

Юнькова О.О., к. ф.-м. н., доцент,
професор кафедри математичного моделювання та статистики,
ДВНЗ «КНЕУ імені Вадима Гетьмана»
Кучер П.В. аспірант
кафедри математичного моделювання та статистики,
ДВНЗ «КНЕУ імені Вадима Гетьмана»

Yunkova O.O., PhD in Mathematics,
Professor of the Department of Mathematical Modelling and Statistics
SHEI KNEU named after V. Hetman
Kucher P.V.
Postgraduate student of the Department of Mathematical
Modeling and Statistics,
SHEI KNEU named after V. Hetman

МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ ТОВАРНИХ РИНКІВ НА ОСНОВІ МЕТОДУ ДЕКОМПОЗИЦІЇ

MODELING AND FORECASTING OF COMMODITY MARKETS BASED ON THE DECOMPOSITION METHOD

Анотація. Прогнози становлять значну інформаційну базу для процесу прийняття управлінських рішень. У даній роботі розглянуто процес дослідження та економетричного моделювання продажів різних товарів: два набори даних, що представляють різні снеки, лікарський препарат від грипу та простуди, знеболююче (нестероїдний протизапальний препарат), гель, що застосовуються проти болю в ногах, а також вітамінний комплекс. Часові ряди щодо продажів снеків і гелю представлено місячними даними, а всі інші набори — тижневими даними. Графічний аналіз даних, а також інструментарій перших різниць та автокореляційна функція показали наявність трендової та сезонної складової в даних часових рядах. Метод декомпозиції, який дає змогу виявляти як довгострокові тенденції розвитку процесів, так і виділити сезонні зміни, застосовано для побудови моделей часових рядів. Для реалізації цього методу вибрано середовище R, в якому передбачено широкий спектр послуг, інструментів і пакетів для аналізу даних. У даній роботі проведено аналіз даних і виділено комбінацію трендової та сезонної складової даних часових рядів. Декомпозицію ряду реалізовано методом LOESS (Seasonal Transformation using LOESS — “STL”). Алгоритм STL реалізує основні засади методу декомпозиції, а саме згладжує часові ряди за допомогою LOESS у двох циклах. У внутрішньому циклі виконується згладжування сезонних коливань і згладжування тренду, а зовнішній цикл мінімізує вплив викидів. При цьому у внутрішньому циклі спершу обчислюється сезонна компонента, а після її видалення обчислюється трендова складова. Залишки обчислюються шляхом віднімання сезонного та трендового компонентів із часових рядів. Виділені залишки перевіряються на відповідність нормальному закону розподілу. Всі зазначені кроки автоматично виконуються за допомогою мови R. Результати економетричного моделювання дозволили виявити головні тенденції, а також здійснити прогноз у короткостроковій перспективі. У результаті дослідження розроблено адекватну модель прогнозування для часових рядів, які можуть виникати на різних товарних ринках. Верифікація моделі показала достатню то-

чність і достовірність виконаних на її основі прогнозів, тому її можна застосовувати у подальших дослідженнях для аналізу продажів продуктів різних категорій і прогнозування їхніх майбутніх значень.

Ключові слова: моделювання, часовий ряд, метод декомпозиції часового ряду, прогнозування, товарні ринки, тренд, сезонність

Abstract. Forecasts are a significant information base for the management decision-making process. This study examined the process of research and econometric modeling of sales of different products: two sets of data representing different snacks, a drug for flu and colds, painkillers (nonsteroidal anti-inflammatory drug), a gel used for pain in legs, as well as a vitamin complex. The periodicity of the time series of snacks and gel was on a monthly basis, all other sets of data had a weekly seasonality. Data analysis was performed and a combination of trend and seasonal components of time series data were identified. Decomposition was performed by the LOESS method (Seasonal Transformation using LOESS — “STL”).

The algorithm of STL will realize basic principles of method of decoupling, namely smooths out sentinel rows by means of LOESS in two cycles. Smoothing of seasonal vibrations and smoothing of trend is executed in the inner loop, and an outer loop minimizes influence of extrass. Thus in the inner loop a seasonal component is calculated at first, and after her moving away the constituent of trend is calculated. Bits and pieces are calculated by deduction of seasonal and trend components from sentinel rows. The distinguished bits and pieces are checked for accordance to the normal law of distribution. All these steps are automatically performed using the R language. In addition, the analysis of the allocated balances for compliance with the normal distribution law was performed.

Prognostication of future values of indexes will carry out on the basis of the built model. For quality of prognosis control part of data (last periods) is excluded from an analysis, to have the opportunity to check adequacy of the got results when compared to the noticed values. The results of econometric modeling allowed to identify the main trends, as well as to make a forecast in the short term.

As a result of the study using the described procedure developed an adequate forecasting model for time series of different product markets, the verification of which showed sufficient accuracy and reliability of forecasts based on it, so it can be used in further research or analysis of products of different categories

Keywords: modeling, time series, time series decomposition method, forecasting, commodity markets, trend, seasonality

Постановка проблеми. Сучасний стан економіки, який характеризується динамізмом, високим ступенем невизначеності навколишнього середовища, суттєвими технологічними перетвореннями, потребує кількісного оцінювання розвитку соціально-економічних процесів. Сьогодні кількість інформації зростає із неймовірною швидкістю, тому застосування нових програмних продуктів для аналізу та узагальнення цих даних є нагальною потребою. Здатність передбачати майбутні перспективи підвищує успішність економічної діяльності будь-якого підприємства. У зв'язку з цим прогнозування продажів того чи того продукту завжди залишається актуальною проблемою. Використання різних економіко-математичних методів і моделей, а також сучасних засобів обробки інформації дозволяють з високою надійніс-

тю і точністю прогнозувати економічні та фінансові показники підприємства, оптимізувати систему управління тощо.

Аналіз останніх досліджень і публікацій. Значну увагу різним аспектам моделювання економічних процесів та явищ, прогнозуванню їхнього розвитку присвятили у своїх роботах П. Черненко [2], О. Мартинюк [2], А. Е. Шуліков [1], М. А. Голованова [1], З. Каїра [3], Р. Хіндмана [4], Дж. Афанасопулоса [4], В. Кочкода-на [5], С. Мастицького [11] та інших.

Метою дослідження є аналіз різних наборів часових рядів, які відображають обсяги продажів продуктів з різних сегментів товарних ринків, моделювання цих рядів і побудова прогнозів із застосуванням декомпозиційного методу і середовища R.

Аналіз і статистичний опис динаміки якого-небудь істотного коливного показника (ціни) починається з виявлення форми його тренда й статистичної оцінки його параметрів. Відповідно до визначення тренда, форма його об'єктивна й відбиває закономірності розвитку досліджуваного процесу [2]. Один з методів аналізу часових рядів є виокремлення складових часового ряду (тренду і періодичних коливань), які представляють його певну регулярну компоненту. Подібна процедура ідентифікації називається декомпозицією. Кожна компонента ідентифікується окремо. Потім внески кожної компоненти комбінуються з метою отримання прогнозів майбутніх значень часових рядів. Методи декомпозиції використовуються як для короткочасних, так і для довгострокових прогнозів. З їхньою допомогою також можна відображати закономірність, що лежить в основі тренду, або коригувати значення ряду, додаючи іншу компоненту.

Наприклад, при розгляді продажів різних товарів (чи то лікарських препаратів, чи будь-яких інших продуктів), дуже важливим є розуміння того, яка частка змінювання продажів, спричинена пенними тенденціями на ринку, а яку можна пояснити сезонним попитом на товар. Метод декомпозиції дає змогу виявляти чотири складові часового ряду:

- тренд (T);
- циклічна компонента (C);
- сезонна компонента (S);
- випадкова змінна (I) [5, с. 50-52].

Тренд (T) показує зміни у довгостроковому періоді, які спостерігаються в часовому ряді за умови, що інші компоненти виключені (циклічна, сезонна, випадкова). У більшості випадків передбачається, що тренд можна представити у вигляді прямої лінії (або експоненціальної, чи логарифмічної кривої) [4, с. 52-55].

Циклічна (C) компонента присутня не завжди. Вона відображає спади і («хвилі») в часовому ряді. Циклічні коливання, як правило, проявляються протягом досить тривалого часу — приблизно від трьох до п'яти років. Для деяких товарів (наприклад, для хліба) відзначаються незначні циклічні коливання, водночас продажі інших (наприклад, будівництво житла) зазнають істотних змін.

Сезонний фактор (S) часто спостерігається у кварталних, місячних або тижневих даних. Під сезонними варіаціями розуміють змінювання рівнів показника з більш-менш стабільною структурою в межах одного року.

Випадкова (I) компонента відображає коливання рівнів ряду, які спостерігаються після виключення впливу тренду, циклічного та сезонного факторів і спричинені впливом випадкових чинників [6, с. 32-33].

Функцію вихідного ряду можна представити в адитивній (1) або мультиплікативній формі (2):

$$y(t) = T(t) + C(t) + S(t) + I(t), \quad (1)$$

$$y(t) = T(t) * C(t) * S(t) * I(t). \quad (2)$$

Вибір однієї з двох моделей проводиться на основі аналізу структури сезонних коливань [7, с. 111-113]. Модель адитивних компонент застосовується в тому разі, коли аналізований часовий ряд має приблизно однакові зміни протягом всієї тривалості спостережень, тобто сезонні коливання є сталими і значення часового ряду при цьому мають певну тенденцію до зростання чи спадання. Мультиплікативна модель ефективніша, коли і коливання і часовий ряд збільшуються або ж навпаки зменшуються з часом [8, с. 21].

Моделювання та прогнозування часових рядів на основі декомпозиції

Для здійснення прогнозу нам потрібно провести передусім декомпозицію часового ряду. У нас є доступні 6 наборів даних: 2 ряди продажів снєків («Snack1», «Snack2»), вітамінний комплекс («Vitamin»), гель від від болю в ногах («Gel»), засіб від грипу («Flu») і жарознижуюче («Head»). «Snack1», «Snack2» і «Gel» — представлені помісячно за кілька років, а от «Vitamin», «Flu» та «Head» — потижнево.

Побудова прогнозової адитивної або мультиплікативної тренд-сезонної моделі передбачає такі кроки:

- згладжування даних методом ковзної середньої;
- визначення різниць між початковими даними і центрованими середніми, які характеризують сезонний чинник, а також середніх значень ряду;
- вилучення сезонної компоненти з початкового часового ряду;
- аналітичне вирівнювання десезоналізованого ряду для оцінювання трендової складової;
- обчислення невинуваткової складової ряду (тренд і сезонність);
- обчислення випадкової складової ряду, аналіз залишків і перевірка адекватності моделі;
- прогнозування за моделлю та перевірка якості прогнозу [9, с. 1-5].

Усі зазначені кроки автоматично виконуються за допомогою мови *R*. Для проведення прогнозу використано адитивні підходи, а саме методику *STL*. Її можна застосувати до будь-якого набору даних, але значущі результати будуть отримані тільки в тому разі, якщо в даних є тимчасова закономірність, що повторюється (наприклад, препарати від кашлю є популярніші у холодну пору, а от солодка вода краще продається влітку, або є онлайн-шопінг зростає листопаді-грудні, коли настає Чорна П'ятниця чи починається передріздвяна гонитва за подарунками). Закономірність відображається у результаті функції *STL* як сезонна компонента.

Алгоритм *STL* згладжує часові ряди за допомогою *LOESS* у двох циклах; внутрішній цикл виконує ітерацію між сезонним згладжуванням і згладжуванням тренду, а зовнішній цикл мінімізує вплив викидів. У ході внутрішнього циклу сезонна компонента обчислюється перша та видаляється для обчислення компонента тренду. Залишок обчислюється шляхом віднімання сезонного та трендового компонентів із часових рядів.

Результати здійсненої декомпозиції часових рядів представлено на рис. 1.

Проведена декомпозиція підтверджує, що всі ряди мають тренди: — зростаючі, а третій — нелінійний і всі ряди мають яскраво виражену сезонність [10, с. 89].

Після розкладання часових рядів на складові необхідно перевірити залишки на відповідність закону нормального розподілу. Гістограми частот залишків часових рядів представлені на рис. 2.

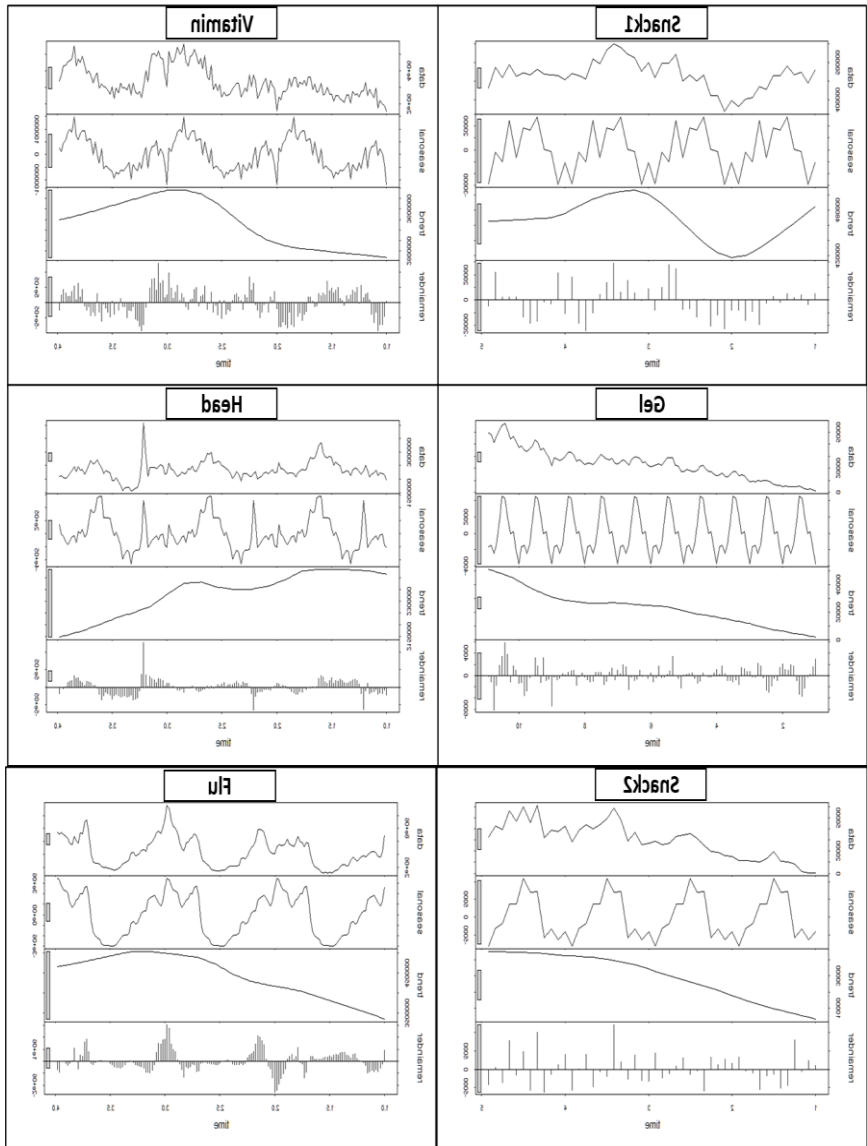


Рис. 1. Результати декомпозиції обсягів продажу часових рядів

Джерело: побудовано авторами на основі даних підприємства

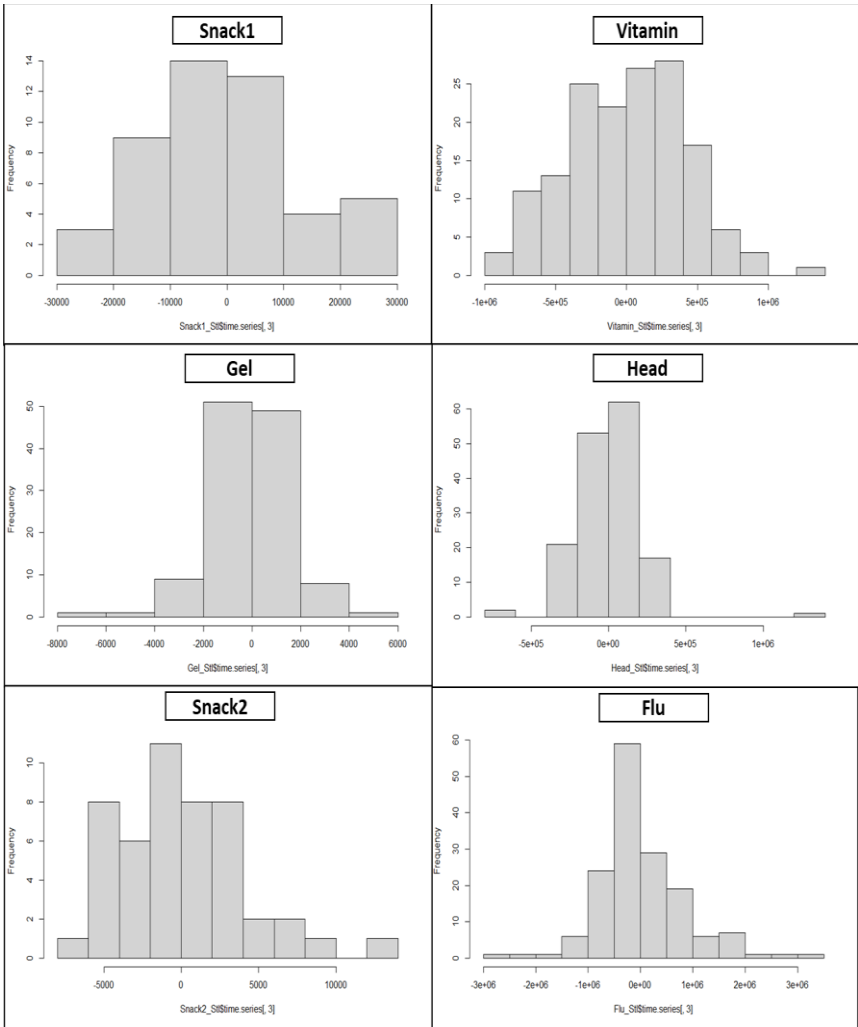


Рис. 2. Розподіл залишків часових рядів

Джерело: побудовано авторами на основі даних підприємства

Залишки моделей, згідно діаграм, представлених на рис. 2, мають розподіл, близький до нормального, отже моделі є адекватними і можуть бути застосовані для передбачення обсягів майбутніх продажів означених лікарських засобів.

Прогнозування майбутніх значень показників будемо здійснювати на підставі побудованої моделі. Для перевірки якості прогнозу частину даних (останні періоди) виключено з аналізу, щоб мати змогу перевірити адекватність отриманих результатів при порівнянні зі спостереженими значеннями. Отримані результати відображено на рис. 3.

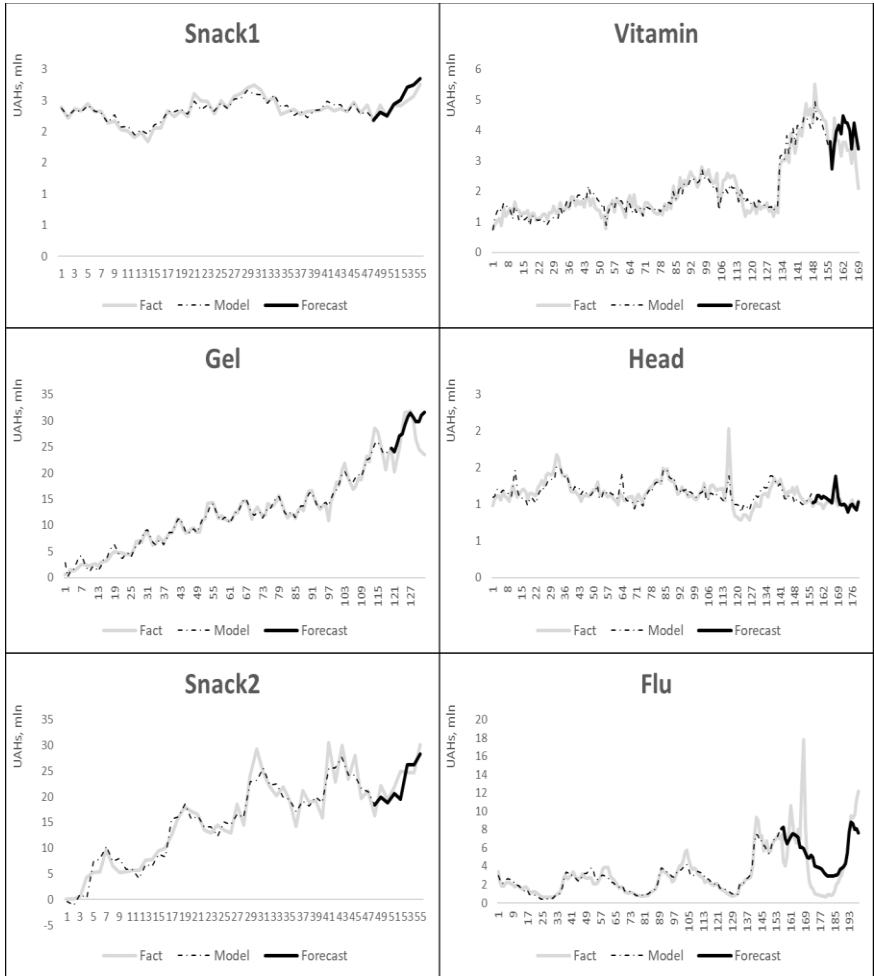


Рис. 3. Результати прогнозування обсягів продажу часових рядів

Джерело: побудовано авторами на основі даних підприємства

Як засвідчують наведені графіки, розбіжності між розрахованими та спостереженими даними для більшості показників є незначними. Але варто зазначити, що для препарату від грипу/простуди не вдалося точно відобразити стрибок у продажах, спричинений форс-мажорною ситуацією, що пов'язана з пандемією коронавірусу у 2020 році. Подібна ситуація спостерігається і для даних щодо продажів гелю від болю в ногах: модель не змогла досконало відтворити суттєве зростання рівнів показника у досліджуваній період. Інші ж чотири набори даних спрогнозовані досить точно, що засвідчує високі прогностичні властивості методу декомпозиції при передбаченні продажів різних продуктів. Однак, зважаючи на те, що у двох із шести випадків метод не спрацював ідеально, варто паралельно застосовувати декілька методів для підвищення точності прогнозу. Оскільки окрім тренд-сезонної складової є багато інших чинників, які можуть впливати на рівень продажів, наприклад, попит на продукцію, рівень дистрибуції, реклама тощо, моделі прогнозування слід комбінувати при здійсненні прогнозу (особливо для довгострокових передбачень).

Висновки з проведеного дослідження. У статті розглянуто особливості моделювання і прогнозування часових рядів на основі використання методу декомпозиції у статистичному середовищі R. Було виокремлено складові шести різних часових рядів (тренду і періодичних коливань) і проведено дослідження. Для 4 рядів метод спрацював доволі точно і дав хороші результати. На підставі проведеної роботи можна стверджувати, що використання такого підходу дає змогу зменшувати вплив невизначеності при моделюванні динамічних процесів і досить точно прогнозувати економічну ситуацію на довільному ринку. Але важливим є розуміння того, що метод не є ідеальним і для підвищення точності результату його слід комбінувати з іншими способами передбаченнями. Запропоновані в роботі підходи можуть використовуватись аналітиками рекламних компаній, корпорацій для аналізу ринку, прогнозування ситуації в різних сегментах тощо.

Бібліографічні посилання

1. Статистичне дослідження часових рядів цін на B2B-ринку / А. Е. Шуліков, М. А. Голованова. *Економіка та управління підприємствами машинобудівної галузі*. 2011. № 2. С. 125–135. Режим доступу: http://nbuv.gov.ua/UJRN/eupmg_2011_2_13

2. Декомпозиція добового графіка електричного навантаження енергосистеми та моделювання його складових під час короткострокового прогнозування / П. О. Черненко, О. В. Мартинюк. *Вісник Вінницького політехнічного інституту*. 2017. № 6. С. 86–94. Режим доступу: http://nbuv.gov.ua/UJRN/vvpi_2017_6_15

3. Каїра З. С. Прогнозування сезонного попиту в логістиці та маркетингу / З. С. Каїра. *Вісник Національного університету "Львівська політехніка"*. 2006. № 554: Проблеми економіки та управління. С. 134–138.

4. Forecasting: Principles and Practice. / Hyndman, Robin John; Athanasopoulos, George. 2nd ed. OTexts, 2018. 384 p.

5. Кочкодан В. Б. Теоретичні аспекти прогнозування структурних характеристик капіталу підприємства / В. Б. Кочкодан. *Інноваційна економіка*. 2014. № 2. С. 48-53.

6. Прогнозування та аналіз часових рядів. Методичні вказівки до практичних занять та самостійної роботи студентів спеціальності 051 «Економіка» освітня програма «Економічна кібернетика», «Економічна аналітика» / Укл.: Юрченко М. Є. Чернігів: ЧНТУ, 2018. 88 с.

7. Тебуева Ф.Б., Перепелица В.А., Кабиняков М.Ю. Декомпозиция и прогнозирование временных рядов с долговременными корреляциями. *Известия Южного федерального университета. Технические науки*. №1. 2013. С. 111 –120.

8. Кизбикенов, Кажимурат Оспанович Прогнозирование и временные ряды: учебное пособие / К. О. Кизбикенов; Алтайский гос. пед. ун-т. Барнаул: АлтГПУ, 2017. 113 с.

9. Шитиков В. Прогнозирование бинарных временных рядов [Электронный ресурс] / Владимир Шитиков. *Электронный журнал "Jahrbuch für EcoAnalytic und EcoPatologic"*. 2019. Режим доступу до ресурсу: http://www.ievbras.ru/ecostat/Kiril/R/Blog/13_BTS.pdf.

10. Венэблз Д. М. Заметки по R: среда программирования для анализа данных и графики (Перевод с английского) / Д. М. Венэблз, У. Н. Смит. Москва, 2013. 109 с.

11. Мاستицкий С. Э. (2020) Анализ временных рядов с помощью R. — Электронная книга, адрес доступа: <https://ranalytics.github.io/tsa-with-r>.

Статтю подано до редакції 09.11.2021