



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Modelování a simulace

Epidemiologické modely - modely na makroúrovni

2020/2021

Matej Otčenáš, xotcen01
Mário Gažo, xgazom00

November 15, 2020

Obsah

1	Úvod	2
2	Riešenie	2
2.1	Implementácia	2
2.2	Voľba modelu	2
2.3	SIR model	2
2.4	SEIRD model	3
3	Odhad parametrov	4
3.1	Parameter α	4
3.2	Parameter β	4
3.3	Parameter σ	5
3.4	Parameter ω	5
4	Experimenty	5
4.1	Experiment 1	6
4.2	Experiment 2	9
4.3	Experiment 3	13
4.4	Experiment 4	17
5	Záver	21
	Zdroje	22

1 Úvod

Našou úlohou bolo vytvoriť model epidémie na makro úrovni, na určitom väčšom území, preskúmať jej dopady na chod kraja. Ďalej máme preskúmať možné scenáre priebehu samotnej epidémie za určitých okolností, teda napr. počet úmrtí, maximálny počet nakazených jedincov v jeden moment, odhad konca epidémie. Taktiež máme zistiť aký vplyv by mal na priebeh angažovanosť moci štátnych zástupcov, rôzne nariadenia, čiastočné, či úplné uzavretie.

2 Riešenie

2.1 Implementácia

Pri tvorbe modelov sme použili jazyk C++ a v ňom vytvorenú triedu model, ktorá sa stará o spracovanie argumentov programu, spúšťanie experimentov a samotnej simulácie, vypisovanie hodnôt do súborov vo formáte CSV v priečinku *statistics* a na štandardný výstup (stdout). Pri vizualizácii dát z CSV súborov sme použili jazyk Python, na spracovanie súborov knižnicu *pandas* a na samotnú vizualizáciu *matplotlib*.

2.2 Voľba modelu

Pre modelovanie tejto situácie nie je možné modelovať Petriho sieťou, keďže sa nejedná o systém hromadnej obsluhy (SHO), ale o spojitý systém. Na výber je viacero existujúcich možností. Po preskúmaní každej a konzultácii s profesorkou Monikou Heiner z Brandenburg University of Technology Computer Science Institute sme sa rozhodli pre SIR model a jeho rozšírenie SEIRD model.

2.3 SIR model

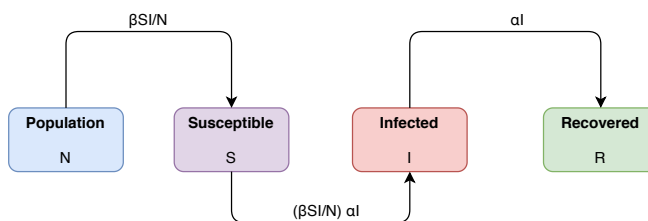
Jedná sa o model šírenia epidémie, ktorý je založený na trojici diferenciálnych rovníc. Pri každom kroku simulácie sú vyrátané početnosti jednotlivých skupín, pričom v každom momente musí platiť, že $S + I + R = N$, a že $dS + dI + dR = 0$. Model zjednodušuje realitu v tom, že tí ktorí sa nakazili a vyliečili, sa nemôžu opätovne nakaziť. Ak ním chceme modelovať epidémiu, musíme zadať početnosť populácie, infekčných, tempo ochorenia a vyliečenia. Model SIR je možné popísať nasledovnými diferenciálnymi rovnicami[1].

$$\frac{dS}{dt} = -\frac{\beta IS}{N} \quad (1)$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I \quad (2)$$

$$\frac{dR}{dt} = \gamma I \quad (3)$$

Premenná S predstavuje množinu ľudí, ktorých je možné nakaziť (angl. *susceptible*), I je množina infekčných ľudí, R množina ľudí, čo ochorenie prekonali a získali imunitu, resp. nákaze podľahli (angl. *removed*). Celkovú populáciu reprezentuje množina N . Rovnice obsahujú taktiež koeficienty, ktoré ovplyvňujú samotný model. Koeficient β udáva rýchlosť šírenia nákazy (angl. *transmission rate*), ten sa odvíja od súčinu reprodukčného čísla¹ a koeficientu α . Koeficient α je naopak hodnota rýchlosti zotavenia z ochorenia (angl. *recovery rate*).



Obrázok 1: Návrh systému SIR

2.4 SEIRD model

Na rozdiel od SIR modelu, SEIRD model počíta aj s určitou mierou mortality a taktiež s tým, že nie každý, kto bol v kontakte s nakazenou osobou sa nakazí. Jedná sa o viac realistický model, no taktiež je komplexnejší, pretože je potreba zvoliť viacero parametrov, teda okrem tých pre SIR aj tempo nakazenia a pravdepodobnosť úmrtia. Model SEIRD je taktiež možné, rovnako ako model SIR, popísať diferenciálnymi rovnicami[2].

$$\frac{dS}{dt} = -\frac{\beta IS}{N} \quad (4)$$

$$\frac{dE}{dt} = \frac{\beta IS}{N} - \sigma E \quad (5)$$

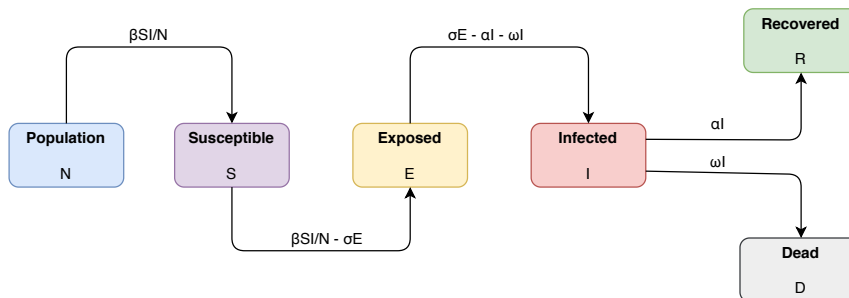
$$\frac{dI}{dt} = \sigma E - \alpha I - \omega I \quad (6)$$

$$\frac{dR}{dt} = \alpha I \quad (7)$$

$$\frac{dD}{dt} = \omega I \quad (8)$$

¹<https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1a-epidemiology/epidemic-theory>

Rovnako ako model SIR, aj model SEIRD pozostáva z identických premenných a koeficientov, je však o niečo rozšírený. Premenná E je množina ľudí, ktorí boli vystavení infekčnej osobe a oni sami sú infikovaní, nie však infekční (angl. *exposed*). Premenná D je množina ľudí, ktorí nákaze podľahli. Koeficient σ je úroveň skrytých prenášačov, ktorí sa stanú infekčnými a ω^2 je úroveň úmrtnia na dané ochorenie.



Obrázok 2: Návrh systému SEIRD

3 Odhad parametrov

Dôležitými parametrami v oboch modeloch sú pravé tie, ktoré vystupujú ako koeficienty v uvedených diferenciálnych rovniciach. Na ich odhad by bolo potrebné zozbierať veľké množstvo dát a tie ďalej spracovávať určitými postupmi. Je však možné tieto hodnoty získať z rôznych vedeckých publikácií, ktoré sa týmito výpočtami zaoberali detailnejšie[3][4][5][6].

3.1 Parameter α

Koeficient α je hodnota vyjadrujúca priemernú dobu liečby z ochorenia *SARS-CoV-2* a je možné ho vyjadriť ako $\frac{1}{\gamma}$, kde priemerný čas na liečbu je odhadovaný na $\gamma = 18$ dní. V takom prípade je $a = \frac{1}{18} = 0.556$.

3.2 Parameter β

Udáva rýchlosť šírenia nákazy, a teda ako rýchlo sa bude prenášať kvapôčkový patogén medzi jednotlivými osobami. Tento koeficient je možné získať vyjadrením z rovnice $R_0 = \frac{\beta}{\alpha} \implies \beta = R_0 * \alpha$. V našom prípade bude hodnota $\beta = 0.556 * 5.6015 = 3.114434$. Je to hodnota, ktorá je odhadnutá pre čínsku provinciu Hubei, ale nezostane rovnaká po celú dobu simulácie, pretože je to

²IFR - Infection Fatality Rate

hodnota dynamická a veľmi často sa môže meniť napríklad na základe epidemiologických opatrení či zvýšenej mobility ľudí v danej oblasti.

3.3 Parameter σ

Hodnota parametru σ je hodnotou, ktorá odhaduje priemerný čas inkubácie a je možné ju vyjadriť ako $\frac{1}{\gamma}$, kde $\gamma = 5.2$ dní. Tak získame hodnotu $\sigma = \frac{1}{5.2} = 0.1923$.

3.4 Parameter ω

Rovnako ako parameter β aj ω je parametrom dynamickým a je ovplyvňovaný vedľajšími faktormi ako je napríklad situácia, že nastane preťaženie zdravotného systému či naopak vznikne tzv. kolektívna imunita, čo ale v prípade ochorenia *SARS-CoV-2* nebolo oficiálne potvrdené. V našom modeli ale budeme uvažovať tento parameter ako konštantu, a teda jeho hodnota bude nemenná. Hodnota bola zvolená podľa dostupných vedeckých publikácií a je nastavená na hodnotu $\omega = 0.0034$.

4 Experimenty

Vytvorený simulačný model popisuje pomocou experimentov priebeh vzniku ochorenia *SARS-CoV-2*, ktoré bolo ako prvé zaznamenané v čínskej provincii Hubei. Ako prvý prípad nákazy bol evidovaný dňa 31.12.2019 čo predstavuje aj počiatok simulácie. Simulácie končí dňom 7.12.2020[7].

Odhadovaná populácia v tejto provincii je približne 58 miliónov ľudí pričom každý experiment ráta s prvotným počtom infekčných zvolenom na 27 čo je počet ľudí, ktorí boli dňa 31.12.2019 identifikovaní ako chorí na neznámy vírus. Experimenty predovšetkým demonštrujú efekt radikálnych opatrení voči šíreniu nákazy, aké boli v tejto provincii, ale aj prevažne v celej Číne nasadené. Boli to opatrenia vo forme úplného lockdownu všetkých miest na čele s mestom Wuhan, ktorý bol samotným epicentrom nákazy v tomto kraji, čo znamenalo uzatvorenie všetkých letísk či odstavenie hromadného transportu. Tieto opatrenia boli oficiálne nariadené v provincii Hubei dňa 24.1.2020. Faktory, ktoré vo veľkom ovplyvnili počiatočný exponenciálny nárast ochorení boli predovšetkým vysoká mobilita ľudí, keďže sa v tomto období práve začínali oslavy Čínskeho nového roku, čo je považované v Číne ako za jednu z najvýznamnejších udalostí v roku. V tomto období nastal súčasne aj nárast nami sledovaného reprodukčného čísla R_0 ³, ktoré dňa 23.1.2020 stúplo z hodnoty 5.6015 na 6.6037 čo avšak pretrvávalo iba do dňa 27.1.2020 a následne kleslo pomerne rapídne na hodnotu 3.7732. Tento výrazný pokles bol docielený vďaka rýchlemu konaniu a uvedeniu oblasti do úplnej karantény. Hodnota $R_0 = 3.7732$ v provincii bola evidovaná do dňa 11.2.2020, kedy sa ochorenie ešte stále šírilo pomerne rýchlo no s utíchajúcou

³Všetky uvedené hodnoty za určité obdobia sú hodnotami priemernými a nie sú dynamicky upravované počas uvedeného obdobia, čo môže spôsobovať isté numerické odchýlky

tendenciou. Dňa 12.2.2020 je evidované reprodukčné číslo na hodnote $R_0 < 1$, kde $R_0 = 0.2020$, čo znamená že samotné ochorenie sa ďalej už nebude šíriť. Táto hodnota pretrváva až do konca simulácie. R_0 ovplyvňuje samotnú rýchlosť šírenia nákazy[8][9][10][11][12].

4.1 Experiment 1

Tento experiment popisuje vývoj ochorenia pomocou modelu SIR a poukazuje na vývoj v provincii Hubei, ktorý by nastal v prípade, že by nebol nariadený žiadny protiepidemiologický krok. V tomto prípade sa nedá oprieť o nejaký dostupný model, pretože v danom kraji boli nariadené radikálne opatrenie v pomerne krátkom čase.

Vstupné hodnoty:

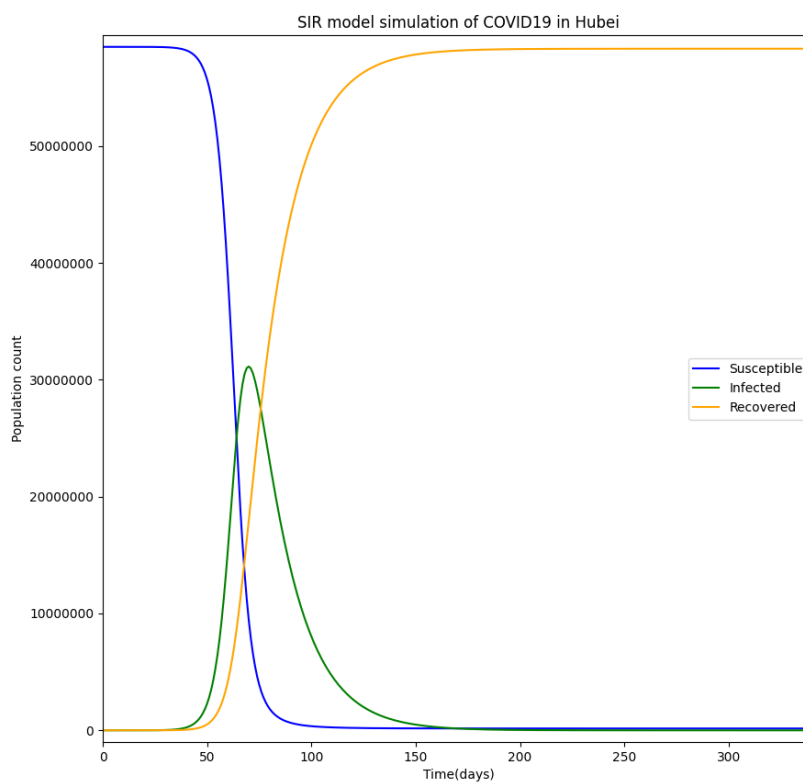
- Populácia = 58,500,000
- Infikovaní = 27
- Koeficient $\beta = 0.31144$
- Koeficient $\alpha = 0.05560$

Výstupné hodnoty:

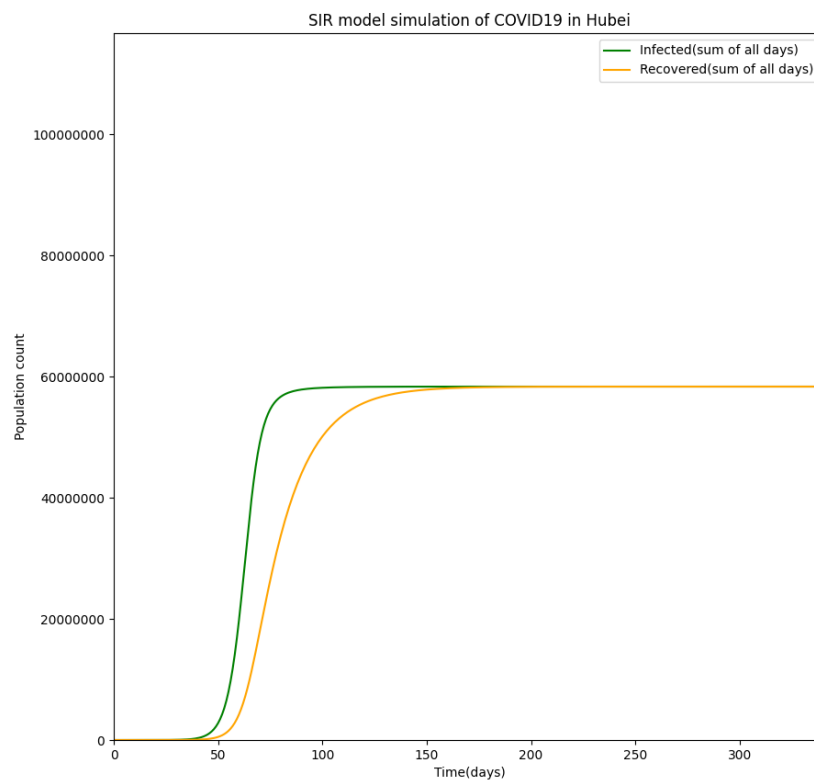
- Celkový počet infikovaných = 58,450,634 (99.91561%)
- Celkový počet zotavených⁴ = 58,450,628 (99.91560%)
- Najväčší denný prírastok = 3,129,556 (deň č. 57)
- Najväčší počet infikovaných v jeden moment = 3,422,7490 (deň č. 64)

Výstup z experimentu jasne poukazuje na to, že v prípade nezavedenia opatrení v akejkoľvek podobe reprodukčné číslo R_0 nezačne nikdy klesať, z čoho vyplýva že počet infikovaných bude neustále exponenciálne rásť. Výstup tohto modelu je na obrázku 3. Simulácia zobrazuje rovnako tak aj priebeh ochorenia s výstupnými hodnotami v podobe inkrementovaného súčtu za každý deň na obrázku 4 a nakoniec je zobrazená simulácia bez skupiny nakaziteľných ľudí na obrázku 5.

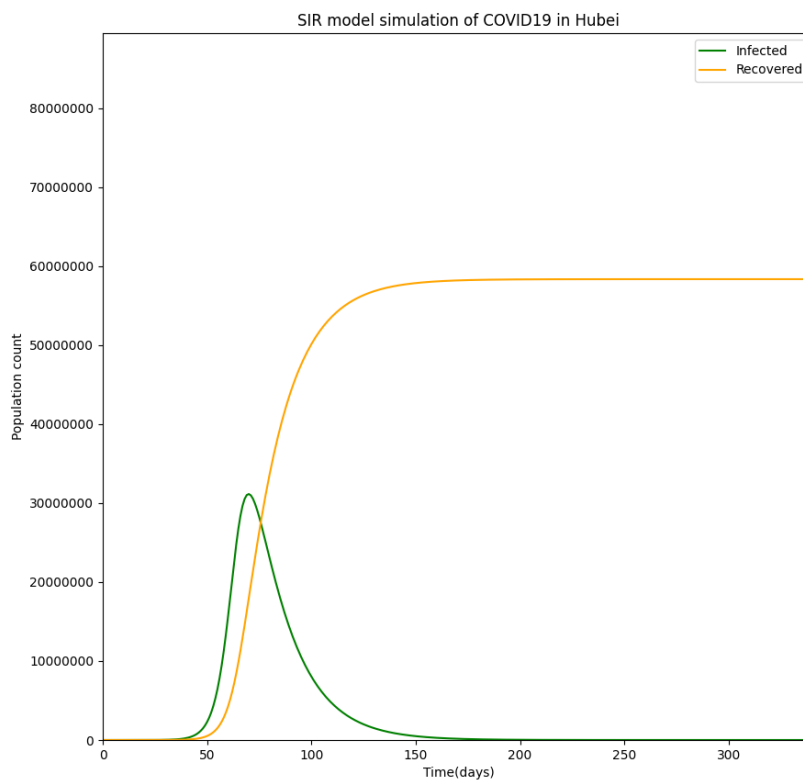
⁴V tomto prípade sú tam zaradení aj tí čo ochoreniu podľahli



Obrázok 3: Výstup experimentu modelu SIR bez opatření



Obrázok 4: Výstup experimentu modelu SIR bez opatrení ako celkový súčet



Obrázok 5: Výstup experimentu modelu SIR bez opatrení so zanedbaním nakazitelných ľudí

4.2 Experiment 2

Tento experiment rovnako ako *Experiment 1* popisuje vývoj ochorenia pomocou modelu SIR, ale v simulácii boli použité predom spomínané opatrenia. Výstup sa stále nepribližuje dostatočne k realite a k presnému odhadu priebehu ochorenia, to má však za úlohu model SEIRD, ktorý je použitý v nasledujúcich experimentoch[13].

Vstupné hodnoty:

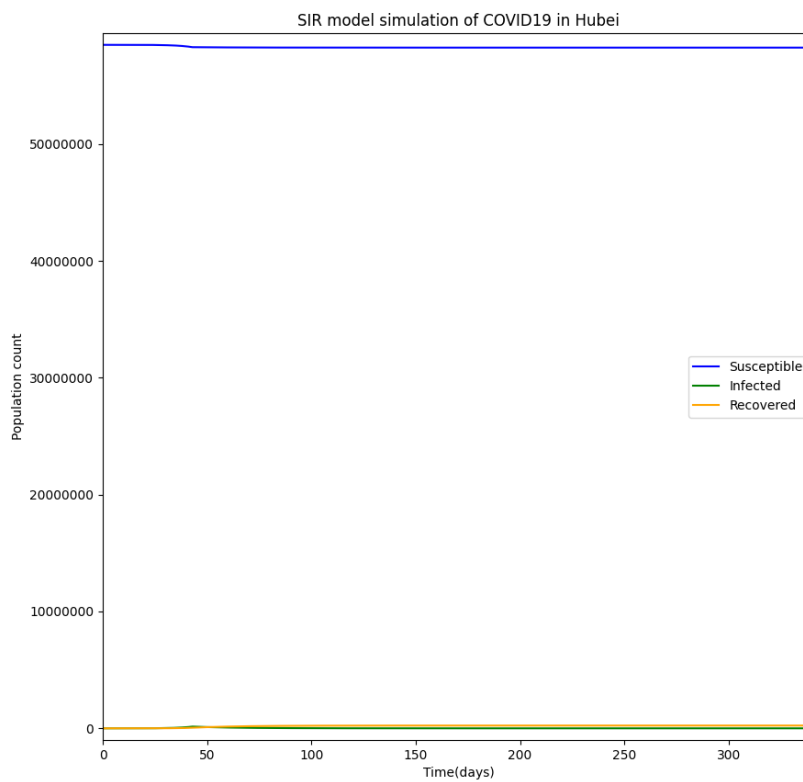
- Populácia = 58,500,000

- Infikovaní = 27
- Koeficient $\beta = 0.31144$
- Koeficient $\alpha = 0.05560$

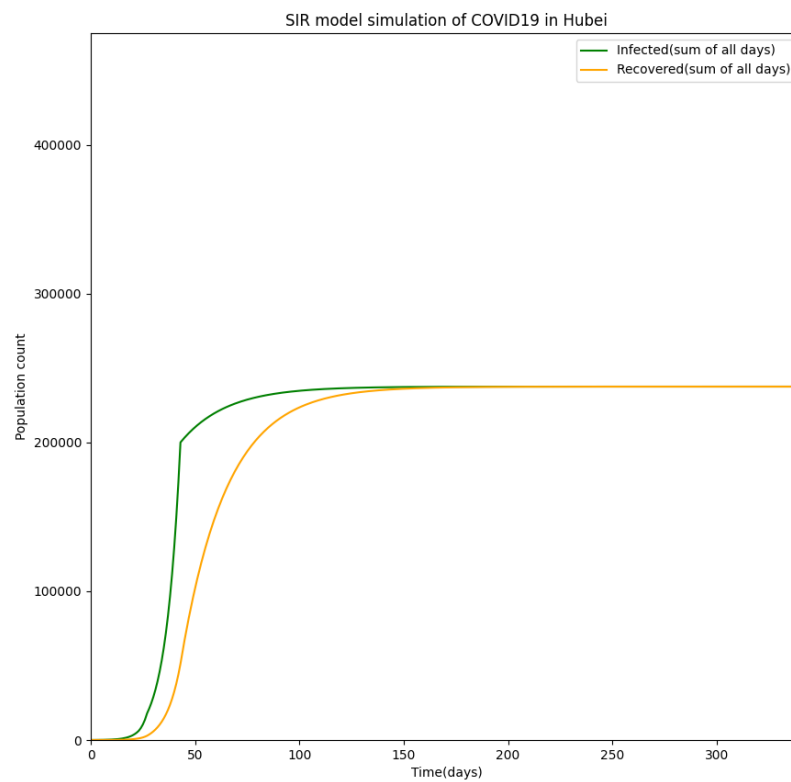
Výstupné hodnoty:

- Celkový počet infikovaných = 237,602 (0.40616%)
- Celkový počet zotavených = 237,602 (0.40616%)
- Najväčší denný prírastok = 19,816 (deň č. 44)
- Najväčší počet infikovaných v jeden moment = 14,8858 (deň č. 44)

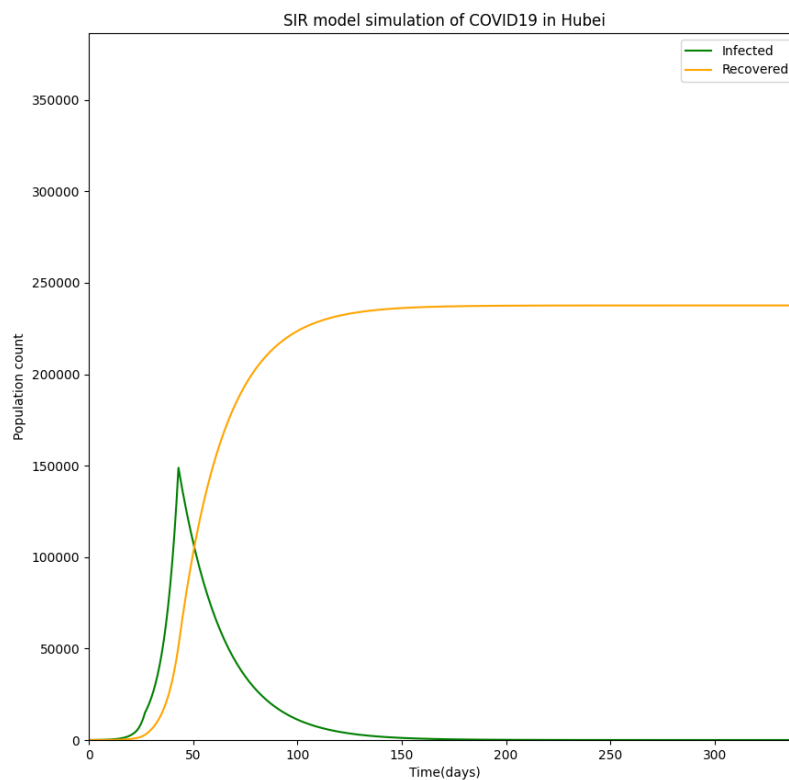
Napriek určitej nepresnosti modelu SIR, je aj takýmto experimentom badateľné, aký efekt majú spomínané protiepidemiologické opatrenia. Počet nakazených sa zredukoval približne o 99.5% čo vykresľuje obrázok 6. Na obrázku 7 je veľmi dobre vidieť, že exponenciálny nárast sa takmer úplne zmenil a skokovo prešiel do logaritmického vývoja, a práve tieto zmeny sú ovplyvnené spomínanou zmenou reprodukčného čísla R_0 , ktoré časom klesne pod hodnotu 1. Obrázok 8 je v podstate iba veľmi detailne priblížený obrázok 6, vďaka čomu vidíme maximálny bod nárastu ochorenia a jeho zmenu.



Obrázok 6: Výstup experimentu modelu SIR s opatreniami



Obrázok 7: Výstup experimentu modelu SIR s opatreniami ako celkový súčet



Obrázok 8: Výstup experimentu modelu SIR s opatreniami so zanedbaním nakazitelných ľudí

4.3 Experiment 3

Experiment je popísaný pomocou modelu SEIRD, čo bude mať isté výhody v presnosti simulácie. Rovnako ako v *Experimente 1* ani tu neboli použité protiepidemiologické opatrenia. V tomto modeli však stojí za povšimnutie vstupná hodnota tzv. *infekčných* (*angl. exposed*), ktorá bola volená na základe vedeckej publikácie, ktorá uvádza, že počet $E = 20 * I$ [3].

Vstupné hodnoty:

- Populácia = 58,500,000
- Infikovaní = 27

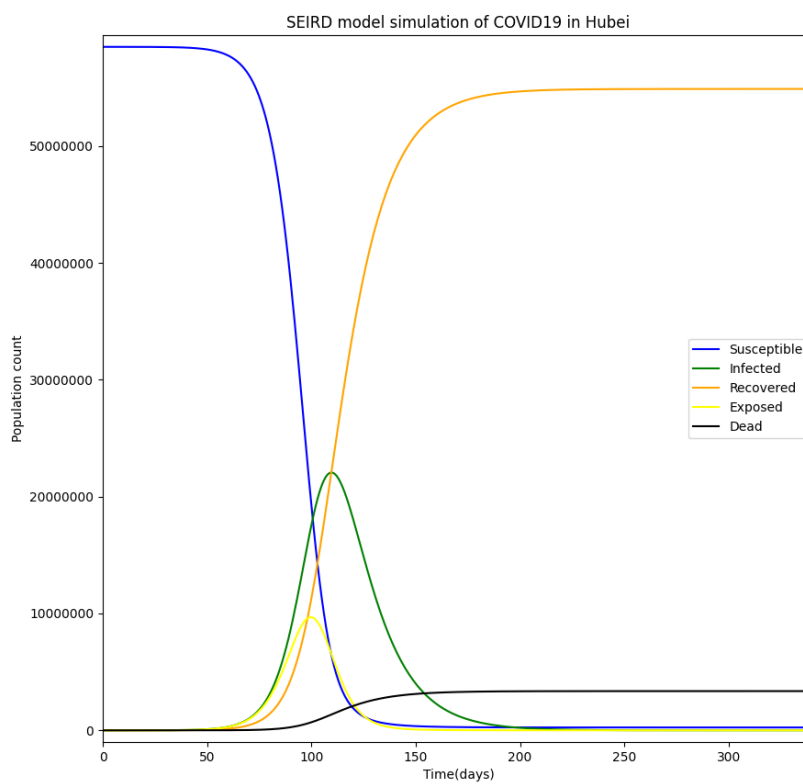
- Infekční⁵ = 540
- Koeficient $\beta = 0.31144$
- Koeficient $\alpha = 0.05560$
- Koeficient $\sigma = 0.19230$
- Koeficient $\omega = 0.0034$

Výstupné hodnoty:

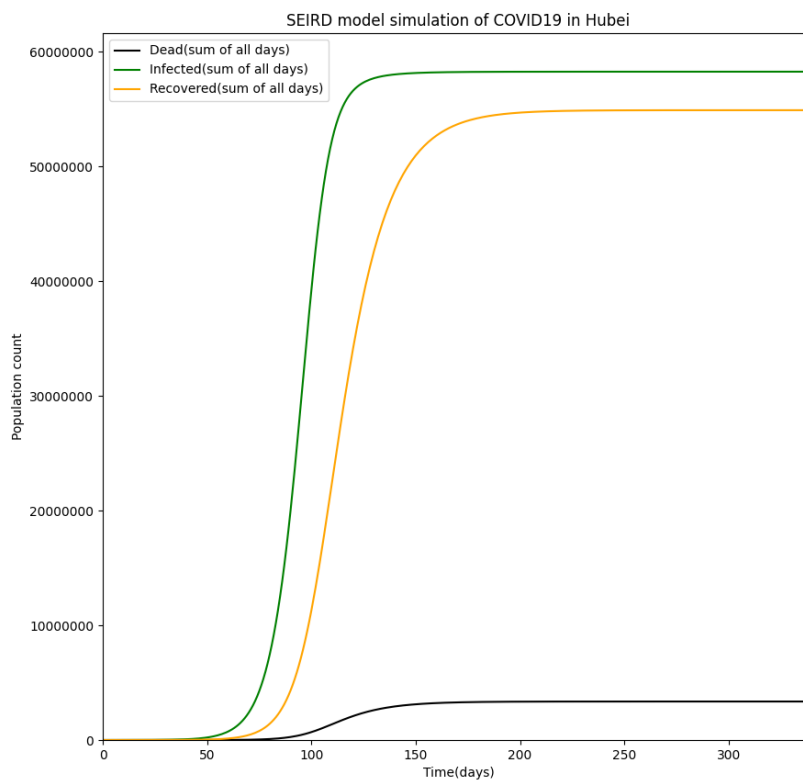
- Celkový počet infikovaných = 58,415,480 (99.85552%)
- Celkový počet zotavených = 55,049,651 (94.10197%)
- Najväčší denný prírastok = 1,177,013 (deň č. 89)
- Najväčší počet infikovaných v jeden moment = 24,172,919 (deň č. 101)
- Počet úmrtí = 3,366,346

Výsledkom experimentu bol podľa očakávaní znovu drastický dopad šírenia nákazy v provincii Hubei. V tomto prípade máme oveľa detailnejší pohľad dopadu ochorenia, ktorým je predovšetkým počet úmrtí. V takomto prípade zistíme, že aby sa predošlo k takémuto pandemickému vyvrcholeniu, je nutné zaviesť opatrenia čo bude popisovať *Experiment 4*. Obrázok 9 nám zobrazuje každú skupinu z modelu SEIRD a jej vývoj počas uvedeného obdobia. Obrázok 10 približuje rovnako ako obrázok 4 súčet ľudí v každej skupine čo sa nakazili, vyliečili a tí ktorí podľahli ochoreniu *SARS-CoV-2*. Obrázok 11 neobsahuje množinu ľudí, ktorí sa môžu nakaziť.

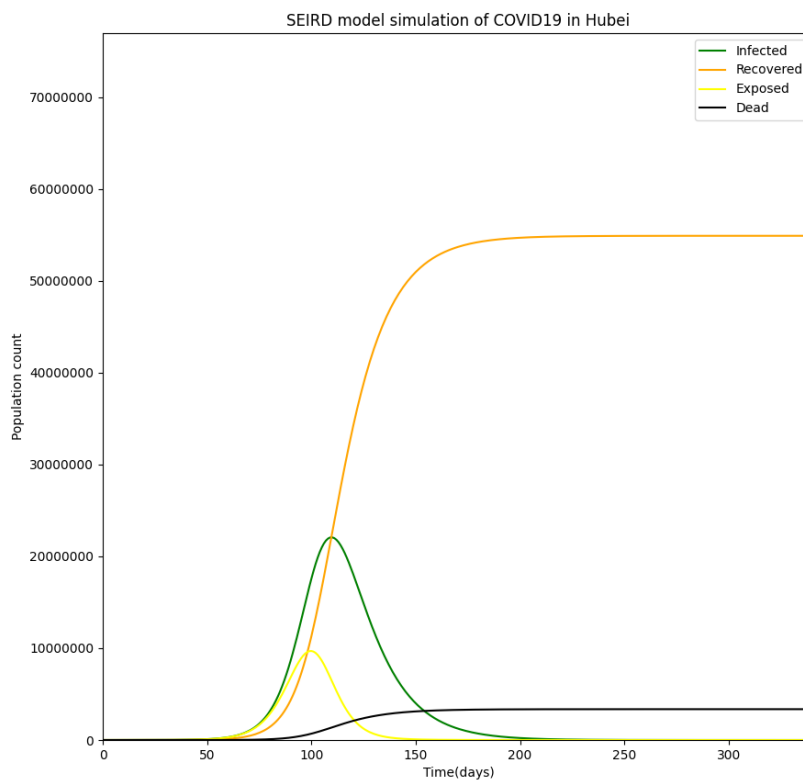
⁵Je to skupina ľudí z množiny E(exposed)



Obrázok 9: Výstup experimentu modelu SEIRD bez opatrení



Obrázok 10: Výstup experimentu modelu SEIRD bez opatrení ako celkový súčet



Obrázok 11: Výstup experimentu modelu SEIRD bez opatrení so zanedbaním nakaziteľných ľudí

4.4 Experiment 4

Posledný experiment sa takmer identicky približuje k priebehu ochorenia *SARS-CoV-2* v čínskej provincii Hubei. Zvolený model SEIRD je schopný pomerne presne pracovať so správne navolenými vstupnými hodnotami a parametrami, čo nám dokazuje aj zhoda so skutočnosťou[14].

Vstupné hodnoty:

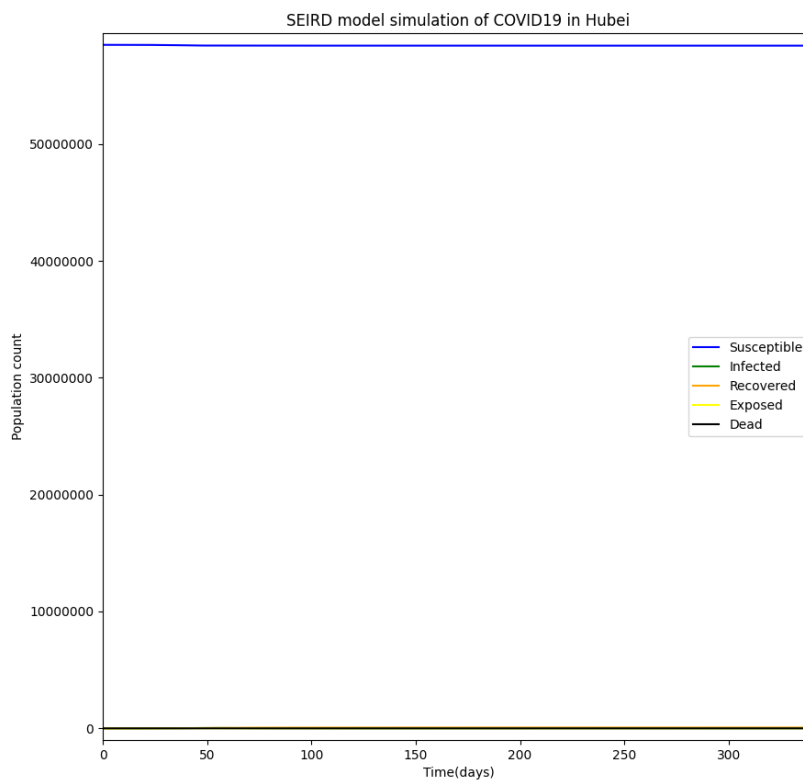
- Populácia = 58,500,000
- Infikovaní = 27

- Infekční = 540
- Koeficient $\beta = 0.31144$
- Koeficient $\alpha = 0.05560$
- Koeficient $\sigma = 0.19230$
- Koeficient $\omega = 0.0034$

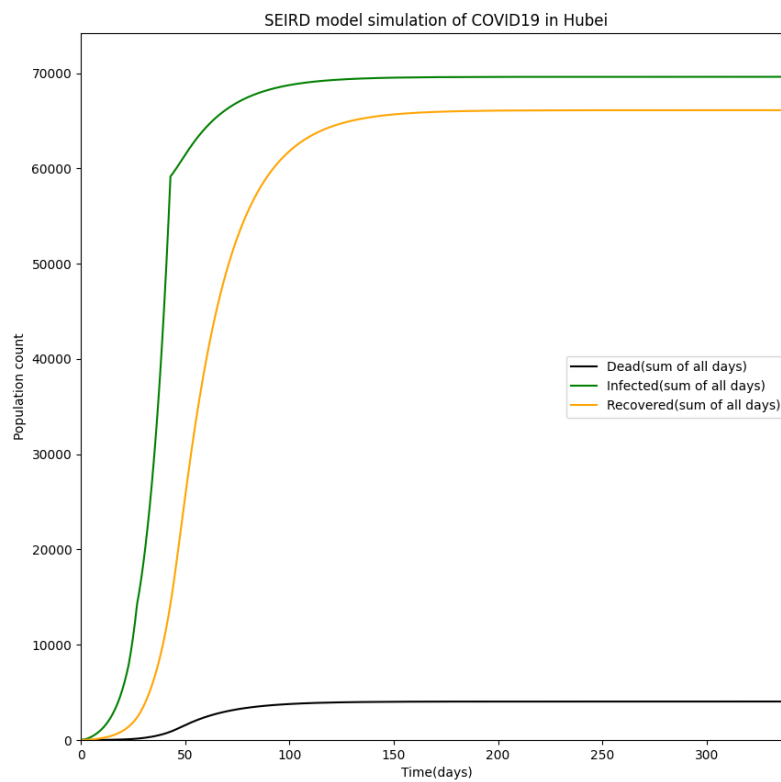
Výstupné hodnoty:

- Celkový počet infikovaných = 69,625 (0.11902%)
- Celkový počet zotavených = 66,122 (0.11303%)
- Najväčší denný prírastok = 2183 (deň č. 45)
- Najväčší počet infikovaných v jeden moment = 30,116(deň č. 48)
- Počet úmrtí = 4043

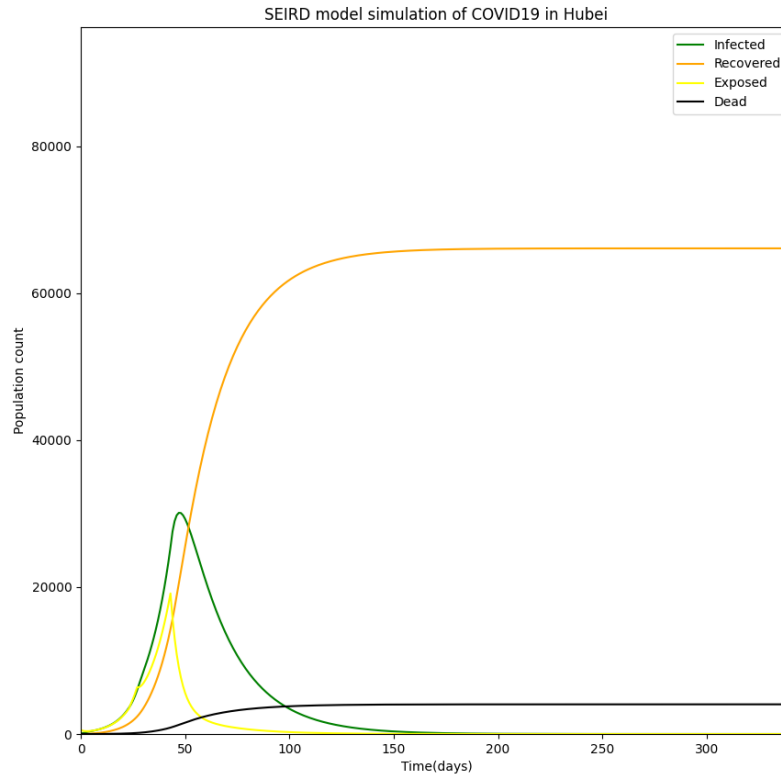
Radikálny prístup k opatreniam dokázal s veľkou pravdepodobnosťou predísť k odhadovaným veľkým fatalitám v populácií, na ktoré sme narazili v *Experimente 3*. Počet infikovaných sa znížil takmer o 99.9% čo je veľmi pozitívny výsledok, a rovnako tak aj počet fatalít, ktorý sa zredukoval cez viac ako 3 milióny ľudí. Obrázok 12 znázorňuje na 58 miliónovej vzorke ľudí takmer zanedbateľné počty nakazených alebo mŕtvych. Obrázok 13 zobrazuje celkové súčty nakazených, uzdravených a mŕtvych za obdobie takmer celého roku. Rovnako si môžeme všimnúť prerušenie exponenciálneho rastu nakazených, čo spôsobil ten istý krok v opatreniach ako v predošlom *Experimente 2*. Obrázok 14 detailnejšie ukazuje vývoj ochorenia na jednotlivých skupinách ľudí modelu SEIRD, avšak zanedbáva skupinu náchylných ľudí (*angl. susceptible*).



Obrázok 12: Výstup experimentu modelu SEIRD s opatreniami



Obrázok 13: Výstup experimentu modelu SEIRD s opatreniami ako celkový súčet



Obrázok 14: Výstup experimentu modelu SEIRD s opatreniami so zanedbaním nakaziteľných ľudí

5 Záver

Odhliadnuc od vedľajších faktorov protiepidemiologických opatrení, akými sú napríklad ekonomický dopad v oblasti Hubei, ktoré boli v skúmaných experimentoch zanedbané je zjavné, že zakročenie vlády pre úplnú karanténu provincie a dôkladný monitoring ľudí, ktorí sa v provincii presúvali bolo základom na zastavenie šírenia nákazy. Naopak je možné vyvodiť, že zvýšená mobilita ľudí, aká nastala počas obdobia slávností Čínskeho nového roku veľmi negatívne ovplyvnila reprodukčné číslo ochorenia *SARS-CoV-2*, čo malo za následok prudký vzostup nakazených ľudí čo paralelne prispievalo k zvýšenej mortalite. Z po-

hľadu ochrany ľudí pred spôsobenou fyzickou ujmom z dôvodu ochorenia bolo zavedenie rýchleho lockdownu s veľkou pravdepodobnosťou tou najlepšou alternatívou.

Zdroje

- [1] Ian Cooper, Argha Mondal, and Chris G. Antonopoulos. “A SIR model assumption for the spread of COVID-19 in different communities”. In: *Chaos, Solitons Fractals* 139 (2020), p. 110057. ISSN: 0960-0779. DOI: <https://doi.org/10.1016/j.chaos.2020.110057>. URL: <http://www.sciencedirect.com/science/article/pii/S0960077920304549>.
- [2] Ivan Korolev. “Identification and estimation of the SEIRD epidemic model for COVID-19”. In: *Journal of Econometrics* (July 2020). DOI: [10.1016/j.jeconom.2020.07.038](https://doi.org/10.1016/j.jeconom.2020.07.038). URL: <https://www.sciencedirect.com/science/article/pii/S0304407620302621>.
- [3] Tae Wuk Bae, Kee Koo Kwon, and Kyu Hyung Kim. *Mass Infection Analysis of COVID-19 Using the SEIRD Model in Daegu-Gyeongbuk of Korea from April to May, 2020*. Aug. 2020. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7458854/>.
- [4] M. Liu et al. *Modelling the evolution trajectory of COVID-19 in Wuhan, China: experience and suggestions*. May 2020. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0033350620301542>.
- [5] Z. Shi L-F. Wang et al. *Mathematical models for devising the optimal SARS-CoV-2 strategy for eradication in China, South Korea, and Italy*. Sept. 2020. URL: <https://translational-medicine.biomedcentral.com/articles/10.1186/s12967-020-02513-7>.
- [6] D. Klinkenberg JA. Backer et al. *Transmission patterns of COVID-19 in the mainland of China and the efficacy of different control strategies: a data- and model-driven study*. July 2020. URL: <https://idpjournals.biomedcentral.com/articles/10.1186/s40249-020-00709-z>.
- [7] *Archived: WHO Timeline - COVID-19*. Apr. 2020. URL: <https://www.who.int/news/item/27-04-2020-who-timeline---covid-19>.
- [8] Nourah S. Altakarli. “China’s Response to the COVID-19 Outbreak: A Model for Epidemic Preparedness and Management”. In: *Dubai Medical Journal* 3.2 (2020), pp. 44–49. DOI: [10.1159/000508448](https://doi.org/10.1159/000508448). URL: <https://www.karger.com/Article/FullText/508448>.
- [9] S. Pei R. Li et al. *Basic reproduction number and predicted trends of coronavirus disease 2019 epidemic in the mainland of China*. July 2020. URL: <https://idpjournals.biomedcentral.com/articles/10.1186/s40249-020-00704-4>.

- [10] Kiesha Prem et al. “The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study”. In: *The Lancet Public Health* 5.5 (Mar. 2020). DOI: [10.1016/s2468-2667\(20\)30073-6](https://doi.org/10.1016/s2468-2667(20)30073-6). URL: [https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667\(20\)30073-6/fulltext](https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(20)30073-6/fulltext).
- [11] Wenjun Du et al. “Epidemic update of COVID-19 in Hubei Province compared with other regions in China”. In: *International Journal of Infectious Diseases* 95 (Apr. 2020), pp. 321–325. DOI: [10.1016/j.ijid.2020.04.031](https://doi.org/10.1016/j.ijid.2020.04.031).
- [12] Fenglin Liu et al. *Predicting and analyzing the COVID-19 epidemic in China: Based on SEIRD, LSTM and GWR models*. Aug. 2020. URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0238280>.
- [13] *Hubei Timeline*. URL: <https://coronavirus.jhu.edu/data/hubei-timeline>.
- [14] Nov Thomala Published by Lai Lin. *Greater China: coronavirus statistics by region*. Nov. 2020. URL: <https://www.statista.com/statistics/1090007/china-confirmed-and-suspected-wuhan-coronavirus-cases-region/>.