

ПРОГНОЗУВАННЯ ЧАСОВОГО РЯДУ ЗАЛИШКІВ ЗНАЧЕНЬ ІНДЕКСУ ГЕОМАГНІТНОЇ АКТИВНОСТІ З ВИКОРИСТАННЯМ ЛАНЦЮГІВ МАРКОВА

Запропоновано метод прогнозування часових рядів геомагнітної активності на основі їх попередніх відліків з використанням моделей Маркова. Подано приклад застосування моделі та розв'язання задачі прогнозування на конкретних даних.

Ключові слова: *марківська модель, ланцюг Маркова, геомагнітна активність, Dst-індекс, прогнозування космічної погоди.*

Предложен метод прогнозирования часовых рядов геомагнитной активности на основании их предыдущих замеров с использованием моделей Маркова. Приведен пример использования модели и решения задачи на конкретных данных.

Ключевые слова: *марковская модель, цепь Маркова, геомагнитная активность, Dst-индекс, прогнозирование космической погоды.*

The paper proposes forecasting method of geomagnetic activity index time series based on their previous measurements using Markov model. An example of model using and task solving was given.

Keywords: *Markov model, Markov chain, geomagnetic activity, Dst-index, forecasting of space weather.*

Постановка проблеми. Процеси формування небезпечних явищ погоди, як свідчать сучасні дослідження, супроводжуються фізичними ефектами, що локально проявляються у земних сферах, які прийнято називати передвісниками або предикторами подій. За характеристиками цих передвісників можна судити про величину і час майбутньої ймовірної події. Відтак, здійснюючи глобальний моніторинг предикторів на поверхні і в атмосфері Землі з використанням космічних засобів дистанційного зондування Землі спільно з мережею спеціалізованих наземних станцій, можливе наукове обґрунтування та прогнозування небезпечних явищ космічної погоди.

Окрім впливу на електроніку та роботу супутників, космічна погода здійснює вплив на розповсюдження радіохвиль та сигнали сучасних космічних радіосистем – навігаційних систем (GPS), систем супутникового зв'язку та ін., призводячи до порушення супутникової навігації, втрати космічного зв'язку та космічних даних. Розповсюджуються через збурену іоносферу та атмосферу, ці сигнали піддаються розсіюванню та послабленню на іоносферних неоднорідностях, що виникають у періоди сильних геомагнітних збурень та бур [1].

Супутниковий зв'язок, що забезпечує радіо- та телемовлення, Інтернет, передачу інформації та даних, піддається впливу космічної погоди через стан середовища, в якому він функціонує. Цей вплив полягає у дії космічної погоди безпосередньо на самі супутники (радіація, аномальне гальмування), у дії факторів космічної погоди на наземний сегмент (відключення електроенергії у періоди магнітних бур) та на сигнали супутникового зв'язку під час їх розповсюдження через верхню та нижню атмосферу Землі.

Наведені факти впливу сонячно-земних зв'язків на космічну техніку та технології, на космічну діяльність у цілому, а також зареєстровані екстремальні прояви космічної погоди свідчать про те, що розвиток технічних систем сягнув такого рівня, коли вплив на них з боку космічної погоди здатен призвести до серйозних катастроф та збитків. Це ставить задачу створення засобів моніторингу космічної погоди, розробки методів її прогнозування та заходів пом'якшення її впливу, а також проведення поглиблених фундаментальних досліджень сонячної активності та її дії на Землю.

Враховуючи складність внутрішньої природи явищ магнітосфери, що відображають послідовність взаємопов'язаних фізичних процесів, не завжди можна побудувати адекватну математичну модель, що буде повною мірою описувати основні динамічні властивості. Саме тому доцільно використовувати нелінійні математичні моделі, структура яких від початку невідома, а параметри залежать від експериментальних даних спостережень.

Індекс *Dst* є ключовим параметром, що характеризує турбулентний стан магнітосфери під час магнітної бурі [2]. Усереднена за декількома станціями варіація отримала назву *storm-time variation*. Поле *Dst* складається з декількох джерел: поля струмів, що течуть уздовж поверхні магнітосфери Землі *DCF*, поля кільцевого струму *DR* та поля струмів, що протікають в іоносфері *DP*. Таким чином, цей індекс визначається трьома складовими: $Dst = DCF + DR + DP$; границі

вимірювання: $-100 < Dst \leq 50$ нТл для помірних магнітних бур, $Dst \leq -100$ нТл для сильних магнітних бур.

Дані для статистичної обробки про заміри Dst -індексу були отримані з бази даних Space Physics Data Facility (SPDF), проекту Геліосферного наукового підрозділу (HSD) NASA, що містить у собі інформацію про основні характеристики стану навколосезонного магнітного поля, сонячного вітру і параметри плазми з декількох космічних апаратів [3].

Аналіз публікацій та постановка задачі. Відомі моделі прогнозування, що базуються на використанні методів штучного інтелекту, не дозволяють явно розрахувати очікувану величину параметра, а лише вказують на його приналежність до якої-небудь заздалегідь визначеної області. Ще одним недоліком подібних методів є недостатній рівень надійності математичного апарату. В сучасних умовах перспективними є моделі, що базуються на теорії марківських процесів [4]. Марківські моделі мають ряд переваг у порівнянні з іншими стохастичними моделями [5]: їх легко будувати на експериментальних даних; вони не потребують розуміння внутрішніх механізмів динаміки змін системи; результати моделювання можуть бути легко представлені графічно; їх застосування не потребує великих витрат машинного часу. До недоліків моделей марківського типу можна віднести умову стаціонарності системи, необхідність отримання достатньо великого за обсягом ряду спостережень для визначення достовірних ймовірностей переходу і відсутність залежності від функціональних механізмів системи.

Ланцюги Маркова ефективно використовуються для моделювання процесів у метеорології: прогнозування погоди [6], оцінки забруднення атмосфери, зміни ландшафту та ін. Застосуванню ланцюгів Маркова при моделюванні метеорологічних процесів присвячені роботи австралійських дослідників [7–9], які теоретично та практично обґрунтували можливість застосування методик до короткострокового прогнозу погоди. Порівняння шести способів оперативного прогнозу показало, що поєднання прогнозів з використанням ланцюга Маркова другого порядку та традиційних чисельних прогнозів погоди значно підвищує точність прогнозування у порівнянні з іншими методами та може бути рекомендовано до використання в оперативній роботі.

Ланцюги Маркова показують хороші результати і при прогнозуванні опадів [10; 11]. У роботах розглянуто три модифікації

стохастичного генератора добової погоди для покращення відтворення високо- та низькочастотної мінливості у рядах метеоелементів.

У роботі [12] запропоновано модель прогнозування циклічних компонент часового ряду значень індексу геомагнітної активності Dst за допомогою сингулярного спектрального аналізу. Отримані результати дозволяють говорити про хорошу якість прогнозу циклічної компоненти, що повільно змінюється.

Метою даної роботи є дослідження ряду залишків та оцінка якості його прогнозування на основі марківської моделі.

Виклад основного матеріалу. Нехай задана послідовність (часовий ряд) $P = \{p_i\}_{i=\overline{1,N}}$, де N – обсяг даних Dst -індексу за визначений часовий проміжок. Подамо вихідний часовий ряд у вигляді:

$$p(t) = p_T(t) + p_3(t),$$

де $p_T(t)$ – циклічна складова; $p_3(t)$ – ряд залишків. Введемо до розгляду систему S , яка відображає поведінку індексу геомагнітної активності за певний період. Система може знаходитися в одному з чотирьох станів S_1 , S_2 , S_3 та S_4 . Стан S_1 показує, що значення індексу знаходиться нижче першого граничного значення, тобто у проміжку $(Dst_{\min}; \alpha_1]$, стан S_2 відповідає проміжку між першим та другим граничним значенням $(\alpha_1; \alpha_2]$, S_3 – між другим та третім $(\alpha_2; \alpha_3]$, а стан S_4 – значення індексу знаходиться вище третього граничного значення, тобто у проміжку $(\alpha_3; Dst_{\max}]$. Система здійснює переходи зі стану у стан у заздалегідь невідомі моменти часу, тож маємо справу з випадковим процесом з неперервним часом. Формування інтервалів часу перебування системи у станах S_1 , S_2 , S_3 та S_4 відбувається наступним чином: від початку відліку часу t_0 до невідомого моменту t_1 (відрізок τ_0) система знаходиться у деякому стані S_i , $i = \overline{1,4}$. Інтервал часу τ_0 має довжину, що виражена кількістю відліків часу знаходження значення Dst -індексу в проміжку $\tau_0 = t_1 - t_0$. Якщо система переходить зі стану S_1 в S_2 , то інтервал часу τ_0 віднесемо до масиву інтервалів $T_{12} = \{\tau_{12,i}, i = \overline{1, N_1}\}$, де N_1 – кількість інтервалів перебування системи s у стані S_1 . Якщо система переходить зі стану S_2 в S_1 , то τ_0 віднесемо до масиву $T_{21} = \{\tau_{21,i}, i = \overline{1, N_2}\}$, де N_2 – кількість

інтервалів перебування системи S у стані S_2 . Аналогічні міркування застосовують і для усіх інших станів та моментів часу (рис.1).

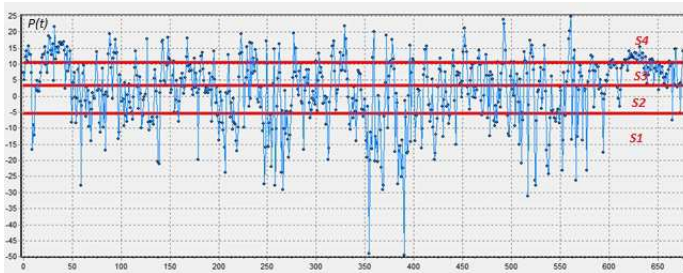


Рисунок 1 – Задання граничного значення та розбиття часового ряду на інтервали, що відповідають станам системи

Граф станів системи S має вигляд, поданий на рис.2, де $\lambda_{i,j}(t)$ – інтенсивності переходу зі стану у стан.

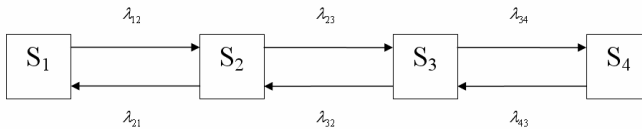


Рисунок 2 – Граф станів системи S

Дослідження поведінки системи S значно спрощується, коли виявиться, що інтенсивності $\lambda_{i,j} = const$. Остання умова є можливою, якщо розподіл інтервалів часу в масивах $T_{12}, T_{21}, T_{23}, T_{32}, T_{34}$ та T_{43} є експоненціальним: $F(\tau) = 1 - e^{-\lambda\tau}$, де τ – інтервал часу; λ – параметр розподілу, оцінку якого здійснюється так:

$$\hat{\lambda} = \frac{N}{\sum_{l=1}^N \tau_l}$$

Аналіз закону розподілу інтервалів часу в масивах $T_{12}, T_{21}, T_{23}, T_{32}, T_{34}$ та T_{43} будемо виконувати в наступній послідовності:

1. Для кожного з масивів $T_{12}, T_{21}, T_{23}, T_{32}, T_{34}$ та T_{43} будується ранжований варіаційний ряд та розраховуються основні статистичні характеристики: математичне сподівання, середньоквадратичне відхилення, коефіцієнти асиметрії та ексцесу [12].

2. Проводиться гістограма оцінка та оцінка емпіричної функції розподілу ймовірностей.

3. По кожному з масивів T_{12} , T_{21} , T_{23} , T_{32} , T_{34} та T_{43} здійснюється відтворення експоненціального розподілу, тобто визначаються оцінки $\lambda_{i,j}$.

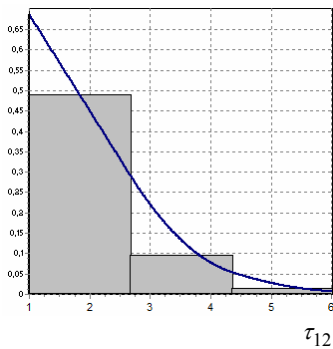
4. Виконується перевірка адекватності експоненціальної моделі розподілу на основі уточненого критерію згоди Колмогорова [13]. При цьому вважається, що модель адекватна, якщо ймовірність узгодження $P_{\text{Колм}}$ для представницьких обсягів вибірок більша, ніж 0,05.

Нижче наведено результати дослідження законів розподілу та прогнозування часового ряду залишків зі щогодинними замірами Dst -індексу. Для оцінки прогнозування часового ряду за допомогою ланцюга Маркова було використано ряд значень індексу геомагнітної активності за період з 01.01.1963 по 31.12.2012 роки, усереднений по 27 днях. Довжина ряду складає 677 позначок. Експеримент проводився для довжин вікна $L=338$. Попередньо з ряду були видалені циклічні компоненти, що показали хорошу якість прогнозу (1–3, 5–8, 11–13 компоненти), на долю яких припадає 79,134 % загальної варіабельності даних [12].

Нехай система S має три граничні межі, значення яких покладемо рівними $\alpha_1 = -4,85$, $\alpha_2 = 1,77$ та $\alpha_3 = 7,96$, що дорівнює ймовірності 0,25, 0,5 та 0,75 відповідно, тобто кількість можливих станів дорівнює чотирьом. Стан S_1 показує, що значення знаходиться нижче першого граничного значення, S_2 – між першим та другим граничним значенням, S_3 – між другим та третім граничним значенням, а S_4 – вище третього граничного значення. λ_{12} – інтенсивність переходу системи зі стану S_1 у стан S_2 , а λ_{21} – зі стану S_2 у стан S_1 , λ_{23} – зі стану S_2 у стан S_3 , λ_{32} – зі стану S_3 у стан S_2 , λ_{34} – зі стану S_3 у стан S_4 , а λ_{43} – зі стану S_4 у стан S_3 .

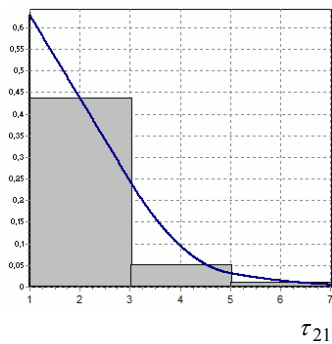
Зі стану S_1 у стан S_2 система переходила 44 рази, тобто масив T_1 містить 44 елементи, зі стану S_2 у стан S_1 – 48 разів і масив T_2 має 48 елементів; зі стану S_2 у стан S_3 - 52 рази, зі стану S_3 у стан S_2 - 59 разів, зі стану S_3 у стан S_4 – 41 раз, зі стану S_4 у стан S_3 – 38 разів. Гістограми відносних частот відповідних масивів подано на графіках (рис. 3).

ϖ_i , відносні частоти



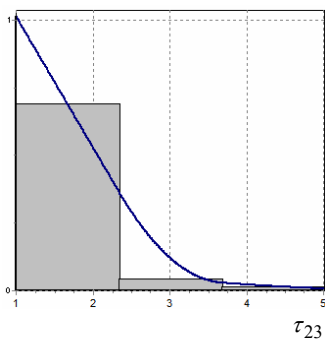
а)

ϖ_i , відносні частоти



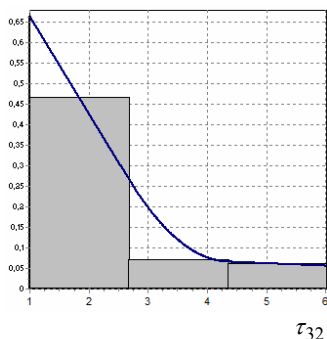
б)

ϖ_i , відносні частоти



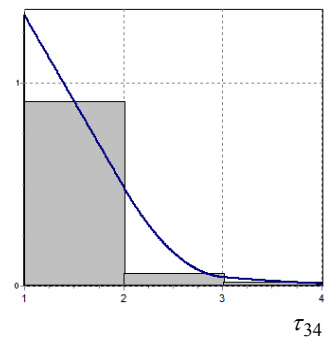
в)

ϖ_i , відносні частоти



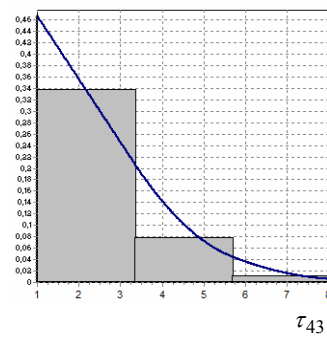
г)

ϖ_i , відносні частоти



д)

ϖ_i , відносні частоти



е)

Рисунок 3 – Гістограми відносних частот

За критерієм Колмогорова та Пірсона підтверджено гіпотезу про експоненціальний розподіл інтервалів часу між переходами системи зі стану у стан, тому для знаходження інтенсивностей переходів системи можемо скористатися формулою $\lambda_{ij} = \frac{1}{\bar{\tau}_{ij}}$, де $\bar{\tau}_{ij}$ – середній час

знаходження у стані. Знайдемо середнє значення для кожного з масивів інтервалів часу та обчислимо шість значень інтенсивностей:

$$\lambda_{12} = \frac{1}{\bar{\tau}_{12}}, \quad \lambda_{21} = \frac{1}{\bar{\tau}_{21}}, \quad \lambda_{23} = \frac{1}{\bar{\tau}_{23}}, \quad \lambda_{32} = \frac{1}{\bar{\tau}_{32}}, \quad \lambda_{34} = \frac{1}{\bar{\tau}_{34}} \quad \text{та} \quad \lambda_{43} = \frac{1}{\bar{\tau}_{43}}, \quad \text{тобто}$$

$$\lambda_{12} = 0,579; \quad \lambda_{21} = 0,552; \quad \lambda_{23} = 0,732; \quad \lambda_{32} = 0,513; \quad \lambda_{34} = 0,804; \quad \lambda_{43} = 0,452;$$

Тепер будемо досліджувати поведінку системи у динамічному та стаціонарному режимах.

Динамічний режим:

Для знаходження ймовірностей станів системи складемо рівняння Колмогорова:

$$\left\{ \begin{array}{l} \frac{dp_1(t)}{dt} = -\lambda_{12}p_1(t) + \lambda_{21}p_2(t), \\ \frac{dp_2(t)}{dt} = \lambda_{12}p_1(t) - (\lambda_{21} + \lambda_{23})p_2(t) + \lambda_{32}p_3(t), \\ \frac{dp_3(t)}{dt} = \lambda_{23}p_2(t) - (\lambda_{32} + \lambda_{34})p_3(t) + \lambda_{43}p_4(t), \\ \frac{dp_4(t)}{dt} = \lambda_{34}p_3(t) - \lambda_{43}p_4(t). \end{array} \right.$$

Гранична умова:

$$p_1(t) + p_2(t) + p_3(t) + p_4(t) = 1.$$

У початковий момент часу значення індексу геомагнітної активності Dst знаходилося вище другого граничного значення та нижче третього, тобто у стані S_3 . Тоді початкові умови будуть наступними: $p_1(0) = 0$; $p_2(0) = 1$; $p_3(0) = 0$; $p_4(0) = 0$;

Отже, з урахуванням значень інтенсивностей переходів, маємо:

$$\begin{cases} \frac{dp_1}{dt} = -0,579 p_1(t) + 0,552 p_2(t) \\ \frac{dp_2}{dt} = 0,579 p_1(t) - 1,284 p_2(t) + 0,513 p_3(t), \\ \frac{dp_3}{dt} = 0,732 p_2(t) - 1,317 p_3(t) + 0,452 p_4(t), \\ \frac{dp_4}{dt} = 0,804 p_3(t) - 0,452 p_4(t). \end{cases}$$

Розв'язок рівняння для динамічного режиму подано на рис. 4.

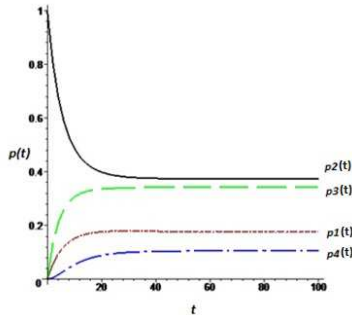


Рисунок 4 – Ймовірності станів системи S у динамічному режимі

Для стаціонарного режиму система рівнянь для обчислення фінальних ймовірностей станів матиме наступний вигляд:

$$\begin{cases} \lambda_{21} p_2 - \lambda_{12} p_1 = 0, \\ \lambda_{32} p_3 - \lambda_{23} p_2 = 0, \\ \lambda_{43} p_4 - \lambda_{34} p_3 = 0, \\ p_1 + p_2 + p_3 + p_4 = 1. \end{cases}$$

Підставимо в неї значення інтенсивностей переходів та знайдемо фінальні ймовірності:

$$\begin{cases} 0,552 p_2 - 0,579 p_1 = 0 \\ 0,513 p_3 - 0,732 p_2 = 0 \\ 0,452 p_4 - 0,804 p_3 = 0 \\ p_1 + p_2 + p_3 + p_4 = 1 \end{cases}$$

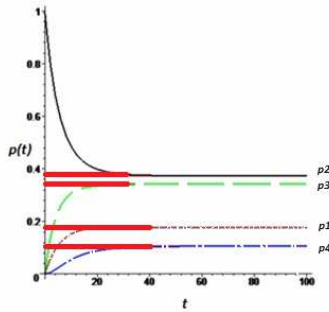


Рисунок 5 – Ймовірності станів та фінальні ймовірності системи S

У результаті отримаємо: $p_1 = 0,173$, $p_2 = 0,377$, $p_3 = 0,346$, $p_4 = 0,104$. Тобто можна зробити висновок, що у наступний момент часу найбільш вірогідним є перехід системи у стан S_2 . Загальний розв'язок системи можна подати у вигляді:

$$p_{зал, l}(t+l) = \begin{cases} -0,579 p_1(l) + 0,552 p_2(l), l=1,2... \\ 0,579 p_1(l) - 1,284 p_2(l) + 0,513 p_3(l), l=1,2... \\ 0,732 p_2(l) - 1,317 p_3(l) + 0,452 p_4(l), l=1,2... \\ 0,804 p_3(l) - 0,452 p_4(l), l=1,2... \end{cases}$$

Тоді прогноз для $\forall t$ на l кроків вперед буде виглядати наступним чином: $p(t+l) = p_T(t+l) + p_{зал}(t+l)$, де $p_T(t+l)$ визначається згідно з роботою [12].

Висновок. У роботі запропоновано метод та побудовано алгоритм дослідження моделі системи на основі ланцюгів Маркова, яка описує поведінку індексу геомагнітної активності Dst. Досліджено поведінку моделі в динамічному та стаціонарному режимах: особливості формування масивів інтервалів часу, побудови систем рівнянь Колмогорова та їх розв'язання, знаходження фінальних ймовірностей. Подальші дослідження будуть спрямовані на апробацію експерименту для визначення сильних ($Dst < -50$) та дуже сильних ($Dst \leq -100$) магнітних бур.

Бібліографічні посилання

1. Сулейменов И. Э. Воздействие на процессы в атмосфере и проблематика геофизических вооружений. / И. Э. Сулейменов – Алматы : Изд-во Казахского нац. ун-та, 2007.

2. **Schindler K.** Physics of Space Plasma Activity. / K. Schindler – New York: Cambridge Univ. Press, 2007. – 508 p.

3. Архив параметров солнечной активности [Электронный ресурс]. – Режим доступа : <http://omniweb.gsfc.nasa.gov/form/dx1.html>

4. **Романовский В. И.** Дискретные цепи Маркова / В. И. Романовский. – М. : Гостехиздат, 1949. – 436 с.

5. **Джефферс Дж.** Введение в системный анализ: применение в экологии / Дж. Джефферс. – М. : Мир., 1981. – 252 с.

6. **Вагер Б. Г.** Оценивание и прогнозирование колебаний метеорологических элементов с использованием марковских цепей. / Б. Г. Вагер // Межвуз. сб. тр.: Численные методы в математическом моделировании гидродинамических и технологических процессов. – Л. : ЛИСИ, 1989. – С. 90–93.

7. **Praedrich K.** Combining predictive schemes in short-term forecasting / K. Praedrich, L. Leslie // Mon. Weather Rev. – 1987. – Vol. 115. – № 8. – P. 1640–1644.

8. **Praedrich K.** Evaluation of techniques for the operational, single station, short-term forecasting of rainfall at a medlatitude station / K. Praedrich, L. Leslie // Mon. Weather Rev. – 1987. – Vol. 115. – № 8. – P. 1645–1654.

9. **Thyer M.** A hidden Markov model for modeling longterm persistence in multi-siterainfall time series / M. Thyer, G. Kuczera // J. Hydrol. – 2003. – 275. – № 1–2. – P. 11–16.

10. **Dubrovsky M.** Might-frequency and lowfrequency variability in stochastic daily weather generator and its effect on agricultural and hydrologic modeling / M. Dubrovsky, J. Buchtele, Z. Zalud. // Clim. change. – 2004. – 63. № 1–2. – P. 145–179.

11. **Greene A.** Downscaling projection of Indian rainfall using a non-homogeneous hidden Markov model / A. Greene, A. Robertson, P. Smyth, S. Triqlia // J. Meteorol. Soc. – 2011. – 137. – № 655. – P. 347–359.

12. **Шаповал Г.А.** Оцінка якості прогнозування циклічних компонент часових рядів значень індексу геомагнітної активності Dst з використанням сингулярного спектрального аналізу / Г.А. Шаповал. – Наукові технології. – 2015. – № 1 (25). – С. 93–98.

13. **Приставка П.О.** Аналіз даних: навч. посіб. / МОН України. / П.О. Приставка, О.М. Мацуга. – Д. : Вид-во ДНУ, 2008. – 92 с.

Надійшла до редколегії 04.11.2015