

# O modelovaní šírenia CoVid-19

*“Essentially, all models are wrong, but some are useful.”*  
G. Box, Empirical Model-Building and Response Surfaces,  
Wiley, 1987, page 424.

V súčasnej situácii v súvislosti s šírením CoVid-19 všetkých zaujíma otázka ohľadne možnosti predikcie budúceho vývoja. Preto je prirodzené pokúsiť sa o matematické modelovanie tohto procesu s tým vedomím, že aj keď žiadny matematický model nemôže vystihovať realitu úplne, môže byť užitočný pre celkový odhad situácie v krátkodobom aj dlhodobom časovom horizonte — modelovanie počasia je príklad, ktorý pozná asi každý.

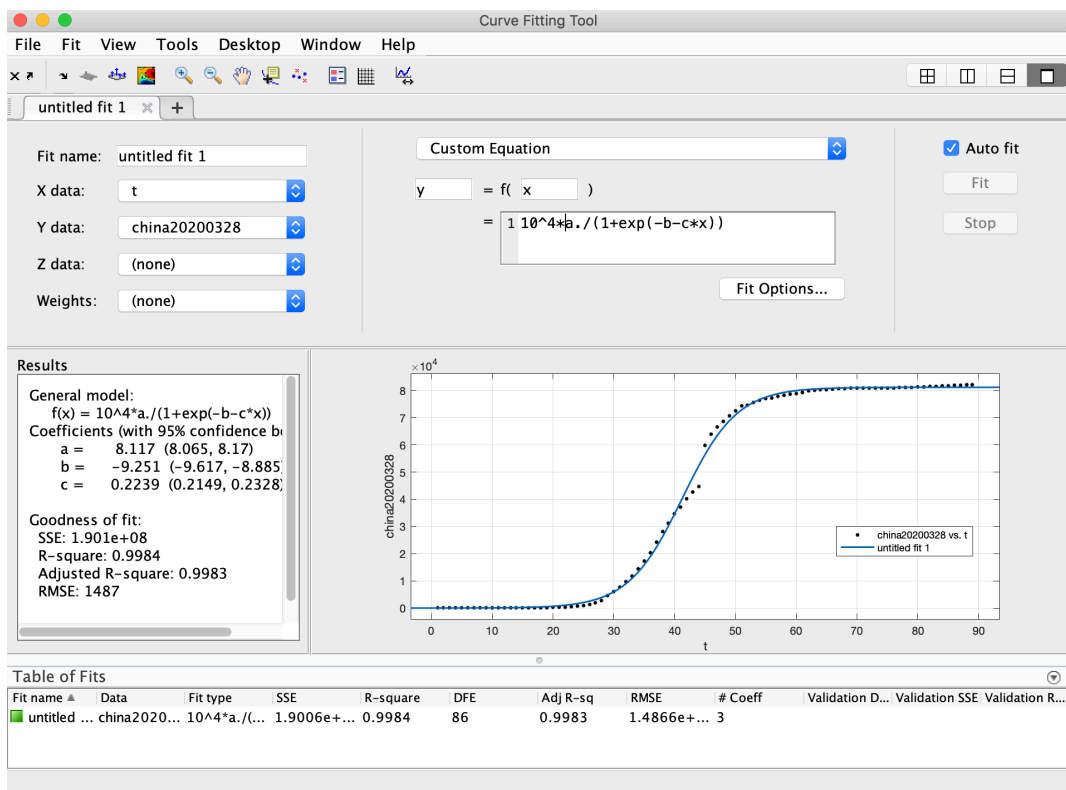
## Jednoduchá aproximácia celkového počtu nakazených po skončení procesu

Keď je k dispozícii úplný súbor dát a proces šírenia infekcie sa zastavil (napríklad Čína), výsledná krivka má tvár pripomínajúci písmeno S a dá sa dobre aproximovať tzv. “logistickou krivkou”:

$$y(t) = \frac{a}{1 + e^{-b-ct}}$$

kde  $t$  je čas,  $y(t)$  je celkový počet infikovaných v čase  $t$ , a koeficienty  $a$ ,  $b$ ,  $c$  sa určujú tak, aby krivka  $y(t)$  bola čo najbližšie k bodom zobrazujúcim aktuálne dáta.

Ak na výpočty použijeme softvér MATLAB a hotový nástroj “Curve Fitting Tool”, ľahko môžeme získať nasledujúcu aproximáciu (na obrázku je uvedený prípad Číny):



Ale toto je situácia typu “po bitke je každý generál”, lebo význam parametra  $a$  je celkový finálny počet infikovaných, ktorý v prípade Číny už v podstate poznáme, ale v prípade krajín, kde sa infekcia ešte šíri, potrebujeme predpovedať práve celkový počet a k tomu aj dobu trvania procesu. Musí byť použitý iný model — dynamický, v podobe diferenciálnych rovníc.

## Model typu SIR

Skratka “SIR”, alebo “S–I–R”, je odvodená od začiatkových písmen slov “susceptible, infected, recovered” (nakazitelní, infikovaní, uzdravení). Daný model predstavuje sústavu troch rovníc o troch neznámych funkciách  $S(t)$ ,  $I(t)$ ,  $R(t)$ :

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta}{N}IS \\ \frac{dI}{dt} &= \frac{\beta}{N}IS - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}$$

kde sa predpokladá, že  $S(t)+I(t)+R(t)=N$ , pričom  $N$  je celková veľkosť populácie. Parameter  $\beta$  je intenzita kontaktov medzi nakazitelnými a infikovanými subjektmi (niečo ako koeficient rýchlosti šírenia infekcie); doba infekčnosti jedinca je  $1/\gamma$ . K týmto trom diferenciálnym rovniciam treba pridať začiatkové podmienky:

$$S(0) = S_0, \quad I(0) = I_0, \quad R(0) = R_0.$$

Tento model bol aktuálne implementovaný v prostredí MATLAB profesorom mechaniky Milanom Batistom (<http://www.fpp.edu/~milanb/>) z Univerzity v Ljubljane a je uverejnený na MATLAB Central File Exchange (<https://www.mathworks.com/matlabcentral/fileexchange/74658>) pod názvom “fitVirusCOVID19”. Tento model je teraz aktualizovaný a vylepšovaný prakticky denne.

Aktuálne dáta pre výpočet sú preberané zo stránky “Coronavirus Source Data” (<https://ourworldindata.org/coronavirus-source-data>), ktorá je aktualizovaná denne na základe údajov ECDC (European Centre for Disease Prevention and Control). Preto Prof. Batista dňa 30.3.2020 pridal aj funkciu na import dát; na použitie tejto funkcie si treba stiahnuť “Full dataset”, uložiť ho vo formáte XLSX a použiť funkciu “importData.m”, ktorá vygeneruje dáta podľa jednotlivých krajín. (Poznámka: je možné očakávať, že v ďalších verziách balíka “fitVirusCOVID19” bude tento pomocný krok automatizovaný.)

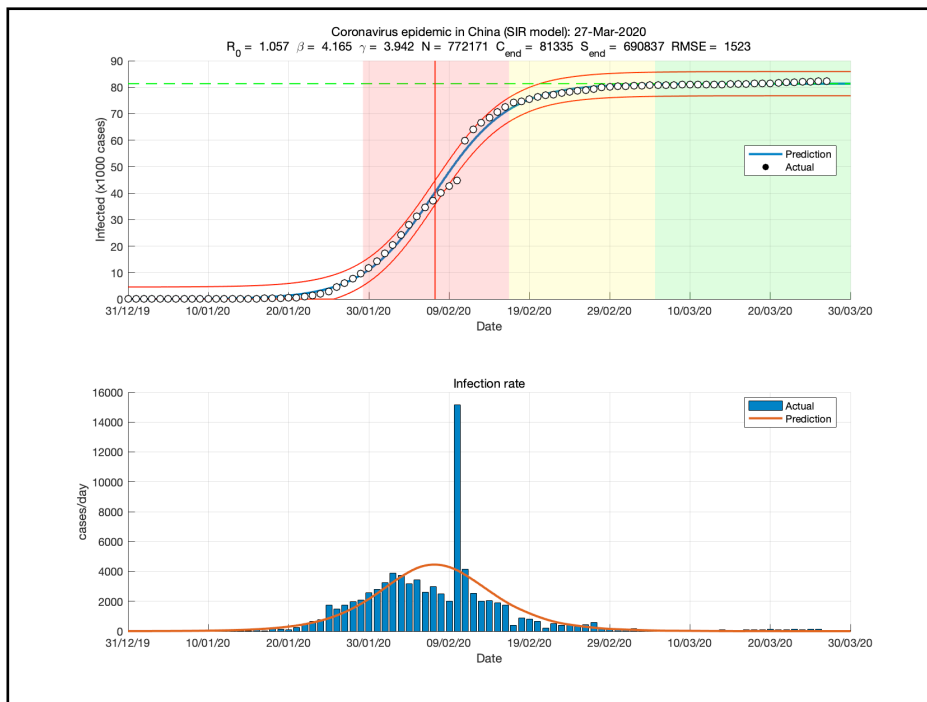
Parametre (koeficienty) modelu pre každú krajinu sú samozrejme iné a preto sú určované iteračnou optimalizačnou metódou na základe aktuálnych dát pre danú krajinu; s pribúdajúcimi dátami sa parametre modelu pre každú krajinu upresňujú. Daný model umožňuje odhadnúť celkový počet infikovaných a celkovú dobu trvania šírenia infekcie.

Samozrejme, tento klasický model výborne implementovaný profesorom Milanom Batistom teoreticky umožňuje rôzne zdokonalenia. Ale v súčasnej situácii celkom dobre vystihuje to, čo už vieme, a umožňuje odhad situácie do budúcnosti, pričom tento odhad je aktualizovaný a upresňovaný na základe pribúdajúcich čísel — podobne predpovediam počasia, ktoré boli spomenuté v úvode.

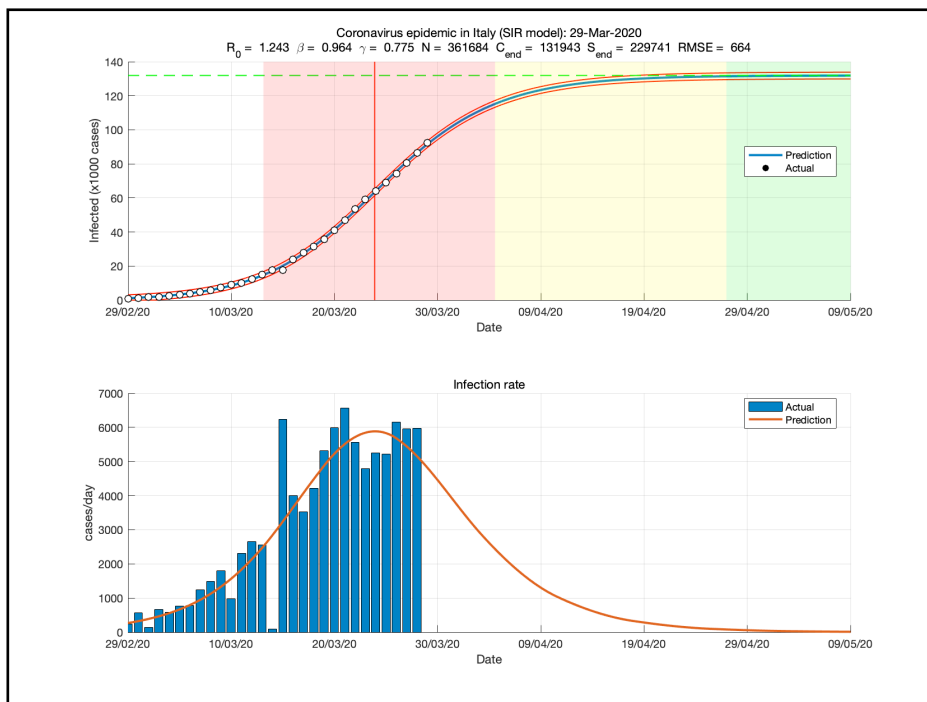
## Príklady k 30.3.2020

Uvedieme príklady výstupov generovaných balíkom “fitVirusCOVID19” od profesora Milana Batistu v dnešnej verzii (30.3.2020) a s dnešnými dátami.

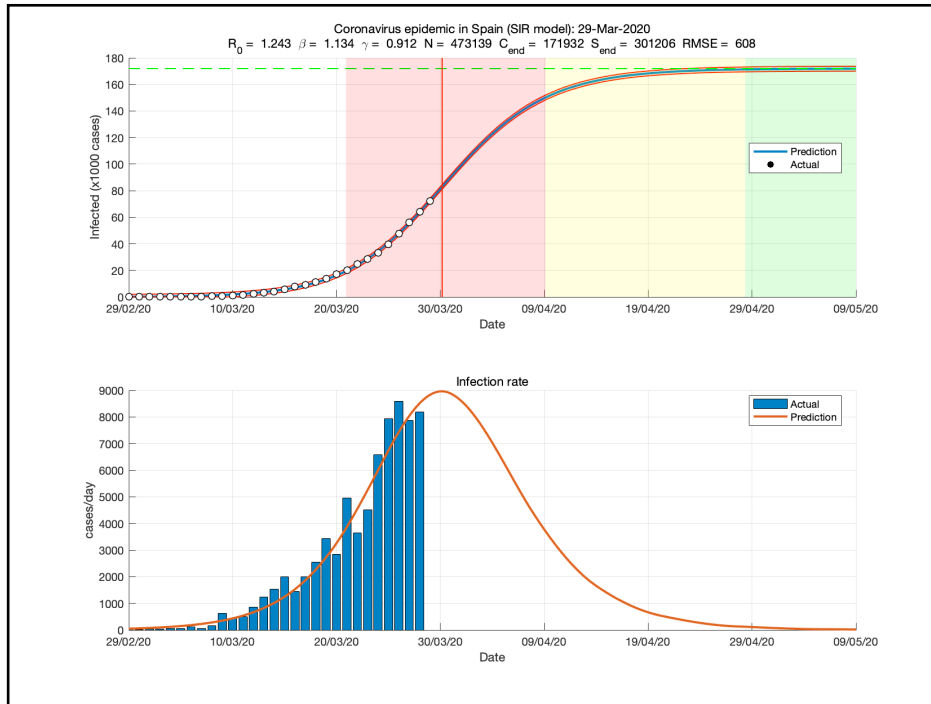
Ako vidieť na príklade Číny, výsledok SIR modelu výborne vystihuje tú “logistickú funkciu”, ktorá bola uvedená na obrázku na začiatku — inými slovami, SIR model vystihuje aj finálny stav:



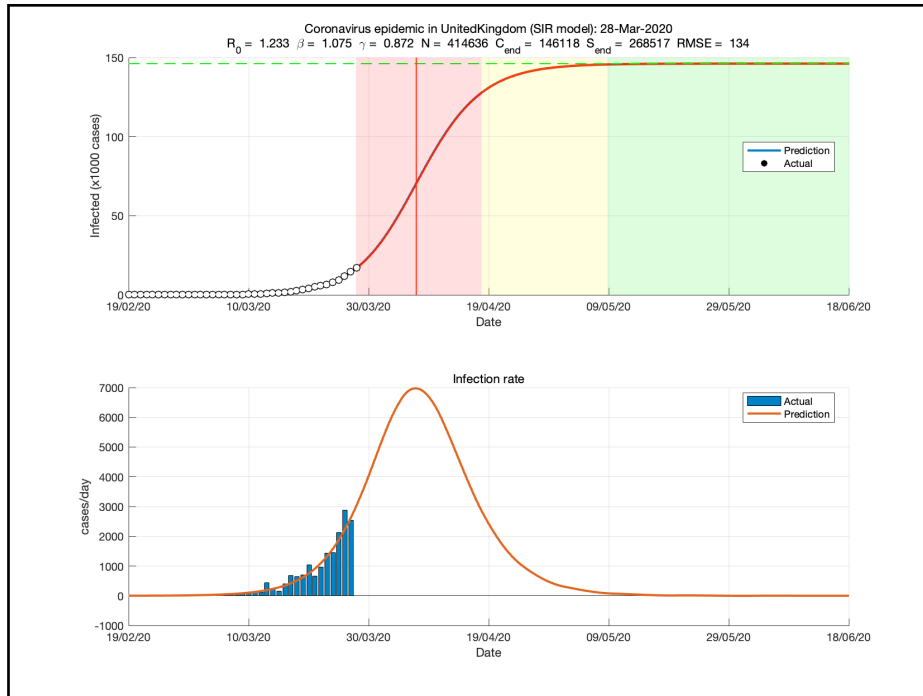
Taliansko a Španielsko sú približne v polovici lineárnej fázy. Podľa SIR modelu, v danej chvíli je možné očakávať celkový počet infikovaných v Taliansku okolo 135 tisíc:



V prípade Španielska je aktuálne odhadovaný finálny počet infikovaných okolo 175 tisíc:

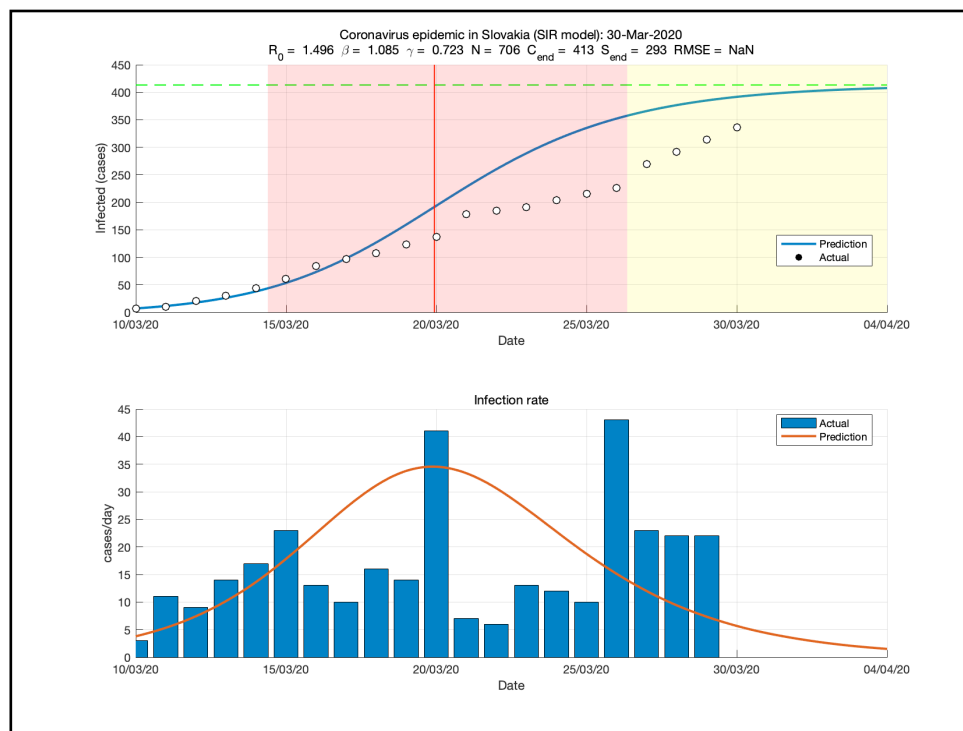


Veľká Británia je na začiatku lineárnej fázy, s aktuálnym očakávaným finálnym počtom infikovaných okolo 150 tisíc obyvateľov:



Na týchto príkladoch vidieť niektoré typické fázy. Prvá fáza je v podstate exponenciálny rast počtu infikovaných. Snaha zabrániť tomuto prudkému nárastu (reštrikcie, obmedzenia kontaktov, ...) má viesť k prechodu do fázy lineárneho nárastu (ružová oblasť na týchto grafoch), po ktorej treba dosiahnuť spomaľovanie (žltá oblasť) a následne zastavenie procesu (zelená oblasť).

Predpoveď v prípade Slovenska nie je zatiaľ celkom jasná, zrejme kvôli nepravidelným prírastkom počtu infikovaných a kvôli malému počtu infikovaných oproti príkladom uvedeným vyššie (nech to tak bude aj naďalej!):



V prípade Belgicka však program prof. Batistu zatiaľ nedokázal vypočítať potrebné konštanty, preto nedokázal ani vytvoriť predikciu, takže zostáva čakať.

Pripravil: prof. RNDr. Igor Podlubný, DrSc.  
30.3.2020

### Odkazy:

1. "Logistic function". From Wikipedia, [https://en.wikipedia.org/wiki/Logistic\\_function](https://en.wikipedia.org/wiki/Logistic_function) (Retrieved March 30, 2020).
2. Smith, D., and Moore, L.: "The SIR Model for Spread of Disease". *Journal of Online Mathematics and its Applications*, December 2001, <https://www.maa.org/press/periodicals/loci/joma/the-sir-model-for-spread-of-disease-introduction> (Retrieved March 30, 2020).
3. Batista, M.: fitVirusCOVID19: Estimation of coronavirus COVID-19 epidemic evaluation by the SIR model. *MATLAB Central File Exchange*, <https://www.mathworks.com/matlabcentral/fileexchange/74658-fitviruscovid19> (Retrieved March 30, 2020).
4. Coronavirus Source Data, <https://ourworldindata.org/coronavirus-source-data> (Accessed March 30, 2020).