

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/344711334>

Some approaches to the use of stochastic models of epidemiology to the problem COVID-19

Preprint · May 2020

CITATIONS

0

READS

145

8 authors, including:



P. S. Knopov

National Academy of Sciences of Ukraine

220 PUBLICATIONS 577 CITATIONS

SEE PROFILE



Vladimir I. Norkin

National Academy of Sciences of Ukraine

111 PUBLICATIONS 1,732 CITATIONS

SEE PROFILE



K.L. Atoyev

V. M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukr...

104 PUBLICATIONS 84 CITATIONS

SEE PROFILE



Vasyl Gorbachuk

National Academy of Sciences of Ukraine

189 PUBLICATIONS 171 CITATIONS

SEE PROFILE

Деякі підходи використання стохастичних моделей епідеміології до проблеми COVID-19

Вступ. Пандемія Covid-19 ставить перед світовим суспільством серйозні виклики, що потребують адекватних методів та засобів її контролю. З одного боку, при відсутності у населення імунітету, вакцини та ліків від коронавірусу, зважаючи на тяжкі наслідки захворювання, епідемія являє собою дуже значну загрозу для життя і здоров'я людей. З другого боку, введення довготермінового карантину та заходів по її обмеженню спричиняє колосальні економічні збитки, зупиняючи економічне життя країн та континентів. Тому рішення по контролю за розповсюдженням хвороби потребують особливої виваженості, адже з одного боку терезів стоїть життя і здоров'я великої кількості людей, з іншого – дуже значні економічні втрати та потенційне зuboжіння населення.

В таких умовах зростає необхідність в засобах моделювання і підтримки прийняття рішень, що спираються на точні обрахунки їх наслідків. Такі засоби повинні включати різноманітні моделі по прогнозу епідеміологічної ситуації і потреб медичної допомоги населенню, моделі прогнозування економічних наслідків урядових (регіональних) рішень по обмеженню епідемії, тощо. Також необхідно враховувати різноманітні ризики та невизначеності, що виникають при моделюванні подібних складних процесів зі стохастичною (невизначеною) природою їх компонент. Це потребує для моделювання відповідної математичної техніки, зокрема, випадкових процесів і полів, стохастичних диференційних рівнянь, регресій спеціального вигляду, тощо. Враховуючи недетермінований характер пов'язаних економічних збитків, для їх оцінювання видається важливим використання сучасного апарата мір ризику.

В якості початкових в даному тексті обрані SIR (SEIR) подібні епідеміологічні моделі, що дозволяють прогнозувати вплив на динаміку процесу розповсюдження хвороби обмежувальних мір по його контролю. В них головним фактором є коефіцієнт реплікації хвороби, що суттєво залежить від таких мір. Так, в [1] вплив на щоденну швидкість розповсюдження захворювання оцінювався для: закриття шкіл; закриття робочих місць; відміни публічних подій; заборони суспільного транспорту; інформаційних компаній; обмеження внутрішніх та міжнародних переміщень.

Основною проблемою для подібних моделей є труднощі їх налаштування (ідентифікації) по реальних даних. Більш детальні моделі потребують більш повних даних щодо профілю захворювання та його розповсюдження. В детермінованих моделях більшість їх параметрів описує середні значення тих, чи інших величин. Вони враховують стохастичність процесів в середньому, успішно апроксимуючи ситуацію в великих популяціях в однорідному середовищі. Стохастичні моделі, на відміну від їх детермінованих аналогів, дозволяють більш адекватно відобразити поведінку процесів, особливо в локальних чи перехідних процесах.

В планах наших досліджень основну увагу буде приділено вивченню саме стохастичних моделей епідеміологічних процесів та оцінювання пов'язаних економічних збитків. Передбачається використання апарата випадкових процесів і полів, стохастичних диференційних рівнянь, регресій спеціального вигляду та мір ризику, які активно розвиваються в Інституті кібернетики. В цьому тексті ми викладемо деякі з висновків, отриманих з початкового розгляду детермінованих епідеміологічних моделей і застосування регресій спеціального вигляду щодо кількості інфікованих осіб.

Регресії з перемиканнями. Передусім розглянемо застосування до спостережень числа інфікованих коронавірусом техніки регресій з перемиканнями, в яких точки перемикання невідомі, тому повинні бути оцінені.

Суть регресії з перемиканнями полягає в тому, що її регресії постійні не на всьому інтервалі спостереження, а лише на його підінтервалах, що розділяються один від одного точками перемикання. Оцінивши точки перемикання, можна визначити періоди часу, в

яких відбувалися структурні зміни об'єкта.

В [2] введено новий клас регресій з перемиканнями в неперервному часі, який потім було поширено на дискретному випадок. Зауважимо, що регресії з перемиканнями на відміну від класичних регресій, чи нейронних мереж, наприклад з [3], дозволяють працювати з нестационарними спостереженнями, вловлюючи зміни структури даних в точках перемикання. Випадкові точки перемикання можуть відповідати, наприклад, локальним спалахам епідемічного процесу.

Розглянемо використання цього класу регресій для статистичного аналізу поширення коронавірусної інфекції в Україні за даними про кількість інфікованих з 23.03.20 по 24.04.20. Через наявність двох аномальних спостережень була створена робастна регресія з перемиканнями, представлена на рис. 1.

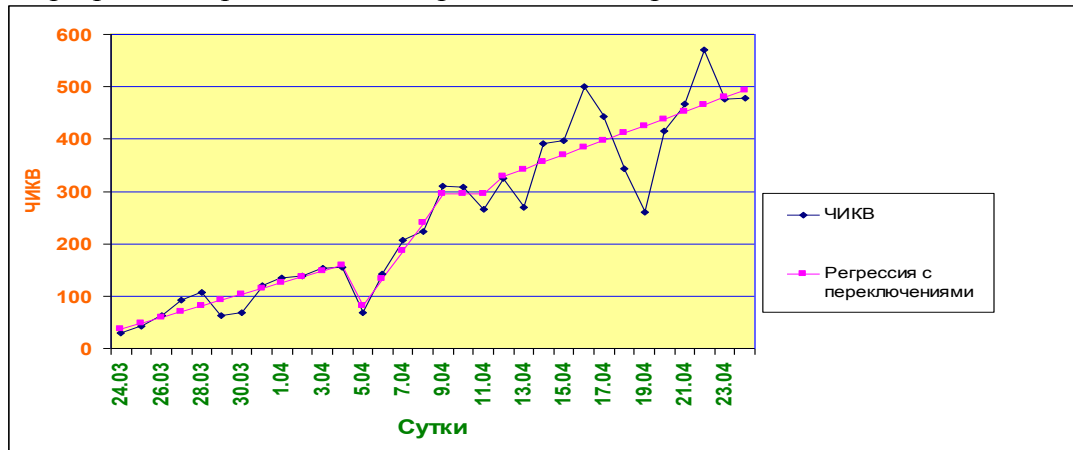


Рис. 1. ЧИКВ - число інфікованих коронавірусом (дані МОЗ України) і його апроксимація регресією з перемиканнями: 24.03 - 24.04.

На ньому видно, що може бути істотна різниця між ЧИКВ в двох сусідніх добах. Тому при побудові регресії припускалися розриви лінії регресії в точках перемикання, якими виявилися: 4.04.20; 8.04.20; 11.04.20.

Кут нахилу прямої регресії на інтервалі часу [12.04; 24.04] виявився незначущим на 1% -му рівні і значущим на 5% -му рівні, це може говорити про існування точки перемикання на цьому інтервалі часу, яку не вдалося виявити через нестачу даних. Проте, результат прогнозу на 8 діб за отриманою регресії виявився досить точним, див. рис. 2.

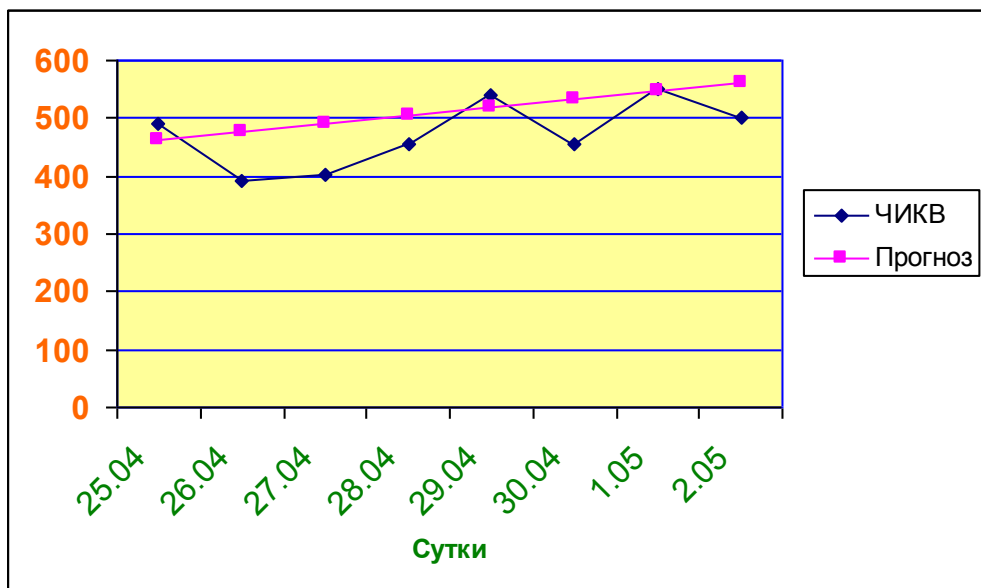


Рис. 2. Фактичні величини ЧИКВ та їх прогноз після 24.04.

Фактичне середнє квадратичне відхилення (с.к.в.) прогнозу дорівнює 56.51, що приблизно на 30% менше теоретичного с.к.в. прогнозу, розрахованого по регресії.

Згідно рис. 2, схоже, що швидкість зростання числа інфікованих почала зменшуватися в порівнянні з тією, яка була на початку 20 – х чисел квітня.

Потім отримана модель уточнювалась за даними про кількість інфікованих включно до 9.05.20. Для цього початок регресії з перемиканнями з точки 23.03 було зсунуто в точку 12.04 – початок останнього інтервалу постійності параметрів регресії, наведеної на рис. 1. Таким чином, розглядався інтервал спостережень [12.04; 9.05].

З початку було висунуто припущення, що на цьому інтервалі триває зафіксований раніше слабкий підйом, потім слідує «плато» і далі спад. Таким чином, регресія повинна мати 2 точки перемикання. Розрахунки показали, що є слабке зростання, а точки перемикання є точками розриву прямої у випадкові періоди часу. У них через велику варіацію даних відбувалося різке зменшення числа інфікованих. Тому, припустивши, що є тільки одна точка перемикання, після якої починається або спад, або «плато», були проведені розрахунки, які дали регресію, наведену на рис.3. С.к.в. цієї регресії виявилось меншим, ніж у регресії з двома точками перемикань.

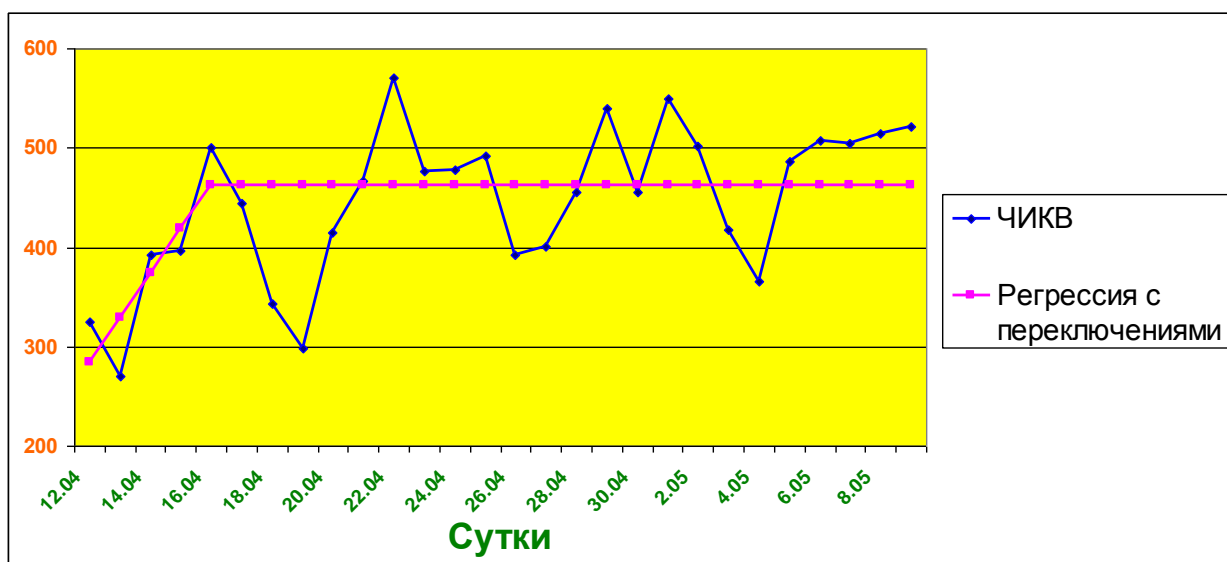


Рис. 3. ЧИКВ – число інфікованих коронавірусом (дані МОЗ України) і його апроксимація регресією з перемиканнями: 12.04 – 9.05.

Згідно рис. 3 регресія є сплайн з горизонтальною ділянкою – «плато», яка починається з точки перемикання 16.04. Перехід від підйому до «плато» відбувається плавно. Прогноз в даному випадку є екстраполяцією горизонтальної прямої на майбутні періоди часу, розрахункове с.к.в. прогнозу 68,31. Результати прогнозу наведені в табл. 1.

| Доба | ЧИКВ, y_t | Прогноз, \hat{y}_t | Похибка прогнозу $y_t - \hat{y}_t$ |
|-------|-------------|----------------------|------------------------------------|
| 10.05 | 416 | 463.0901 | -47.09 |
| 11.05 | 375 | 463.0901 | -88.09 |

Таблиця 1

Фактичне с.к.в. прогнозу 70.63, що трохи більше його розрахункової величини. Помилка прогнозу на дві доби – негативна, що може бути непрямим доказом спаду числа інфікованих, що починається.

Його надійне визначення можна буде зробити аналогічно попередньому кроку: зсунути регресію вправо в уже відому точку перемикання і знову побудувати регресію з перемиканнями. При цьому на кожному кроці, крім прогнозу, визначається тренд числа інфікованих, що важливо для прийняття рішень щодо посилення або ослаблення

карантину.

Описана вище процедура обчислень може бути автоматизована, що дозволить обробляти дані в реальному масштабі часу.

Регресію з перемикальними можна буде також використовувати для визначення тривалості лікування інфікованих коронавірусом, а також для управління протіканням різних епідемій. Можливо, її вдасться використовувати для визначення моментів втручання при лікуванні, коли використовуються різні схеми лікування.

Зауваження. Регресійні та інші подібні моделі (моделі прогнозування в вигляді нейронних мереж, є загально прогностичними моделями і не завжди враховують специфіку епідемічних процесів і тому можуть застосовуватися тільки для коротко часового прогнозування епідемії. Більш специфічні регресійні моделі можуть прогнозувати так зване репродуктивне число епідемії як функцію параметрів популяції та інтенсивності вживаних протиепідемічних заходів.

Опис епідеміологічних моделей. Згідно [4] модель **SIR** розподіляє населення на категорії (compartments): *S* (**Susceptible**, сприйнятливі), *I* (**Infectious**, інфіковані) та *R* (**Removed**, особи, що перехворіли), динаміка яких описуються у вигляді відповідних систем диференціальних рівнянь:

$$\begin{aligned}\frac{dS}{dt} &= -\beta \frac{1}{N_0} SI \\ \frac{dI}{dt} &= \beta \frac{1}{N_0} SI - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}$$

де β та γ – так звані швидкості передачі та видалення (відновлення) відповідно, а N_0 – кількість населення, що знаходиться під загрозою хвороби.

Часто в якості змінних обираються не абсолютні величини *S*, *I*, *R*, а їх пропорції щодо N_0 , тобто $S_* = S / N_0$; $I_* = I / N_0$; $R_* = R / N_0$. Тоді система переписується у компактнішому вигляді.

Перейдемо до більш звичних позначень $T_{\text{inf}} = 1 / \gamma$, $R_0 = \beta / \gamma$, де T_{inf} – середній період, під час якого інфікована персону розповсюджує хворобу, а R_0 – відповідно коефіцієнт реплікації захворювання, тобто середня кількість заражень однією хворою людиною. Тоді $\gamma = 1 / T_{\text{inf}}$, $\beta = R_0 \gamma = R_0 / T_{\text{inf}}$, отже, маємо рівняння

$$\begin{aligned}\frac{dS}{dt} &= -\frac{R_0}{T_{\text{inf}}} \frac{1}{N_0} SI \\ \frac{dI}{dt} &= \frac{R_0}{T_{\text{inf}}} \frac{1}{N_0} SI - \frac{1}{T_{\text{inf}}} I \\ \frac{dR}{dt} &= \frac{1}{T_{\text{inf}}} I\end{aligned}\tag{1}$$

Зрозуміло, що розподіл населення по відповідним категоріям задовольняє загальному балансу $N_0(t) = S(t) + I(t) + R(t)$. Тому на середньо- та довгостроковому періоді при моделюванні необхідно враховувати динаміку кількості населення $N_0(t)$ (демографію), зважаючи на кількість народжуваних за одиницю часу b та коефіцієнт смертності населення μ :

$$\begin{aligned}
\frac{dS}{dt} &= -\frac{R_0}{T_{\text{inf}}} \frac{1}{N_0} SI + b - \mu S \\
\frac{dI}{dt} &= \frac{R_0}{T_{\text{inf}}} \frac{1}{N_0} SI - \frac{1}{T_{\text{inf}}} I - \mu I . \\
\frac{dR}{dt} &= \frac{1}{T_{\text{inf}}} I - \mu R
\end{aligned}
\tag{2}$$

Модель **SEIR** вводить до розгляду додаткову категорію E (**Exposed**) для осіб, що перебувають в інкубаційному періоді, коли спочатку особа тільки інфікувалась, але не є заразною. До попередніх рівнянь, крім динаміки категорії $E(t)$, додається ще й параметр T_{inc} – середній інкубаційний період, а система має вигляд:

$$\begin{aligned}
\frac{dS}{dt} &= -\frac{R_0}{T_{\text{inf}}} \frac{1}{N_0} SI \\
\frac{dE}{dt} &= \frac{R_0}{T_{\text{inf}}} \frac{1}{N_0} SI - \frac{1}{T_{\text{inc}}} E \\
\frac{dI}{dt} &= \frac{1}{T_{\text{inc}}} E - \frac{1}{T_{\text{inf}}} I \\
\frac{dR}{dt} &= \frac{1}{T_{\text{inf}}} I
\end{aligned}
\tag{3}$$

В таку модель можна ввести змінну D , що описує смертність, спричинену інфікуванням [5]. Тоді третє рівняння (3) змінюється на наступне

$$\frac{dI}{dt} = \frac{1}{T_{\text{inc}}} E - \frac{1}{T_{\text{inf}}} (1 + \delta) I$$

та додається ще п'яте

$$\frac{dD}{dt} = \frac{1}{T_{\text{inf}}} \delta I ,$$

в яких параметр $\delta = Cfp / (1 - Cfp)$, де Cfp описує середню пропорцію фатальних (летальних) випадків для інфікованих.

Описані моделі можуть в подальшому бути уточнені шляхом додавання нових категорій. Наприклад, додаванням категорій осіб, що підлягають госпіталізації та потребують апаратів ШВЛ, тощо.

Проте детальні моделі потребують відповідного налаштування. Навіть проста модель (3) потребує даних по розподілу інфікованих персон на категорії E та I , тобто на тих, хто перебуває в інкубаційному періоді, та тих, хто вже вийшов з нього.

Оскільки тестування осіб щодо наявності інфікування не можна вважати достатнім, в модель (1) для налаштування можна ввести ще один параметр. Скажімо, офіційні дані відображають лише певну частку P_{inf} інфікованих осіб, а замість реальних величин I відомі лише $I_* = P_{\text{inf}} I$. Тоді, вводячи позначення $I = I_* / P_{\text{inf}}$, отримаємо

$$\begin{aligned}\frac{dS}{dt} &= -\frac{R_0}{T_{\text{inf}}} \frac{1}{N_0} S \frac{1}{P_{\text{inf}}} I_* \\ \frac{dI_*}{dt} &= \frac{R_0}{T_{\text{inf}}} \frac{1}{N_0} S I_* - \frac{1}{T_{\text{inf}}} I_* \\ \frac{dR}{dt} &= \frac{1}{T_{\text{inf}} P_{\text{inf}}} I_*\end{aligned}\quad (4)$$

Тепер параметри $R_0, T_{\text{inf}}, P_{\text{inf}}$ системи (4) можна оцінити за близькістю траєкторії $I_*(t)$ та $S(t)$ до відповідних даних спостережень. Оскільки ми можемо оцінити їх лише за траєкторією $I_*(t)$, параметр P_{inf} ні на що не буде впливати. Отже, (4) не відрізнятиметься від системи (1), тому будемо вивчати саме останню.

Зауваження. Описані SIR моделі є спрощеними версіями більш загальної моделі, що враховує історію хвороби (age-of-infection models), див. (Братусь та ін. (2010)). Ці моделі отримані із більш загальної моделі при наступних суттєвих припущеннях:

число контактів індивіда в одиницю часу постійно і дорівнює $\beta = R_0/T_{\text{inf}}$;

проміжок часу, коли хворий може заражати інших індивідів (тобто час хвороби) нескінченний та розподілений за експоненційним законом з середнім значенням T_{inf} .

Проте SIR подібні моделі загалом вірно описують загальний хід епідемій і широко застосовуються в епідеміологічному моделюванні [1,4,7-9].

Налаштування моделі SIR. Звернемося до системи (1), яку будемо налаштовувати за даними спостережень активних інфікованих $y_i, i=1, \dots, n$, які визначаються як добова “загальна кількість інфікованих” мінус “загальна кількість тих, що перехворіли”. Оскільки досліджуваний процес динамічно змінюється, в якості критерія будемо розглядати суму зваженого модуля відхилення траєкторії від точок спостережень. Величини ваг підбираються в залежності від плину часу, зі збільшенням їх від старих спостережень до нових. Так, для спостережень $y_i, i=1, \dots, n$ такий критерій має вигляд:

$$\text{Crit} = \sum_{i=1}^n w_i |I(t) - y_i|, w_i = 2t / (n(n+1)). \quad (5)$$

Зафіксуємо чисельність населення України на час моделювання на рівні $N_0 = 41858$ тис. чол. (за деякими оцінками на 01.03.2020 року). Можна було б врахувати динаміку чисельності населення згідно моделі (2), проте це не важливо.

Зауважимо, що параметр T_{inf} описує середній період, під час якого інфікована особа розповсюджує хворобу. Цей параметр характеризує вірус та організм людини, тому не може довільно змінюватись. За оцінками Оксфордської моделі для Covid-19 $T_{\text{inf}} \approx 4.5$. Тому ми обмежили підбір параметрів умовами $4.4 \leq T_{\text{inf}} \leq 4.6$.

Підбір параметрів R_0, T_{inf} шляхом мінімізації критерію (5) виконувався за допомогою стандартних процедур для обмеженої глобальної оптимізації.

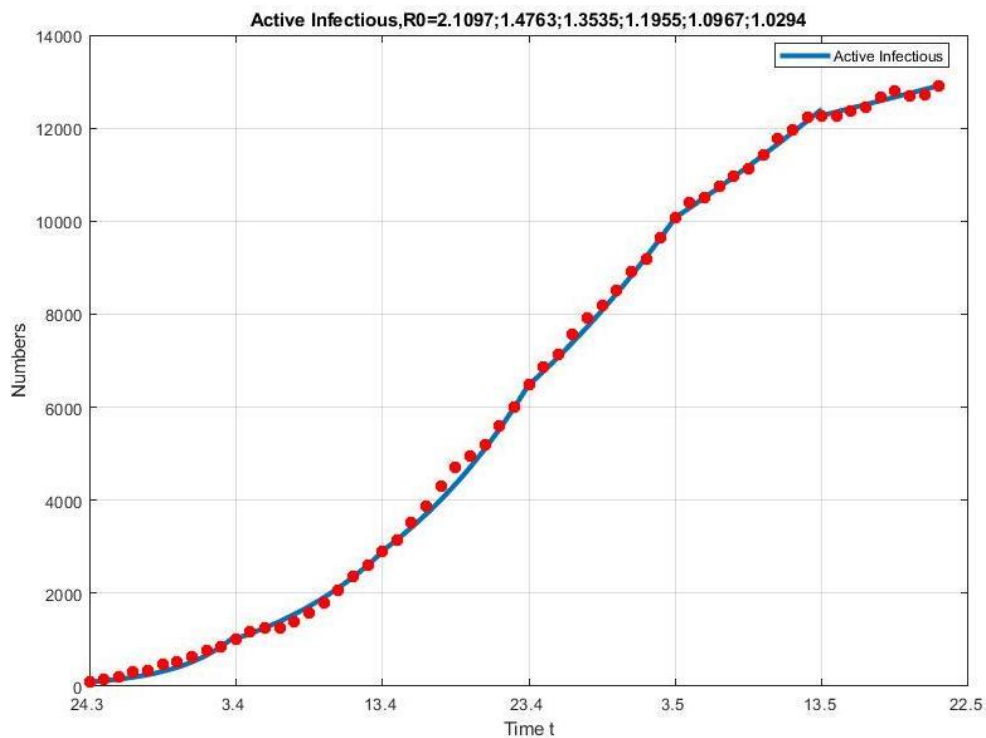
Результати розрахунків. Період спостережень з 24 березня по 20 травня був розподілений на інтервали по 11 днів, для яких були оцінені параметри R_0, T_{inf} системи (1). Вони наведені в наступній таблиці:

| Інтервал | R_0 | T_{inf} | Crit |
|-------------|--------|------------------|----------|
| 24.03–03.04 | 2.1097 | 4.4992 | 78.1169 |
| 03.04–13.04 | 1.4763 | 4.5807 | 57.9253 |
| 13.04–23.04 | 1.3535 | 4.4014 | 116.0376 |
| 23.04–03.05 | 1.1955 | 4.4121 | 64.1395 |
| 03.05–13.05 | 1.0967 | 4.5106 | 62.9927 |

| | | | |
|-------------|--------|--------|---------|
| 13.05–20.05 | 1.0294 | 4.4997 | 65.9745 |
|-------------|--------|--------|---------|

Таблиця 1.

Відповідні часові відрізки траєкторії $I(t)$ з наведеними в таблиці параметрами системи (1) та добові дані щодо активних інфікованих представлені на наступному графіку.

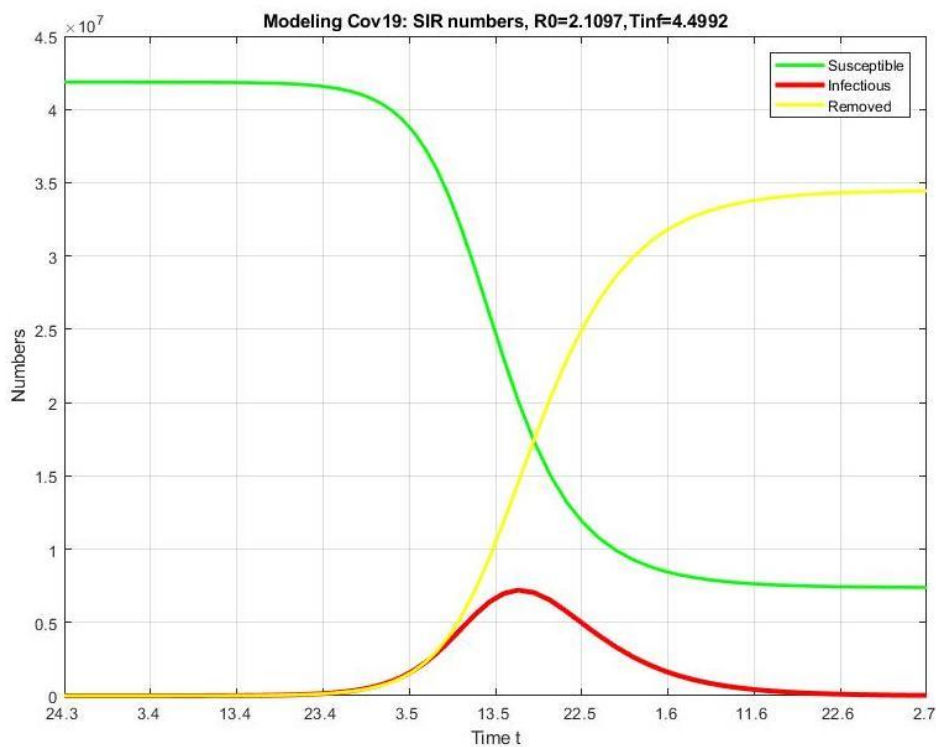


Графік 1

Постановою Кабміну України з 12 березня було введено карантин, який з 6 квітня було посилено. Його дія почала проявлятися з певним часовим лагом, що призвело до поступового зниження коефіцієнта кожного наступного періоду. Такому зниженню також сприяла робота по тестуванню населення та ізоляції потенційно інфікованих осіб.

За допомогою (1) легко змоделювати розповсюдження захворювання без карантину, вважаючи сталими параметри R_0, T_{inf} , що характерні для 1-го періоду. Зрозуміло, що таке моделювання має дуже умовний характер, адже при епідемії люди, уникаючи загрози, навіть без офіційного карантину самостійно обмежують свої контакти і пересування, використовують засоби захисту для себе і родини, тощо. Тому базовий параметр R_0 з плином часу значно зменшується навіть за відсутності карантину.

Результати такого моделювання представлені на наступному графіку.



Графік 2

За графіком населення динамічно інфікується до моменту набуття ним колективного імунітету, потім швидкість інфікування спадає і епідемія поступово згасає. Проте її наслідки, незважаючи на штучність моделювання, могли виявитись важкими, зважаючи на небезпечність коронавірусу, відсутність для нього ліків та велику кількість інфікованих.

В свою чергу, якщо процес захворювань стримується (спадає) лише за рахунок карантину, то при послабленні його умов, що впливають на R_0 , відсутність колективного імунітету у населення загрожує перезапуском процесу розповсюдження хвороби.

Звернемося знову до таблиці 1. Як неважко бачити, коефіцієнт реплікації R_0 за рахунок карантину в кожний наступний період поступово знижується. За останній період він достатньо близький до 1, а саме 1.0294. На разі встановлюється хитка рівність R_0 близько 1.

Відношення між коефіцієнтами реплікації наших часових відрізків становлять:

$$\frac{R_0^2}{R_0^1} = 0.9168; \frac{R_0^3}{R_0^2} = 0.8833; \frac{R_0^4}{R_0^3} = 0.9174; \frac{R_0^5}{R_0^4} = 0.9386,$$

тому при збереженні тенденції наступним значенням може бути $R_0^6 = R_0^5 \times 0.9386 = 0.9662$. При $R_0 < 1$ це призведе до початку затухання процесу розповсюдження захворювання.

Таким чином, при збереженні поточних тенденцій вже в наступний десятиденний період може розпочатись поступове зменшення кількості активних інфікованих.

З іншого боку, планове пом'якшення умов карантину, що розпочалося, збільшить параметр R_0 , отже, може перезапустити процес розповсюдження хвороби.

Модель SIQR. Розглянемо ще один варіант моделей типу SIR, яка не враховує інкубаційний період захворювання, проте додає до останньої ще кілька категорій. SIQR розподіляє населення на наступні основні категорії: S (*Susceptible*, сприйнятливі), I (*Infectious*, інфіковані), Q (*Quarantine*, на карантині, тобто ізольовані або на лікуванні) та R (*Removed*, особи, що перехворіли чи померли). Додатковими категоріями є H (*Hospitalized*, госпіталізовані) та D (*Dead*, померлі). Величини S, I, Q, R, H, D залежать

від часу. В моделі вважається, що число інфікованих та заразних I може бути значно більше (в декілька разів), ніж число виявлених хворих.

Припущення моделі. 1. Число нових інфікованих в одиницю часу (доба, кілька днів, тиждень, тощо) пропорційне добутку $\frac{I}{N_0}S$ з коефіцієнтом α (інтенсивність зараження).

2. Нові інфіковані відразу стають заразними. Певна доля β нових інфікованих в одиницю часу йде на карантин (в результаті тестування, чи епідемічного розслідування, чи в лікарню як хворі), а доля $(1 - \beta)$ нових інфікованих залишається не виявленою.

3. Певна доля γ інфікованих та карантинних в одиницю часу залишає ці категорії осіб та переходить до категорії, що вибули. Цей коефіцієнт може бути різним для інфікованих і карантинних, тобто можуть бути два коефіцієнта γ_I та γ_Q . Вважається, що ті, що вибули, отримують імунітет до хвороби.

4. Вимірюваною величиною є число Q карантинних індивідів.

5. Відомі є початкові значення $S_0 = S(0)$, $Q_0 = Q(0)$, $N_0 = const$ (загальна чисельність населення), а також I_0 (наприклад, $I_0 = Q_0/\beta$, якщо коефіцієнт β вважається відомим, наприклад, в разі масштабного сліпого тестування).

6. Параметри α , γ невідомі і оцінюються зі статистичних даних $Q(t_k)$, $0 < t_1 < \dots < t_k < T$.

7. Додатковими відомими параметрами являються доля госпіталізованих ε від числа карантинних, $H = \varepsilon Q$ (hospitalized, наприклад, приблизно вважаємо $\varepsilon \approx 0.2$), та доля δ померлих від загального числа госпіталізованих, $D = \delta H$ (наприклад, приблизно вважаємо $\delta \approx 0.03$).

Величини S, I, Q, R задовольняють систему звичайних диференціальних рівнянь:

$$\begin{aligned}\frac{dS}{dt} &= -\alpha S \frac{I}{N_0}, \\ \frac{dI}{dt} &= \beta \alpha S \frac{I}{N_0} - \gamma_I I, \\ \frac{dQ}{dt} &= (1 - \beta) \alpha S \frac{I}{N_0} - \gamma_Q Q, \\ \frac{dR}{dt} &= \gamma_I I + \gamma_Q Q.\end{aligned}$$

Додатковими рівняннями є

$$\frac{dH}{dt} = \varepsilon Q, \quad \frac{dD}{dt} = \delta H.$$

Введемо нормовані величини $s = S/N_0$, $i = I/N_0$, $q = Q/N_0$, $r = R/N_0$, $h = H/N_0$, $d = D/N_0$. Для цих величин отримуємо наступну систему рівнянь, інваріантну до розміру популяції:

$$\begin{aligned}\frac{ds}{dt} &= -\alpha si, \\ \frac{di}{dt} &= (1 - \beta) \alpha si - \gamma_i i, \\ \frac{dq}{dt} &= \beta \alpha si - \gamma_q q,\end{aligned}$$

$$\frac{dr}{dt} = \gamma_i i + \gamma_q q.$$

$$\frac{dh}{dt} = \varepsilon q, \quad \frac{dd}{dt} = \delta h.$$

Початкові умови для $t = 0$ мають вигляд:

$$s(0) = 1 - i_0 - q_0, \quad i(0) = i_0, \quad q(0) = i_0 / \beta, \quad r(0) = 0, \quad h(0) = 0, \quad d(0) = 0.$$

Керування параметрами моделі. Параметри α, β можуть вважатися керованими змінними моделі. Вони можуть бути сталими величинами, функціями часу t , або функціями стану i епідемічної системи. Априорі можна вважати, що коефіцієнт α залежить від i та спадає з ростом i , наприклад, $\alpha(i) = \alpha_0(1 - \alpha_1 i)$, $\alpha(i) = \frac{\alpha_0}{1 + \alpha_1 i^{\alpha_2}}$ та інше,

що відображає «страх» перед епідемією і відповідне зменшення інтенсивності контактів, збільшення обережності і т.д.

Моделі із запізненням. Більш адекватною може бути модель, яка враховує запізнення інформації, запізнення проявів хвороби, запізнення в отриманні результатів тестування, зміну у ступені заразності по мірі розвитку хвороби, тривалість лікування і таке інше. [7,10]. Так найпростіша модель SIR із запізненням описується наступною системою рівнянь:

$$\frac{ds(t)}{dt} = -s(t) \sum_{\tau=0}^{\tau_{\max}} \alpha_{\tau} (i(t-\tau) - i(t-\tau-1)),$$

$$\frac{di(t)}{dt} = s(t) \sum_{\tau=0}^{\tau_{\max}} \alpha_{\tau} (i(t-\tau) - i(t-\tau-1)) - \sum_{\tau=0}^{\tau_{\max}} \gamma_{\tau} (i(t-\tau) - i(t-\tau-1)),$$

$$\frac{dr(t)}{dt} = \sum_{\tau=0}^{\tau_{\max}} \gamma_{\tau} (i(t-\tau) - i(t-\tau-1)),$$

де коефіцієнт α_{τ} характеризує ступінь заразності хворого, який хворіє вже τ днів, γ_{τ} описує долю одужалих (та вмерлих) після τ днів хвороби, τ_{\max} – максимальний час хвороби. За допомогою коефіцієнтів $\alpha_{\tau}, \gamma_{\tau}$ можна врахувати такі обставини хвороби, як запізнення прояву хвороби чи тривалість хвороби. Для розв'язання цієї системи рівнянь потрібно знати початкову історію епідемії: $i(0), i(1), \dots, i(\tau_{\max} + 1)$. та використання спеціальних методів розв'язання диференціальних рівнянь з запізненням аргументів.

Моделі спалахів епідемії. Епідемічні моделі містять багато невизначених та стохастичних параметрів. Наприклад, число інфікованих за одиницю часу є випадковою величиною, але в розглянутих детермінованих моделях це число замінюється середнім значенням. [10-14]. Випадковим є час хвороби. Випадковими є також спалахи епідемії в часі і просторі. Якщо епідемія трапилася одного разу у великій популяції (мегаполісі, країні), то спочатку вона повільно розвивається за експоненціальним законом і є час спинити її карантинними та іншими заходами. Але спалах може трапитись у порівняно невеликій популяції (гуртожитку, лайнері, містечку і т.д.), тоді ймовірно епідемія захоплює велику частину цієї невеликої популяції і тільки потім спадає. Для опису подібної ситуації потрібно застосовувати моделі у вигляді випадкових процесів і полів, де, наприклад, спалахи трапляються за пуассонівським законом у часі і є точково випадковими у просторі (наприклад, на графі) і розвиваються і спадають локально. Але спалах може розвинути і у пандемію. Можна припустити, що епідемії подібні спалахам лісових пожеж, але більш складні та динамічні за рахунок пересування індивідів.

Дані та ідентифікація моделі. Ідентифікації підлягають всі невідомі параметри

$\alpha, \beta, \gamma_i, \gamma_q, \varepsilon, \delta$ та деякі початкові дані моделі (наприклад, i_0). Статистичні данні для ідентифікації параметрів моделі можна узяти з сайту Національної служби здоров'я України [6]. В експериментах з **SIQR** моделлю вважалось, що $\gamma_i = \gamma_q = \gamma$, оцінювалися параметри α та γ , решта параметрів вважалися відомими (сценарними), зокрема припускалось, що коефіцієнт β є відомим ($2 \leq \beta \leq 50$) та $i_0 = q_0 / \beta$.

У теорії керування лінійними динамічними системами визначають і досліджують такі поняття, як керованість, спостережуваність та ідентифікуємість системи. Епідемічні моделі є нелінійними, тому проблема їх ідентифікації є більш складною. Типовий підхід до ідентифікації епідемічної моделі у вигляді системи диференційних рівнянь полягає у мінімізації по параметрам похибки прогнозу моделі, отриманого після розв'язання рівнянь, порівняно з даними спостережень. Але тут є багато варіантів, що прогнозувати та що спостерігати, як вимірювати похибку моделі, які параметри моделі отримувати безпосередньо з окремих вимірювань та іншого. Більш простим є так званий регресійних підхід, коли наступний стан системи вважається аналітичною функцією (алгебраїчною або у вигляді нейронної мережі) попередніх станів. Зауважимо також, що задача мінімізації похибки може бути багатоекстремальною задачею з багатьма близькими локальними екстремумами [15]. Ще одна потенційна складність задачі ідентифікації впливає з сингулярності проблеми відносно початкових даних – із практично нульових початкових даних виникає пандемія, що охоплює практично всю популяцію.

Числові експерименти з SIQR моделлю.

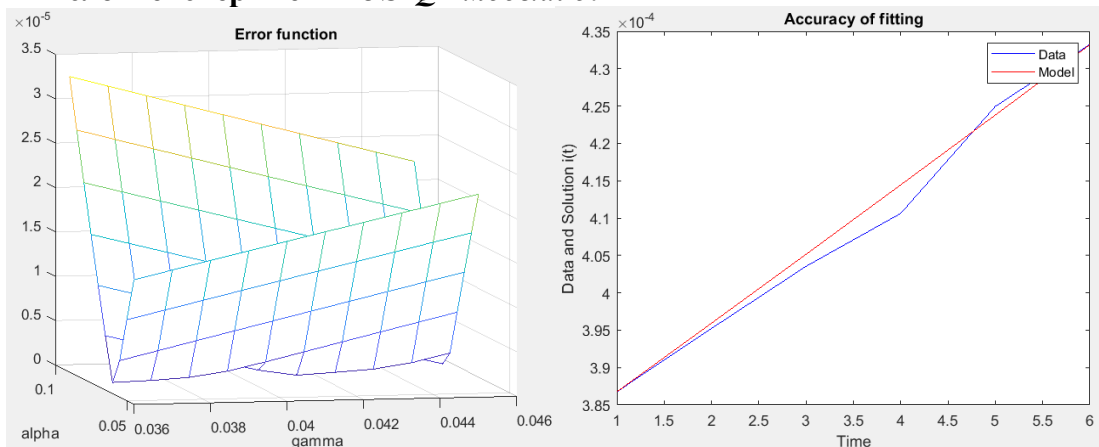


Рис.4 Ілюстрація налаштування моделі при $\beta = 0.1$ на даних про виявлених інфікованих в Україні за 8-14 травня

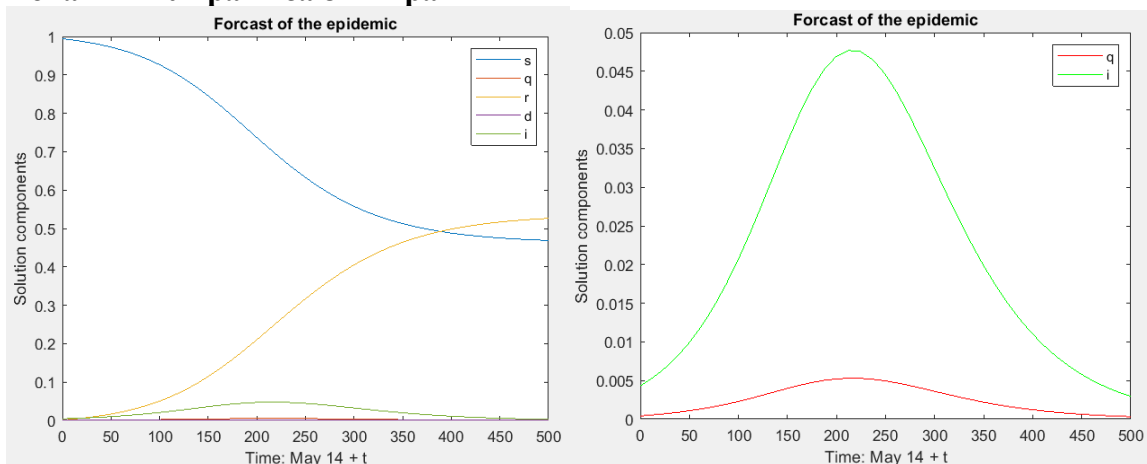


Рис. 5. Прогноз епідемії після 14 травня, якщо число безсимптомних хворих в $10 = 1/\beta$ більше виявлених хворих.

Стохастична модель. Врахування просторової структури розповсюдження COVID -19.

1.3 урахуванням успішного досвіду впровадження моделей класу SEIR в інших країнах Робоча Група при Президії НАН України розробила модель SEIR_U, [1].

Подальший розвиток цієї моделі пов'язан із врахуванням просторово-часової динаміки розповсюдження захворювання.

Як приклад стохастичної епідеміологічної моделі можна запропонувати модифікацію детермінованої моделі типу SEIR, наведеної в [1], з доданим стохастичним білим шумом (узагальненою похідною від стандартного вінерівського процесу) для врахування випадкових збурень. При дослідженні просторово-часової динаміки розповсюдження захворювання будемо враховувати наступні фактори:

1. Існують декілька вогнищ розповсюдження COVID-19. В Європі це гірськолижні курорти Австрії, Швейцарії та Франції, звідки вірус було рознесено по всім континентам відпочиваючими там туристами. Вважається, що таким осередком був і Мілан, в якому відбувся матч Ліги чемпіонів, на якому були присутні носії інфекції. В Україні такими вогнищами є західні регіони, у які повернулось багато заробітчани з Європи (насамперед Чернівецька область), у центрі -це великий транспортний хаб.

2. Оскільки у різних регіонах епідемічна ситуація складається по різному, будемо вважати, що для кожного регіону діє своя епідемічна модель, параметри якій дещо відрізняються від інших регіонів. Таким чином будемо розглядати 25 моделей, між якими будуть існувати перетоки інфікованих, оскільки повністю ізоляція регіонів не була проведена.

Дослідження такої масштабної моделі, яку складають $25 \times 14 = 350$ рівнянь, потребує потужних обчислювальних засобів, в якості яких доцільно використовувати СКІТ - суперкомп'ютерний обчислювальний комплекс Інституту кібернетики ім. В.М. Глушкова НАН України.

Запропоновані моделі дозволяють не тільки прогнозувати динаміку поширення епідемії, але й оцінювати для різних регіонів України найбільш ефективні з точки зору біологічної та економічної безпеки сценарії виходу з карантину, що залишається головним засобом стримування пандемії в світі.

2. 25-ти компартентальна модель .

Модифікуємо модель [1]. та введемо наступні припущення

1. Різні групи інфікованих мають різну вірулентність коронавірусу $R(i)$, тобто мають різні коефіцієнти репродукції (середня кількість заражень, що спричиняє одна хвора людина).
2. Для врахування випадкових збурень використовується білий шум (n-мірний стандартний вінерівський процес).
3. Розглядається декілька компартментів, між якими існує перетик інфікованих та здорового населення. Це дає можливість враховувати існування декількох джерел інфікування, оскільки у різних регіонах України епідемічний процес починався в різні строки, а також різну щільність населення в окремих регіонах України.
4. Модель враховує наявність запізнення в розвитку епідемічного процесу.

Для визначення швидкості зараження $V_{зар}$ використовуються не тільки припущення, що

$V_{\text{зар}} = (\text{здорові}) \times (\text{інфіковані})$, (тобто швидкість пропорційна кількості населення та кількості інфікованих), але і більш складні функції, які враховують деякі фактори управління, що гальмують процес зараження, по прикладу гальмування ферментативних реакцій інгібіторами.

Як приклад стохастичної епідеміологічної моделі можна запропонувати модифікацію детермінованої моделі типу SEIR, наведеної в [1], з доданим стохастичним білим шумом для врахування випадкових збурень. При дослідженні просторово-часової динаміки розповсюдження захворювання будемо враховувати наступні фактори:

Модифікована модель з урахуванням випадкових факторів має наступний вигляд:

$$\frac{d\vec{Y}}{dt} = F(\vec{Y}) + b_i \dot{w}_i,$$

де \vec{Y} - $14 \times 25 = 350$ - ті мірний вектор змінних, що визначають розповсюдження коронавірусу.

$y(1+14xI)$ – здорові, I-го регіону відповідно (IRB); (I,24)

$y(2+14xI)$ – безсимптомні хворі в інкубаційному періоді, коли вони ще не є заразними (IRB);

$y(3+14xI)$ – хворі, в інкубаційному періоді, які після появи симптомів потрапляють на карантин;

$y(4+14xI)$ – безсимптомні IRB;

$y(5+14xI)$ – інфіковані, які після появи симптомів потрапляють на карантин IRB;

$y(6+14xI)$ – легкі симптомні хворі, які одужують без госпіталізації IRB;

$y(7+14xI)$ – важкі симптомні хворі, які попадуть в лікарню та, зрештою, одужають IRB;

$y(8+14xI)$ – критичні симптомні хворі, які потраплять в лікарню та, зрештою, помруть IRB;

$y(9+14xI)$ – важкі симптомні хворі, які знаходяться в лікарні IRB;

$y(10+14xI)$ – критичні симптомні хворі, які знаходяться в лікарні IRB ;

$y(11+14xI)$ – пацієнти, які одужали після безсимптомного протікання хвороби IRB;

$y(12+14xI)$ – пацієнти, які одужали після легкого симптомного протікання хвороби IRB;

$y(13+14xI)$ – пацієнти, які одужали після важкого протікання хвороби і після госпіталізації IRB;

$y(14+14xI)$ – померлі IRB .

$w(t)$ - n-мірний стандартний винеровський процес (білий шум);

b_i ($i=1,350$) - параметри інтенсивності збурень.

a_0 - частина безсимптомних хворих, A_m, A_s, A_f - пропорції легких, важких та критичних хворих. Причому $A_m + A_s + A_f + a_0 = 1$

T_{inc} – латентний, або інкубаційний період

T_{inf} – активний період, в цей час хворий заразний

T_{car} – час до ізоляції після настання симптомів

T_m – середній період одужання легкого симптомного хворого;

T_r – середній період одужання важкого симптомного хворого в лікарні;

T_h – середній час перебування в лікарні;

T_d – середній період до настання смерті в лікарні.

N – загальна кількість населення;

$R(II)$ – коефіцієнт репродукції, середня кількість заражень, що спричиняє одна безсимптомно хвора людина ІКВ;

$R(2I)$ – коефіцієнт репродукції, що спричиняє одна хвора людина, яка після появи симптомів потрапляє на карантин ІКВ;

Розглянемо детерміновану однокомпонентну модель

Перший хворий на коронавірус був зареєстрований 10 березня. 100-й хворий був зареєстрований 25-го березня. Якщо брати за початок моделювання дату 10 березня, то 50-й день буде 29 квітня, 100-й день буде – 18 червня, 150-й -14 серпня. Зменшення параметру K_0 с 60 до 30 зміщує максимум захворюваності з 50-ого дня на 100-й.

Розглянемо, як впливає значення початкових даних на час досягнення максимуму захворюваності. Відомо, що більшість заробітчан верталися в Україну, не підлягаючи епідемічному контролю, тому кількість інфікованих на момент реєстрації першого хворого 10 березня викликає великі сумніви. На рис. 6 наведені результати модельних досліджень. Збільшення кількості інфікованих та хворих в інкубаційному періоді зміщує **пик** захворюваності на більш ранні строки

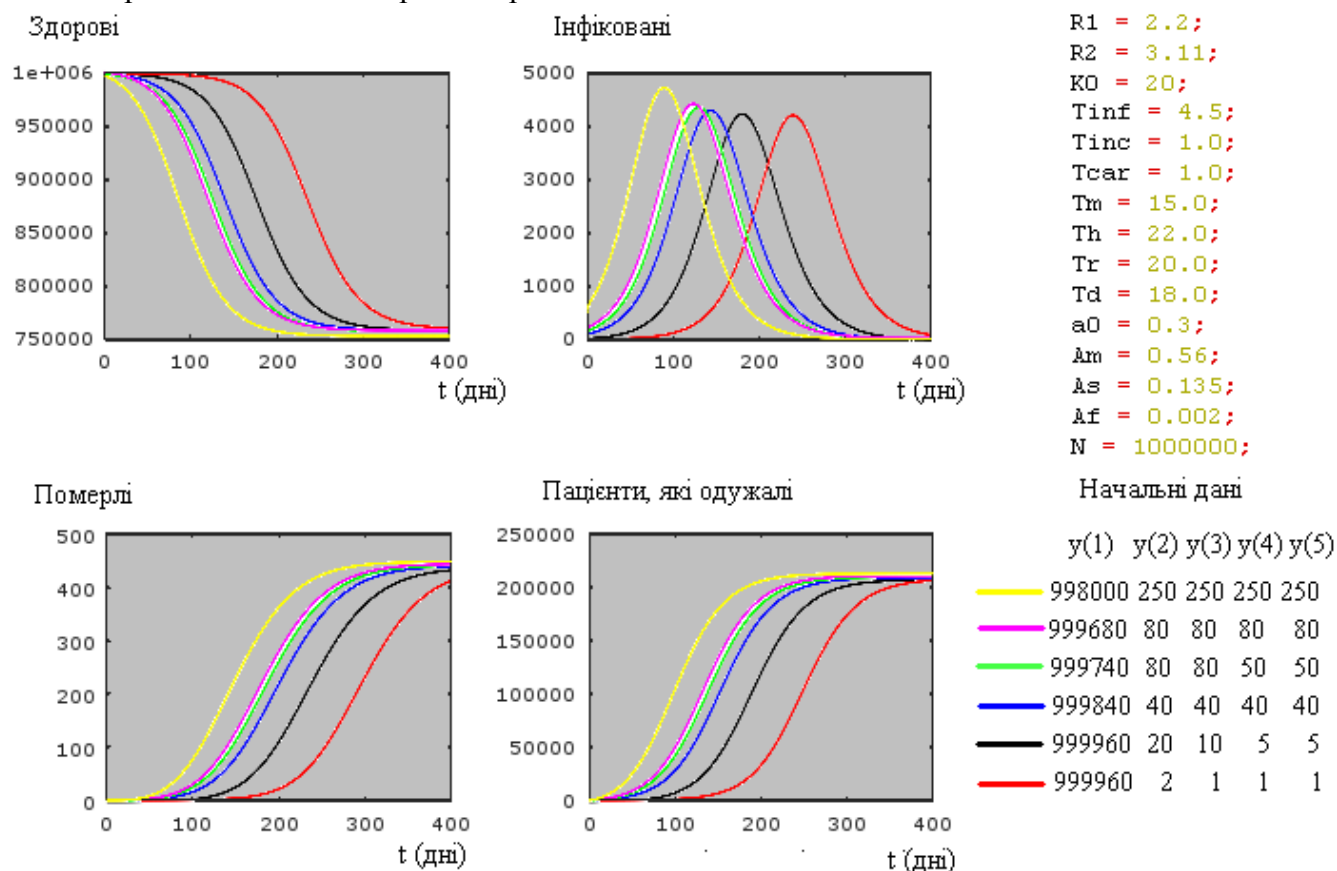


Рис. 6. Вплив початкових даних на динаміку захворюваності.

Розглянемо двокомпонентну модель

Якщо, функції V_1, V_2, V_3, V_4 , які визначають залежність швидкості інфікування від кількості інфікованих, то лінійне наближення функції V_i має наступний вигляд $V_1 = y(4), V_2 = y(5), V_3 = y(18), V_4 = y(19)$.

Із врахуванням гальмування процесу інфікування можуть застосовуватися більш складні функції, наприклад, для $V_1 = k_1 * y(4) / (k_2 + y(4))$, або $V_1 = k_1 * y(4) / (k_2 * (1 + I/k_3) + y(4))$, де I -фактор гальмування процесу зараження, який пов'язаний з профілактичними діями по

попередженню епідемії; k_i – параметри моделі (аналогі констант Михаеліса та інгибування у фк-ерментативної кінетиці),.

На рис 7 наведено результати моделювання для двохкомпонентної моделі для двох регіонів з населенням у 3 та 5 млн. мешканців. Виведено сумарні показники для двох регіонів. Розглянуто лінійне наближення, тобто $V1 = y(4)$, $V2 = y(5)$, $V3 = y(18)$, $V4 = y(19)$.

Як видно з рисунку, існує область параметрів моделі, коли виникає коливальний режим, тобто модель описує появу другої хвилі епідемії.

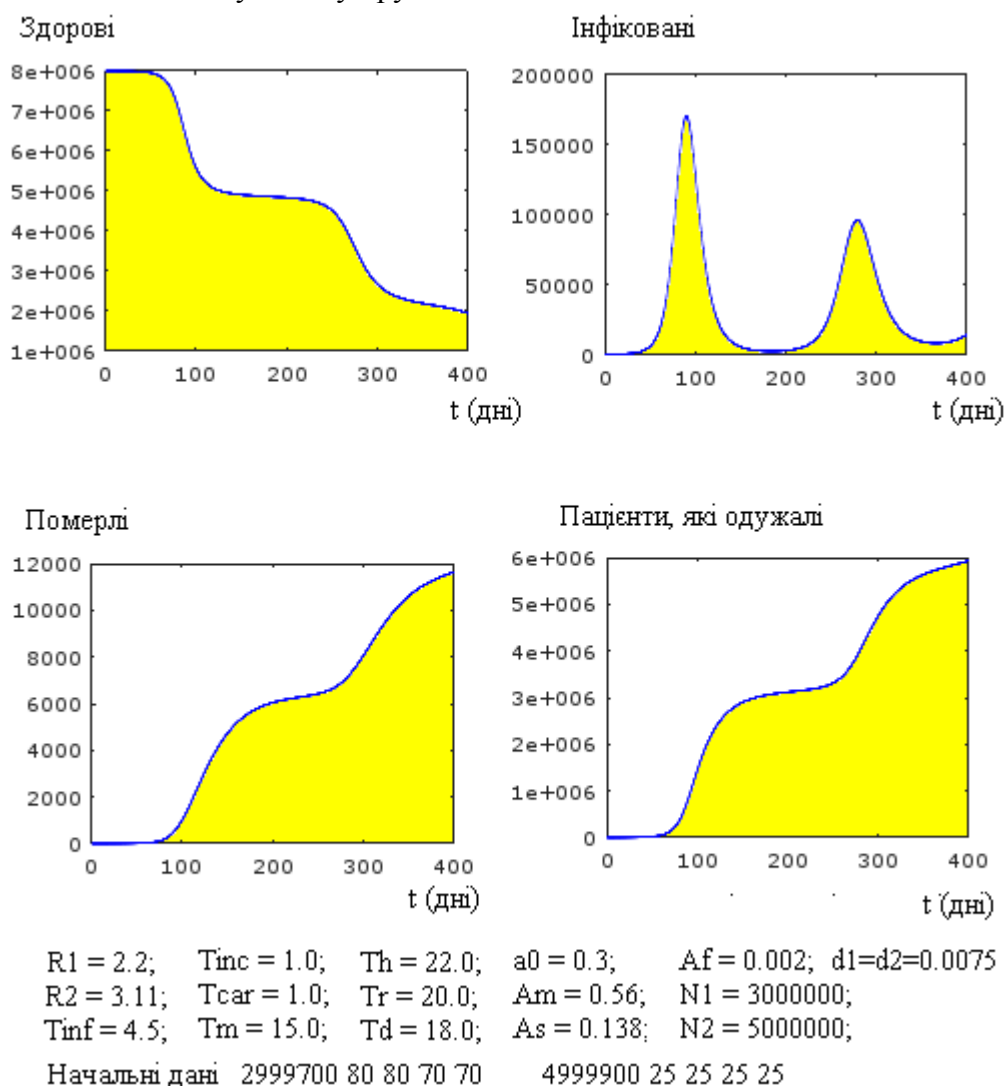


Рис 7.

Перелік посилань

1. Бровченко І. Розробка математичної моделі поширення епідемії COVID-19 в Україні // Світогляд, 2020, №2 (82). – С.2-14.
2. Кнопов П.С., Корхин А.С. Метод построение регрессии с переключениями в непрерывном времени с неизвестными точками переключения.// Кибернетика и системный анализ, 2020, № 1, С. 82 – 96.

3. Згуровський М.З. та ін. Проєкт «ФОРСАЙТ COVID-19» Світового центру даних «Геоінформатика і сталий розвиток». Режим доступу: <http://wdc.org.ua/uk/covid19-project>
4. Keeling M.J., Rohani P. (2008) *Modeling Infectious Diseases in Humans and Animals*, Princeton, NJ: Princeton University Press.
5. Chowell G., Hyman J.M., Bettencourt L.M.A., Castillo-Chavez C. (2009) *Mathematical models in population biology and epidemiology*, Dordrecht: Springer.2009.
6. Національна служба здоров'я України. Моніторинг ситуації із кількістю підозр та підтвердженими випадками захворювання на COVID-19 в Україні. Доступ: <https://nszu.gov.ua/covid/dashboard>
- 7.Братусь А.С., Новожилов А.С., Платонов А.И. Динамические системы и модели в биологии. - М.: ФИЗМАТЛИТ, 2010. - 436 с.
8. Brauer F., Castillo-Chavez C., Feng Z. *Mathematical Models in Epidemiology*. Springer, 2019. 619 p.
- 9.Gerardo Chowell · James M. Hyman Bettencourt L.M.A., Castillo-Chavez C. (Eds.) *Mathematical and Statistical Estimation Approaches in Epidemiology*. Springer, 2009. 10.
- 10.Katriel G. Stochastic discrete-time age-of-infection epidemic models. *International Journal of Biomathematics*. Vol. 6, No. 1 (2013). P.1-20.
- 11.Bettencourt L.M.A., Castillo-Chavez C. (Eds.) *Mathematical and Statistical Estimation Approaches in Epidemiology*. Springer, 2009. 363 p
- 12.Singh H., Dhar J. *Mathematical population dynamics and epidemiology in temporal and spatio-temporal domains*. Apple Academic Press, 2019.
- 13.Allen L.J.S. *Stochastic Population and Epidemic Models. Persistence and Extinction*. Springer, 2015.
- 14..Malchow H., Petrovskii S.V., Venturino E. *Spatiotemporal Patterns in Ecology and Epidemiology. Theory, Models, and Simulation*. Chapman & Hall/CRC, 2008. 443 p.
- 15.Ломов А.А. Вариационные методы идентификации линейных динамических систем и проблема локальных экстремумов, Управление большими системами. 2012, выпуск 39. С. 53-94.

Досвід і доробок авторів

- 1.Atoev K.L. *New Approaches to Epidemiological Risk Assessment Management "Supercourse: Epidemiology, the Internet and Global Health.- Pittsburgh: Disease Monitoring and Tele-communication WHO Collaborating Center,2002, <http://www.pitt.edu/~super1/lecture/lec4871/index.htm>*
2. Управление эпидемическим риском в Украине: новые подходы и стратегия развития /В.Ф.Москаленко, К.Л.Атоев, С.П. Бережнов и др. // *Врачебное дело.* —2002. —№ 3/4. — С. 3–17.
- 3.Оценка эскалации ВИЧ-инфекции и СПИДа в Украине. Информационная технология управления эпидемическим процессом и ранжирования рисков / К.Л.Атоев, Лапушенко, С.П. Бережнов и др. // *Врачебное дело.* —2004. —№ 5/6. — С. 3–17.
4. Sergienko I.V., Yanenko V.M., Atoev K.L. Optimal control of the immune response synchronizing the various regulatory compartments of the immune system. I. Mathematical analysis of the risk of pathological disorders in organism // *Cybernetics and Systems Analysis.* – 1995.-Vol.31.-N2.- P.225-239.
- 5.Sergienko I.V., Yanenko V.M., Atoev K.L. Optimal control of the immune response synchronizing the various regulatory compartments of the immune system. . II. Identification of model parameters and missing data recovery// *Cybernetics and Systems Analysis.* – 1997.- Vol.33.-N1.- P.131-144.

Робота виконана в Інституті кібернетики імені В.М. Глушкова НАН України.

Куратори роботи: Кнопов П.С., Норкін В.І.

Колектив виконавців: Атоєв К.Л., Горбачук В.М., Кирилюк В.С., Біла Г.Д.,
Самосьонок О.С., Богданов О.В.