УДК 519.86, 519.24, 004.942, 332.144

РЕТРОСПЕКТИВНОЕ МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ ГЛОБАЛЬНЫХ ПРОЦЕССОВ С ИСПОЛЬЗОВАНИЕМ ИНТЕГРАЛЬНОЙ МОДЕЛИ ПРИТОКА-ОТТОКА

Ю.Е. Балыкина, В.В. Захаров (Санкт-Петербург)

Введение

В фундаментальной общеметодологической статье [1] констатируется, что в последней четверти XX века наблюдалось изменение характера глобального демографического и экономического развития. Авторы обсуждают результаты исследований по математическому моделированию и прогнозированию мировой динамики, в том числе предложенную когнитивную схему взаимодействия различных сфер жизни на разных этапах исторического развития, базовые динамические уравнения такого взаимодействия. Важное значение в исследовании имели методы нелинейной динамики, экономической синергетики, нацеленные на описание неравновесных процессов, на анализ закономерностей разрушения старых и формирования новых социально-экономических структур, описанные в [2]. По результатам исследований коллектива Междисциплинарной научно-образовательной школы МГУ «Математические методы анализа сложных систем» (руководитель научного коллектива — акад. В. А. Садовничий) был подготовлен инициативный доклад для Римского клуба «Преодолевая пределы роста», приуроченный к 50-летию первого доклада Римскому клубу «Пределы роста» [3, 4]. В нем обобщается взгляд российских ученых на кардинальные изменения, происходящие в мире в последние десятилетия, излагаются методология моделирования и комплексный подход на основе системного взгляда на общество. Среди многочисленных научных публикаций, посвященных моделированию динамики глобальных демографических процессов, можно отметить [5-7]. Глобальные процессы распространения вирусов уже почти сто лет привлекают внимание большого количества исследователей [8-10]. Однако следует отметить, что предложенная Кермаком и МакКендриком общая нелинейная трехкамерная модель, параметры которой не являлись постоянными, в дальнейшем во многих исследованиях трансформировалась в модель SIR с постоянными коэффициентами, что значительно расширило возможности применения традиционных методов математического анализа, однако существенно ограничило возможности ее применения при построении прогнозов.

Важнейшим свойством глобальных процессов является стохастический характер основных измеряемых индикаторов этих процессов и их нестационарность, обусловленная влиянием на их динамику большого количества факторов среды, в которой они эволюционируют. Вследствие этого возможности традиционных методов математической статистики для прогнозирования динамики временных рядов значений этих измеряемых показателей довольно ограничены и не могут обеспечивать достаточно высокую точность построенных среднесрочных и долгосрочных прогнозов глобальных процессов. Следует отметить, что в научной литературе можно найти публикации, посвященные разработке методов прогнозирования нестационарных временных рядов, применение которых позволяет обеспечить приемлемую точность получаемых с их помощью прогнозов, например, при исследовании баз данных транспортных потоков и процессов потребления электроэнергии [11-15].

Одно из направлений научных исследований коллектива Центра аналитики динамических процессов и систем СПбГУ – разработка новых методов и математических моделей для прогнозирования глобальных процессов в здравоохранении и демографии.

Представленные в нашем докладе результаты этих исследований включают в себя описание авторской Интегральной модели притока-оттока и методов, использованных при прогнозировании нестационарных процессов распространения вирусов [16-18] и динамики численности населения [19, 20]. В докладе приводится математическая формализация общей интегральной модели притока-оттока, демонстрируются возможности ее использования для построения краткосрочных и долгосрочных прогнозов динамики этих процессов в мире и в отдельных странах и мегаполисах. Приводятся результаты вычислительных экспериментов по построению таких прогнозов с использованием разработанных алгоритмов. Приводится анализ точности ретроспективного прогнозирования.

Материалы и методы

Пусть динамическая система *X* в момент времени $t = t_0 \geq 0$ состоит из $X\left(t_0\right) \geq 0$ элементов определенного типа. Предположим, что в каждый момент времени $t = t_0 + 1, t_0 + 2, \ldots, t_0 + T \quad \left(1 \leq T < \infty\right)$ задано число $x_{inf}\left(t\right) > 0$ новых элементов того же типа, поступающих в систему, и $x_{of}\left(t\right) > 0$ число элементов, покидающих систему, $x_{inf}\left(t_0\right) = x_{of}\left(t_0\right) = 0$. Временные ряды $\left\{x_{inf}\left(t\right)\right\}$ и $\left\{x_{of}\left(t\right)\right\}$ будем соответственно называть временными рядами притока и оттока. Рассмотрим дискретную модель динамики размера (количества элементов) системы $X\left(t\right)$ при наличии притока $x_{inf}\left(t\right)$ и оттока $x_{of}\left(t\right)$. Очевидно, что состояние системы (размер системы) можно описать дискретным уравнением

$$X\left(t\right) = X\left(t-1\right) + x_{inf}\left(t\right) - x_{of}\left(t\right). \tag{1}$$

Если приток и отток представляют собой недетерминированные динамические процессы, то эволюция системы $X\left(t\right)$ есть случайный процесс, определяемый динамикой временных рядов притока и оттока, а система (1) в общем случае имеет стохастический характер. Если временные ряды не являются стационарными, то построение прогнозов будущей динамики системы с достаточно высокой точностью весьма затруднительно.

Рассмотрим суммы членов временных рядов притока и оттока до момента времени $t \leq T$.

Определение 1. Будем называть $X_{inf}\left(t\right)$ *интегральным объемом притока* в систему *X* с момента времени $t_0 + 1$ до момента времени $t \leq T$ сумму членов временного ряда притока

$$X_{inf}\left(t\right) = \sum_{\tau = t_0}^{t} x_{inf}\left(\tau\right).$$

Определение 2. Будем называть $X_{of}\left(t\right)$ *интегральным объемом оттока* из системы *X* с момента времени $t_0 + 1$ до момента времени $t \leq T$ сумму членов временного ряда оттока

$$X_{of}(t)=\sum_{\tau=t_0}^{t}x_{of}(\tau).$$

Следует заметить, что полученные таким образом временные ряды $X_{inf}(t)$ и $X_{of}(t)$ зависят от всех членов временных рядов $\{x_{inf}(\tau)\}$ и $\{x_{of}(\tau)\}$, $t_0+1\le\tau\le t$, соотвественно.

Уравнение (1) с учетом введенных ограничений можно преобразовать к следующему виду:

$$X(t)=X(t_0)+X_{inf}(t)-X_{of}(t).$$

Введем для любого $t>t_0+1$ и заданных временных рядов интегральных объемов притока и оттока следующие обозначения:

$$r_{inf}(t)=100\times\frac{X_{inf}(t)-X_{inf}(t-1)}{X_{inf}(t-1)}; \tag{2}$$

$$r_{of}(t)=100\times\frac{X_{of}(t)-X_{of}(t-1)}{X_{of}(t-1)}. \tag{3}$$

Величина $r_{inf}(t)$ называется процентным приростом интегрального объема притока $X_{inf}(t)$ в момент времени *t*, а величина $r_{of}(t)$ – процентным приростом интегрального объема оттока $X_{of}(t)$.

Рассмотрим следующую систему дискретных уравнений переменных $X(t)$, $X_{inf}(t)$ и $X_{of}(t)$ с недетерминированными параметрами $r_{inf}(t)$ и $r_{of}(t)$:

$$X_{inf}(t)=\left(1+\frac{r_{inf}(t)}{100}\right)X_{inf}(t-1); \tag{4}$$

$$X_{of}(t)=\left(1+\frac{r_{of}(t)}{100}\right)X_{of}(t-1); \tag{5}$$

$$X(t)=X(t_0)+X_{inf}(t)-X_{of}(t). \tag{6}$$

при $X_{inf}(t_0)=X_{of}(t_0)=0$, $t>t_0+1$.

Систему (4)-(6) будем называть *Интегральной моделью притока-оттока* системы *X* со стохастическими параметрами $r_{inf}(t)$ и $r_{of}(t)$.

Если подставить в уравнения (5) и (6) значения $r_{inf}(t)$ и $r_{of}(t)$, вычисленные по формулам (2) и (3) для заданных значений членов временных рядов $\{X_{inf}(t)\}$ и $\{X_{of}(t)\}$, мы, очевидно, получим равенства. С другой стороны, при использовании других значений этих параметров $\tilde{r}_{inf}(t)$ и $\tilde{r}_{of}(t)$ мы получим другие временные ряды интегральных объемов притока и оттока, а именно $\{\tilde{X}_{inf}(t)\}$ и $\{\tilde{X}_{of}(t)\}$.

Результаты

При исследовании процесса распространения вируса SARS-CoV-2 во время пандемии были использованы фактически статистические данные о количестве ежедневных случаев заболевания ($x_{inf}(t)$) и ежедневное суммарное число выздоровевших и умерших ($x_{of}(t)$). Текущее число болеющих людей в этом процессе равно размеру системы $X(t)$ в момент времени *t*.

Вместо традиционно используемой трехкамерной модели SIR (Susceptible-Infected-Removed) мы рассматривали модель, описываемую при $t > t_0$ системой дискретных уравнений следующего вида:

$$C(t) = \left(1 + \frac{r_{inf}(t)}{100}\right) C(t-1); \tag{8}$$

$$R(t) = \left(1 + \frac{r_{of}(t)}{100}\right) R(t-1); \tag{9}$$

$$I(t) = C(t) - R(t); \tag{10}$$

$$I(t_0) = C(t_0) = R(t_0) = 0.$$

В этой модели переменная $C(t)$ характеризует интегральный приток (общее количество) новых случаев заболевания (Confirmed Cases), а переменная $R(t)$ – интегральный отток суммарного количества выздоровевших и умерших (Removed), начиная с момента времени $t_0 + 1$ до момента времени *t* (включительно):

$$C(t) = \sum_{t=t_0+1}^{t} x_{inf}(t);$$

$$R(t) = \sum_{t=t_0+1}^{t} x_{of}(t).$$

Параметрами модели являются процентный прирост интегрального объема притока (общего количества выявленных случаев заболевания) $r_{inf}(t)$ и процентный прирост интегрального объема оттока (общего количества выздоровевших и умерших пациентов) $r_{of}(t)$.

Будем называть модель динамики эпидемического процесса (8)-(10) *интегральной моделью притока и оттока болеющих людей.*

В соответствии с информацией ООН [7] имеем временные ряды ежегодных значений численности населения $N(t)$, количества родившихся детей $B(t)$ и умерших людей $D(t)$, а также значений чистого миграционного потока $NM(t)$. Приток в систему народонаселения страны обеспечивает годовая рождаемость $B(t)$, отток из системы равен $DNM(t) = D(t) - NM(t)$.

Тогда система дискретных уравнений, описывающих динамику численности населения $N(t)$, будет иметь следующий вид:

$$IntB(t) = \left(1 + \frac{r_{inf}(t)}{100}\right) IntB(t-1); \tag{11}$$

$$IntDNM(t) = \left(1 + \frac{r_{of}(t)}{100}\right) IntDNM(t-1); \tag{12}$$

где

$$IntB(t) = \sum_{t=t_0+1}^{t} B(t);$$

$$IntDNM(t) = \sum_{t=t_0+1}^{t} DNM(t);$$

$$N(t) = N(t_0) + IntB(t) - IntDNM(t). \tag{13}$$

Будем называть дискретную модель (11)-(13) *интегральной моделью притока и оттока населения.*

Заключение

Разработанная Интегральная модель притока и оттока и апробированные методы прогнозирования будущей динамики временных рядов ее стохастических параметров с высокой точностью могут быть использованы:

– при моделировании и прогнозировании эпидемических процессов распространения новых инфекций в России и в других странах;

– при моделировании и прогнозировании численности населения Земли в целом и в различных странах на основе статистических данных ООН.

Построенные в процессе вычислительных экспериментов ретроспективные прогнозы (двухнедельные для пандемии и 10-летние для численности населения) демонстрируют возможность прогнозирования этих глобальных процессов с достаточно высокой точностью. Значения ошибок MAPE (средняя абсолютная ошибка в процентах) построенных ретроспективных прогнозов количества болеющих людей на пиках эпидемии, как правило, не превышает 3-5%, а при прогнозировании численности населения в большинстве случаев не превышала 1%.

Важным результатом применения разработанных методов для долгосрочного прогнозирования является прогноз динамики численности населения Земли до 2050 года: в 2039 году численность населения Земли выйдет на максимальный уровень 8,6 млрд (в 2023 году – 8,1 млрд) и затем начнет медленно уменьшаться до 8,3 млрд в 2050 году. Прогноз Отдела народонаселения ООН на 2050 год – 9,7 млрд человек.

Источник финансирования

Исследование выполнено за счет гранта Российского научного фонда № 23-21-10049 (https://rscf.ru/project/23-21-10049/) и гранта Санкт-Петербургского научного фонда.

Литература

1. **Садовничий В. А., Акаев А. А., Ильин И. В., Коротаев А. В., Малков С. Ю.** Моделирование и прогнозирование глобальной динамики в XXI веке // Вестник Московского университета. Сер. 27. Глобалистика и геополитика. 2022. № 1. С. 5-35.
2. **Малков С.Ю.** Нелинейная динамика нелинейного мира // Экономические стратегии. 2009. № 8. С. 44-51.
3. Преодолевая пределы роста. Основные положения доклада для Римского клуба: монография / под ред. В. А. Садовничего. М.: Изд-во Московского университета, 2023. – 99 с.
4. Reconsidering the Limits to Growth: A Report to the Russian Association of the Club of Rome / Ed. by V. Sadovnichy et al. N.Y.: Springer, 2023.
5. Global dynamics: approaches from complexity science / Ed. by A.G. Wilson. Hoboken, NJ: John Wiley & Sons, 2016.

6. **Капица С.П.** Феноменологическая теория роста населения Земли // УФН. 1996. Т. 166. № 1. 63-80 с.
7. United Nations, Department of Economic and Social Affairs, Population Division (2022). World Population Prospects 2022: Methodology of the United Nations population estimates and projections. UN DESA/POP/2022/TR/NO. 4.
8. **Kermack W.O., McKendrick A.G.** A Contribution to the Mathematical Theory of Epidemics // Proc R Soc (London) A. 1927. Vol. 115. Iss. 772. P. 700-721.
9. **Gecili E., Ziady A., Szczesniak R.D.** Forecasting COVID-19 confirmed cases, deaths and recoveries: revisiting established time series modeling through novel applications for the USA and Italy // PLoS One. 2021. Vol. 16. Iss. 1. e0244173
10. **Anderson R.M., May R.M.** Infectious diseases of humans: Dynamics and control. Oxford: Oxford University Press, 1991.
11. **Fryzlewicz P., Van Bellegem S., Sachs R.** Forecasting non-stationary time series by wavelet process modelling // Ann Inst Stat Math. 2003. Vol. 55. P. 737-764.
12. **Brockwell P. J., Davis R. A.** Time Series: Theory and Methods, 2nd ed., Springer, New York, 1991.
13. **Han H., Liu Z., Barrios M. et al.** Time series forecasting model for non-stationary series pattern extraction using deep learning and GARCH modeling // J Cloud Comp. 2024. Vol. 13. Art. №. 2.
14. **Bhatti U.A., Bazai S.U., Hussain S., Fakhar S., Ku C.S. et al.** Deep learning-based trees disease recognition and classification using hyperspectral data // Computers, Materials & Continua. 2023. Vol. 77. Iss. 1. P. 681-697.
15. **Guokun Lai, Wei-Cheng Chang, Yiming Yang, Hanxiao Liu.** Modeling long-and short-term temporal patterns with deep neural networks // Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, 2018.
16. **Захаров В. В., Балыкина Ю. Е.** Балансовая модель эпидемии COVID-19 на основе процентного прироста // Информатика и автоматизация. 2021. Т. 20. № 5. С. 1034-1064.
17. **Балыкина Ю. Е., Захаров В. В.** Интегральная модель притока и оттока и ее приложения // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2024. Т. 20. Вып. 2. С. 121-135.
18. **Захаров В. В., Балыкина Ю. Е.** Ретроспективный анализ и прогнозирование распространения вирусов в реальном времени: COVID-19 в Санкт-Петербурге и в Москве в 2020-2021 гг. // Вопросы вирусологии. 2024. Т. 69. № 6. С. 500–508.
19. **Захаров В. В.** Принцип динамического баланса демографического процесса и пределы роста населения Земли // Докл. РАН. Матем., информ., проц. упр. 2023. № 513. С. 108-114.
20. **Захаров В. В., Ндиайе С. М.** Прогнозирование численности населения и динамические игры против природы // Математическая теория игр и её приложения. 2024. Т. 16. Вып. 1. С. 17-37.