ad un ulteriore calo delle morti.

La nostra analisi ha portato a determinare almeno 9600 vite salvate dall'introduzione dei vaccini nel periodo in esame che andava da marzo a maggio 2021 (Figura 3).



**Fig. 3**

**Stima per difetto delle vite salvate dalle vaccinazioni fino a metà maggio 2021.**  
Analogamente ai casi precedenti, viene stimato un modello matematico della dinamica delle morti cumulative in Italia, basato sui dati del periodo tra ottobre 2020 e marzo 2021. Supponendo che gli effetti delle vaccinazioni siano osservabili a partire da marzo 2021, la linea rossa tratteggiata mostra una simulazione dei decessi da marzo in poi ipotizzando la dinamica 'senza vaccinazioni'. Tale andamento è poi confrontato con la curva grigia dei dati reali (dinamica con vaccinazioni). Il segmento tratteggiato verde, infine, mette in evidenza la stima per difetto delle vite salvate.

In conclusione, osserviamo che, come intuibile, guadagni e costanti di tempo dei modelli dipendono dalla contagiosità e dalla virulenza della variante in circolo: ci si aspetta che, ad esempio, la variante alpha e la variante omicron presentino parametri strutturalmente diversi. Il modello proposto rimane valido a condizione che si aggiornino i parametri tramite una nuova stima. Per di più, questa peculiarità apre la strada a interessanti direzioni di ricerca: una possibile miglioria consisterebbe nell'abbandonare il modello parametrico esponenziale in favore di una risposta impulsiva  $g$  tempo-variante (approccio modellizzabile per esempio tramite tecniche di identificazione come i metodi kernel). Così facendo, i parametri del modello risulterebbero anch'essi funzioni dipendenti dal tempo e non avrebbero il vincolo di risultare costanti sui brevi periodi.

0

1

0

1

0