УДК 519.246.8 DOI: https://doi.org/10.15407/techned2019.02.038

СТАТИСТИЧНИЙ АНАЛІЗ ПЕРІОДИЧНОЇ АВТОРЕГРЕСІЇ З ВИПАДКОВИМИ КОЕФІЦІЄНТАМИ В ЗАДАЧАХ ОПЕРАТИВНОГО ПРОГНОЗУВАННЯ ЕЛЕКТРОСПОЖИВАННЯ ПІДПРИЄМСТВ

М.Є. Фриз[1*], канд.техн.наук, **Л.М. Щербак**[2**], докт.техн.наук
[1] - **Тернопільський національний технічний університет імені Івана Пулюя,**
вул. Руська, 56, Тернопіль, 46001, Україна, e-mail: mykh.fryz@gmail.com
[2] - **Київський міжнародний університет,**
вул. Львівська, 49, Київ, 03179, Україна, e-mail: prof_scherbak@ukr.net

Охарактеризовано властивості та застосування періодичних умовних лінійних випадкових процесів (УЛВП) у задачах математичного моделювання, аналізу та оперативного прогнозування електроспоживання. Обгрунтовано можливість здійснення статистичного аналізу УЛВП з дискретним часом із використанням моделі авторегресії з випадковими коефіцієнтами. Така модель є частинним випадком УЛВП. Запропоновано метод оцінювання параметрів послідовності періодичної авторегресії з випадковими коефіцієнтами, суть якого полягає у представленні досліджуваної послідовності у вигляді сукупності L (де L – період) стаціонарних і стаціонарно зв'язаних підпослідовностей та застосуванні до кожної з них двоетапного методу найменших квадратів для знаходження відповідних оцінок. Наведено результати комп'ютерного імітаційного моделювання, що підтверджують слушність запропонованих статистичних оцінок. Розглянуто приклад здійснення оперативного прогнозування електроспоживання організації, що належить до класу малих та середніх підприємств, з використанням моделі періодичної авторегресії з випадковими коефіцієнтами. Бібл. 16, рис. 4, табл. 2.
Ключові слова: математична модель, умовний лінійний випадковий процес, період, авторегресія з випадковими коефіцієнтами, оцінювання параметрів, імітаційна модель, прогнозування, електроспоживання.

Вступ. Обгрунтування математичних моделей інформативних стохастичних сигналів, процесів та завад є одним із найбільш важливих етапів розробки інформаційно-вимірювальних та інформаційно-керуючих систем в електроенергетиці [1, 2], а також автоматизованих систем аналізу ресурсоспоживання (електро-, газо-, водоспоживання), інформаційних технологій медичної діагностики. Математична модель є теоретичним підгрунтям структурної, програмної та технічної реалізації розроблюваних систем і технологій, основою алгоритмів опрацювання досліджуваних сигналів та методів прийняття рішень. Тому модель повинна бути адекватною вимірювальному сигналу чи процесу, відображати фізичний механізм його породження, а також бути придатною для здійснення його теоретичного аналізу та вирішення задач ідентифікації та оцінювання інформативних характеристик за результатами експериментів, задач прогнозування та побудови імітаційних моделей.

Наведеним вище вимогам задовольняють моделі сигналів у вигляді лінійних випадкових процесів (ЛВП) та послідовностей [1−3]. У згаданих прикладних областях поширеним є також використання лінійних періодичних випадкових процесів [1, 2], які дають можливість врахувати циклічні властивості досліджуваних сигналів, спричинені, наприклад, ритмічним характером навантажень енерго-, газо- чи водопостачальних систем, вібрацій підшипників електричних машин, імпедансу біологічних тканин живих організмів, електрофізіологічних сигналів (у задачах медичного приладобудування).

Найчастіше ЛВП використовується в задачах математичного моделювання сигналів, які можна зобразити у вигляді суми великого числа стохастично незалежних імпульсів із випадковими параметрами, які виникають у послідовні моменти часу, що утворюють пуассонівський потік [1−3]. Але якщо ці імпульси є стохастично залежними випадковими функціями, то адекватною математичною моделлю буде умовний лінійний випадковий процес (УЛВП) [4−6], який представляє собою стохастичний інтеграл від випадкової функції (ядра) за процесом із незалежними приростами (ЛВП має подібну конструкцію, у якій ядро – невипадкова функція).

ORCID ID: *http://orcid.org/0000-0002-8720-6479, **http://orcid.org/0000-0002-1536-4806

Для задач статистичного оцінюванння, прогнозування та прийняття рішень найчастіше використовується ЛВП із дискретним часом у вигляді послідовності авторегресії, а для циклічних процесів – періодичної авторегресії [2, 7]. Аналогічно, ефективним інструментом статистичного аналізу УЛВП є послідовність авторегресії з випадковими коефіцієнтами (Random Coefficient Autoregressive (RCA) Model) [4, 8]. Властивості та методи оцінювання параметрів моделі RCA у випадку її стаціонарності досліджені багатьма авторами (фундаментальні результати наведено у монографії [8]). У роботі [9] охарактеризовано модель періодичної авторегресії з випадковими коефіцієнтами (Random Coefficient Periodic Autoregressive (RCPAR) Model), встановлено умови періодичності її моментних функцій та багатовимірних функцій розподілу. Автори [10] досліджували частинний випадок моделі RCPAR (детальніше див. далі), де періодичними є лише математичні сподівання відповідних коефіцієнтів, отримано оцінки характеристик такої моделі. Таким чином, актуальною є задача оцінювання параметрів моделі періодичної авторегресії з випадковими коефіцієнтами в загальному випадку.

У роботі [10] побудовано статистичні оцінки параметрів для частинного випадку моделі RCPAR, а саме: відповідні коефіцієнти моделі є незалежними, періодичними є лише їхні математичні сподівання, які змінюються за косинусоїдальним законом.

Метою даної статті є обгрунтування методу статистичного оцінювання параметрів моделі RCPAR для загального випадку та можливостей її застосування для оперативного прогнозування електроспоживання підприємств.

Умовні лінійні випадкові процеси. Умовним лінійним випадковим процесом $\xi(t)$, $t \in (-\infty, \infty)$ називається стохастичний інтеграл виду

$$\xi(t) = \int_{-\infty}^{\infty} \varphi(\tau, t) d\eta(\tau), \tag{1}$$

де $\varphi(\tau,t), \tau, t \in (-\infty,\infty)$ − дійсна *випадкова* функція (ядро), $\eta(\tau), \tau \in (-\infty,\infty)$, $\mathbf{P}(\eta(0)=0)=1$ − дійсний гільбертовий стохастично неперервний випадковий процес із незалежними приростами; $\mathbf{M}\eta(\tau) = a(\tau) < \infty$ і $\mathbf{D}\eta(\tau) = b(\tau) < \infty$, $\forall\tau$; випадкові функції $\varphi(\tau,t)$ і $\eta(\tau)$ є стохастично незалежними.

На відміну від (1) у роботі [6] розглядався УЛВП з однорідним центрованим породжуючим процесом $\eta(\tau)$, тобто $a(\tau) = 0$, $b(\tau) = \text{const}$.

Стохастичний інтеграл (1) існує в розумінні збіжності в середньоквадратичному розумінні послідовності відповідних інтегральних сум тоді й тільки тоді, коли [5]

$$\mathbf{M}\big(\xi(t)\big)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{M}\big(\varphi(\tau_1,t)\varphi(\tau_2,t)\big) da(\tau_1) da(\tau_2) + \int_{-\infty}^{\infty} \mathbf{M}\varphi^2(\tau,t) db(\tau) < \infty. \tag{2}$$

Зауважимо, що формально лінійний випадковий процес відрізняється від умовного лінійного випадкового процесу тим, що для ЛВП [3] ядро $\varphi(\tau,t)$ у виразі (1) є *невипадковою* функцією.

Умовний (умовно) лінійний випадковий процес (1) вперше охарактеризовано у [6] для випадку однорідного та центрованого процесу $\eta(\tau)$. Зокрема, автор [6] досліджував центральну граничну теорему для послідовності лінійних функціоналів від умовно лінійних процесів із неперервним та дискретним часом із застосуванням до задач математичного моделювання радіолокаційних завад.

Якщо $\eta(\tau)$ – простий пуассонівський процес, то вираз (1) можна записати наступним чином:

$$\xi(t) = \sum_{k=-\infty}^{\infty} \varphi(\tau_k, t), \tag{3}$$

де $\ldots < \tau_{k-1} < \tau_k < \tau_{k+1} < \ldots$ – моменти стрибків пуассонівського процесу, що дорівнюють моментам появи випадкових імпульсів $\varphi(\tau_k,t)$ ($\varphi(\tau_k,t) = 0$ при $t < \tau_k$), причому випадкові функції $\ldots, \varphi(\tau_{k-1},t)$, $\varphi(\tau_k,t)$, $\varphi(\tau_{k+1},t), \ldots$ є стохастично залежними.

В основі моделювання навантажень енергосистем у вигляді лінійного періодичного випадкового процесу лежить зображення досліджуваного процесу у вигляді суми великого числа *незалежних* імпульсів (із випадковою тривалістю та амплітудою), що виникають у пуассонівські моменти часу. Автори [11] досліджували дані електроспоживання на рівні окремих домогосподарств, отримані з використанням сучасних інтелектуальних лічильників електроенергії. За результатами експериментальних досліджень в [11] показано, що процеси електроспоживання окремих домогосподарств є стохастично *залежними*. Це підтверджує адекватність моделі УЛВП для задач математичного моделювання та прогнозування процесів електроспоживання на рівні домогосподарств, житлових районів та окремих підприємств.

Наведемо ще дві властивості УЛВП, які є важливими для практичних застосувань. Нехай $\phi(\tau,t)=\mathbf{M}\varphi(\tau,t)$ і $R_{\varphi}(\tau_1,\tau_2;t_1,t_2)=\mathbf{M}\big(\varphi_0(\tau_1,t_1)\varphi_0(\tau_2,t_2)\big)$ − відповідно математичне сподівання та кореляційна функція ядра УЛВП (де $\varphi_0(\tau,t)=\varphi(\tau,t)-\mathbf{M}\varphi(\tau,t)$).

Якщо процес $\eta(\tau)$ є однорідним і виконуються умови $\phi(\tau,t)=\phi(t-\tau)$ і $R_{\varphi}(\tau_1,\tau_2;t_1,t_2)=R_{\varphi}(\tau_1,\tau_2;t_2-t_1)$, то УЛВП (1) є слабо стаціонарним випадковим процесом [5].

Якщо існує число (період) $T>0$ таке, що для процесу $\eta(\tau)$ виконуються умови $da(\tau)=da(\tau+T)$ і $db(\tau)=db(\tau+T)$, а математичне сподівання та кореляційна функція ядра $\varphi(\tau,t)$ є періодичними за сукупністю своїх аргументів, тобто $\phi(\tau,t)=\phi(\tau+T,t+T)$ і $R_{\varphi}(\tau_1,\tau_2;t_1,t_2)=R_{\varphi}(\tau_1+T,\tau_2+T;t_1+T,t_2+T)$, то (1) є періодично корельованим процесом [5].

Умовний лінійний випадковий процес із дискретним часом зображається так [12]:

$$\xi_t=\sum_{\tau=-\infty}^{\infty}\varphi_{\tau,t}\eta_{\tau},\ t\in\mathbf{Z}, \qquad (4)$$

де $\varphi_{\tau,t},\ \tau,t\in\mathbf{Z}$ − випадкова функція (ядро), $\eta_{\tau},\ \tau\in\mathbf{Z}$ − послідовність незалежних випадкових величин із скінченною дисперсією; $\varphi_{\tau,t}$ та η_{τ} − стохастично незалежні, причому у даній роботі ми будемо розглядати тільки випадок $\mathbf{M}\eta_{\tau}=0$, тоді для збіжності в середньоквадратичному розумінні ряду (4) ядро $\varphi_{\tau,t}$ мусить задовольняти умові $\mathbf{M}\sum_{\tau=-\infty}^{\infty}\left|\varphi_{\tau,t}\right|^2<\infty$, $\forall t\in\mathbf{Z}$.

Лінійну випадкову послідовність отримаємо з (4), якщо $\varphi_{\tau,t}$ – невипадкова функція.

Періодична авторегресія з випадковими коефіцієнтами. Оцінювання характеристик. Ідея авторів щодо статистичного оцінювання УЛВП (4) базується на таких міркуваннях. Лінійна випадкова послідовність (ЛВП) може розглядатися як відгук лінійного цифрового фільтра (з невипадковими параметрами, які в загальному випадку змінюються з часом, наприклад, періодично) на вхідний сигнал у вигляді послідовності незалежних випадкових величин (білий шум у вузькому розумінні). Найбільш важливою для практичних застосувань є ЛВП у формі послідовності авторегресії, що є відгуком рекурсивного цифрового фільтра на вхідний білий шум.

УЛВП із дискретним часом (4) можна трактувати як відгук цифрового фільтра з випадковими параметрами (імпульсна реакція $\varphi_{\tau,t}$ є випадковою функцією) на вхідний білий шум η_{τ}. Якщо цей фільтр побудувати так, щоб він мав тільки рекурсивну структуру, то випадковий сигнал на його виході буде представляти собою послідовність авторегресії з випадковими коефіцієнтами.

Послідовність авторегресії з випадковими коефіцієнтами порядку $p\in\mathbf{N}$ зображається наступним чином [8, 9]:

$$\xi_t=\sum_{k=1}^{p}(a_{k,t}+\alpha_{k,t})\xi_{t-k}+\eta_t,\quad t\in\mathbf{Z}, \qquad (5)$$

де η_t − центрована послідовність незалежних випадкових величин з дисперсією $\mathbf{D}\eta_t=\sigma_t^2<\infty$; $\mathbf{a}_t=(a_{1,t},a_{2,t},...,a_{p,t})'$ − послідовність векторів невипадкових коефіцієнтів (тут і далі $\mathbf{A}'$ − матриця, транспонована до матриці $\mathbf{A}$); $\boldsymbol{\alpha}_t=(\alpha_{1,t},\alpha_{2,t},...,\alpha_{p,t})'$ − послідовність незалежних центрованих випадкових векторів, з кореляційними матрицями $\mathbf{R}_t=\mathbf{M}(\boldsymbol{\alpha}_t\boldsymbol{\alpha}_t')$; $\boldsymbol{\alpha}_t$ і η_t є незалежними.

Як вже зазначалося вище, найбільш вивченим є випадок стаціонарної моделі RCA, для якої $\mathbf{D}\eta_t=\sigma_t^2=\sigma^2$, $\mathbf{a}_t=\mathbf{a}=(a_1,a_2,...,a_p)'$, $\mathbf{R}_t=\mathbf{R}$, тобто характеристики не залежать від часу.

Послідовність (5) називається періодичною авторегресією з випадковими коефіцієнтами (RCPAR) [9], якщо існує таке число $L\in\mathbf{N}$ (період), що

$$\sigma_t^2=\sigma_{t+L}^2,\qquad \mathbf{a}_t=\mathbf{a}_{t+L},\qquad \mathbf{R}_t=\mathbf{R}_{t+L}. \qquad (6)$$

Властивості періодичності ймовірнісних розподілів моделі RCPAR обгрунтовано в [9]. Автори [10] досліджували методи оцінювання характеристик моделі RCPAR, однак за певних обмежень, зокрема невипадкові параметри $a_{1,t}, a_{2,t}, ..., a_{p,t}$ розглядаються як косинусоїдальні функції з періодом L (і оцінюються параметри цих функцій), крім того у моделі [10] $\mathbf{D}\eta_t = \sigma^2 = \mathrm{const}$ компоненти вектора $\boldsymbol{\alpha}_t$ є незалежними й також мають постійні дисперсії. Нашим завданням далі є побудова методу статистичного оцінювання характеристик моделі RCPAR для загального випадку.

Зауваження. Далі використовуються деякі спеціальні матричні перетворення [8, 13], які необхідно пояснити. Зокрема, нехай $\mathbf{G}$ − $(m \times n)$-матриця з елементами $g_{ij}, i = \overline{1,m}, j = \overline{1,n}$. Позначимо також $\mathbf{g}_j = (g_{1j}, g_{2j}, ..., g_{mj})'$ − j-й стовпець матриці $\mathbf{G}$. Векторизацією (vectorization) матриці $\mathbf{G}$ називається вектор-стовпець $vec(\mathbf{G}) = (\mathbf{g}'_1 \ \ \mathbf{g}'_2 \ \ ... \ \ \mathbf{g}'_n)'$. Тобто, $vec(\mathbf{G})$ утворюється послідовним розміщенням стовпців матриці $\mathbf{G}$ один під одним, починаючи з першого. Наприклад, якщо $\mathbf{G}$ є (3×3)-матрицею, то $vec(\mathbf{G}) = (g_{11}, g_{21}, g_{31}, g_{12}, g_{22}, g_{32}, g_{13}, g_{23}, g_{33})'$.

Нехай тепер $\mathbf{G}$ − *симетрична* $(p \times p)$-матриця. Позначимо $\mathbf{h}_j = (g_{jj}, g_{j+1,j}, ..., g_{pj})'$, $j = \overline{1,p}$. Напіввекторизацією (half-vectorization) симетричної матриці $\mathbf{G}$ називається вектор-стовпець $vech(\mathbf{G}) = (\mathbf{h}'_1 \ \ \mathbf{h}'_2 \ \ ... \ \ \mathbf{h}'_p)'$. Наприклад, якщо $\mathbf{G}$ є симетричною (3×3)-матрицею, то $vech(\mathbf{G}) = (g_{11}, g_{21}, g_{31}, g_{22}, g_{32}, g_{33})'$. Іншими словами, $vech(\mathbf{G})$ можна утворити з $vec(\mathbf{G})$ «відкиданням» усіх наддіагональних елементів $g_{ij}, i < j$.

Отже, за результатами спостережень послідовності ξ_t потрібно оцінити компоненти векторів $\mathbf{a}_t = (a_{1,t}, a_{2,t}, ..., a_{p,t})', t = \overline{1,L}$, дисперсії $\sigma_t^2, t = \overline{1,L}$, а також елементи кореляційних $(p \times p)$-матриць $\mathbf{R}_t, t = \overline{1,L}$. Зауважимо, що оскільки кожна матриця $\mathbf{R}_t$ є симетричною, то потрібно оцінювати лише $p(p+1)/2$-елементні вектори $\boldsymbol{\gamma}_t = vech(\mathbf{R}_t), \quad t = \overline{1,L}$.

Для зручності подальших викладок запишемо (5) у такій формі [8]:

$$\xi_t = \sum_{k=1}^{p} a_{k,t}\xi_{t-k} + u_t = \mathbf{x}'_{t-1}\mathbf{a}_t + u_t, \tag{7}$$

де $u_t = \sum_{k=1}^{p} \alpha_{k,t}\xi_{t-k} + \eta_t = \mathbf{x}'_{t-1}\boldsymbol{\alpha}_t + \eta_t, \qquad \mathbf{x}'_t = (\xi_t, \xi_{t-1}, ..., \xi_{t+1-p})$.

Нехай $(\xi_{1-p}, ..., \xi_0, \xi_1, ..., \xi_{qL})$ − вибірка (обсяг вибірки дорівнює $qL + p$) спостережень послідовності (7), де $q \in \mathbf{N}$ (кількість циклів) і період L вважається відомим.

У [8] обгрунтовано двоетапний алгоритм оцінювання параметрів стаціонарної моделі RCA методом найменших квадратів. Отримувані оцінки є слушними (тобто вони збігаються за ймовірністю до оцінюваних параметрів при зростанні обсягу вибірки) та асимптотично нормальними. Розроблюваний нами метод оцінювання параметрів моделі RCPAR базується на результатах [8], але з врахуванням періодичних властивостей послідовності (7).

Враховуючи періодичність ймовірнісного розподілу процесу (7) та періодичність його параметрів, розглянемо L вкладених підпослідовностей, а саме

$$\zeta_{l,s} = \xi_{(s-1)L+l} = \mathbf{x}'_{(s-1)L+l-1}\mathbf{a}_{(s-1)L+l} + u_{(s-1)L+l}, \quad l = \overline{1,L}, \quad s = \overline{1,q}. \tag{8}$$

Кожна l-та послідовність із (8) є стаціонарною (як функція від s) послідовністю, крім того, усі вони є стаціонарно зв'язаними.

Таким чином, на першому етапі застосовуємо метод найменших квадратів [8] до (8) для кожного $l = \overline{1,L}$, отримуючи оцінки $\hat{\mathbf{a}}_l = (\hat{a}_{1,l}, \hat{a}_{2,l}, ..., \hat{a}_{p,l})'$, $l = \overline{1,L}$ невипадкових параметрів моделі RCPAR у такому вигляді:

$$\hat{\mathbf{a}}_l = \left(\sum_{s=1}^{q} \mathbf{x}_{(s-1)L+l-1}\mathbf{x}'_{(s-1)L+l-1} \right)^{-1} \sum_{s=1}^{q} \mathbf{x}_{(s-1)L+l-1}\xi_{(s-1)L+l}, \quad l = \overline{1,L}. \tag{9}$$

На другому етапі, використовуючи оцінки $\hat{\mathbf{a}}_l$, $l=\overline{1,L}$ отримуємо наступне:

$$\hat{u}_{(s-1)L+l}=\xi_{(s-1)L+l}-\mathbf{x}'_{(s-1)L+l-1}\hat{\mathbf{a}}_l,\quad l=\overline{1,L}. \tag{10}$$

Відповідно до [8] та з урахуванням (8) статистичні оцінки $\hat{\boldsymbol{\gamma}}_l$ та $\hat{\sigma}_l^2$ параметрів $\boldsymbol{\gamma}_l$ та σ_l^2, $l=\overline{1,L}$ будуть мати вигляд

$$\hat{\boldsymbol{\gamma}}_l=\left(\sum_{s=1}^{q}\left(\mathbf{z}_{(s-1)L+l}-\overline{\mathbf{z}}_l\right)\left(\mathbf{z}_{(s-1)L+l}-\overline{\mathbf{z}}_l\right)'\right)^{-1}\sum_{s=1}^{q}\hat{u}^2_{(s-1)L+l}\left(\mathbf{z}_{(s-1)L+l}-\overline{\mathbf{z}}_l\right),\quad l=\overline{1,L}, \tag{11}$$

$$\hat{\sigma}_l^2=\frac{1}{q}\sum_{s=1}^{q}\hat{u}^2_{(s-1)L+l}-\overline{\mathbf{z}}'_l\hat{\boldsymbol{\gamma}}_l,\quad l=\overline{1,L}, \tag{12}$$

де $\mathbf{z}_t=\mathbf{D}'vec(\mathbf{x}_{t-1}\mathbf{x}'_{t-1})$, $\overline{\mathbf{z}}_l=\frac{1}{q}\sum_{s=1}^{q}\mathbf{z}_{(s-1)L+l}$, $l=\overline{1,L}$; $\mathbf{D}$ − матриця дублювання (duplication matrix) [13], яка має розмір $p^2\times p(p+1)/2$ і будується наступним чином: $\mathbf{D}'=\sum_{i=1}^{p}\sum_{j=1}^{i}\mathbf{v}_{ij}\left(vec(\mathbf{T}_{ij})\right)'$, де $\mathbf{v}_{ij}$ − вектор, що містить $p(p+1)/2$ елементів, серед яких $\left[(j-1)p+i-j(j-1)/2\right]$-й елемент дорівнює одиниці, а інші елементи дорівнюють нулю ($1\le j\le i\le p$); $\mathbf{T}_{ij}$ − $(p\times p)$-матриця, у якої (i,j)-й та (j,i)-й елементи дорівнюють одиниці, а решта елементів дорівнюють нулю.

Результати комп'ютерного імітаційного моделювання. Імітаційне моделювання послідовності періодичної авторегресії з випадковими коефіцієнтами здійснювалося з використанням співвідношення (5). Були задані такі характеристики імітаційної моделі: період $L=3$, порядок моделі $p=2$, розподіл випадкових послідовностей $\alpha_{1,t}$, $\alpha_{2,t}$ і η_t – нормальний, крім того, $\alpha_{1,t}$ та $\alpha_{2,t}$ моделювалися як незалежні випадкові послідовності, тобто матриці $\mathbf{R}_t$ − діагональні. Інші задані параметри імітаційної моделі та результати їхнього оцінювання наведено далі у табл. 1 (у таблиці $S_{1,t}, S_{2,t}$, $t=\overline{1,L}$ − дисперсії випадкових величин $\alpha_{1,t}$ та $\alpha_{2,t}$).

Таблиця 1

Параметр	Задане значення	Середнє	Дисперсія	Середнє	Дисперсія	Середнє	Дисперсія
		$q=200$		$q=500$		$q=1000$	
$a_{1,1}$	0	0.0018	0.0059	0.0042	0.0032	-0.0093	0.0017
$a_{1,2}$	0.1	0.0991	0.0176	0.1014	0.0094	0.094	0.0056
$a_{1,3}$	-0.2	-0.1942	0.0055	-0.1968	0.0021	-0.1965	0.0011
$a_{2,1}$	0.36	0.3562	0.0057	0.3666	0.0031	0.3575	0.0013
$a_{2,2}$	-0.4	-0.4065	0.0079	-0.409	0.004	-0.4006	0.0018
$a_{2,3}$	-0.5	-0.4923	0.0128	-0.4956	0.0062	-0.4902	0.0034
$S_{1,1}$	0.22	0.1826	0.0298	0.2042	0.012	0.2118	0.0086
$S_{1,2}$	0.3	0.2425	0.053	0.2416	0.031	0.2525	0.016
$S_{1,3}$	0.15	0.1337	0.0101	0.137	0.0094	0.1506	0.0055
$S_{2,1}$	0.2	0.1865	0.0207	0.1886	0.0099	0.1944	0.0084
$S_{2,2}$	0.1	0.1003	0.0217	0.0929	0.0111	0.0903	0.0078
$S_{2,3}$	0.25	0.2185	0.0413	0.2124	0.0339	0.2331	0.0284
σ_1^2	0.16	0.1977	0.0423	0.1909	0.0275	0.1707	0.0156
σ_2^2	1	1.0306	0.0428	1.0416	0.0186	1.037	0.0116
σ_3^2	0.49	0.5402	0.0313	0.5337	0.0269	0.4935	0.0188

На рис. 1 показано приклад згенерованої реалізації послідовності періодичної авторегресії з випадковими коефіцієнтами із заданими характеристиками. Тут важко оцінити циклічний характер послідовності ξ_t, але наявність циклічності можна підтвердити статистичним аналізом.

Зокрема, на рис. 2 представлено реалізації оцінок кореляційної функції $R_{t,\tau}$ випадкової послідовності ξ_t: на рис. 2, *а* − для $\tau - t = 0$ ($R_{t,t}$ − це дисперсія послідовності ξ_t), на рис. 2, *б* − для $\tau - t = 1$ і на рис. 2, *в* − для $\tau - t = 2$ (оцінки обчислено для цілочисельних значень t, а з'єднання лініями на рис. 2 виконано для кращої візуалізації результату).

Імітаційне моделювання було проведено для різної кількості циклів q (200, 500 і 1000). При кожному q було згенеровано 100 незалежних реалізацій послідовності RCPAR, обчислено оцінки параметрів моделі для кожної її згенерованої реалізації.

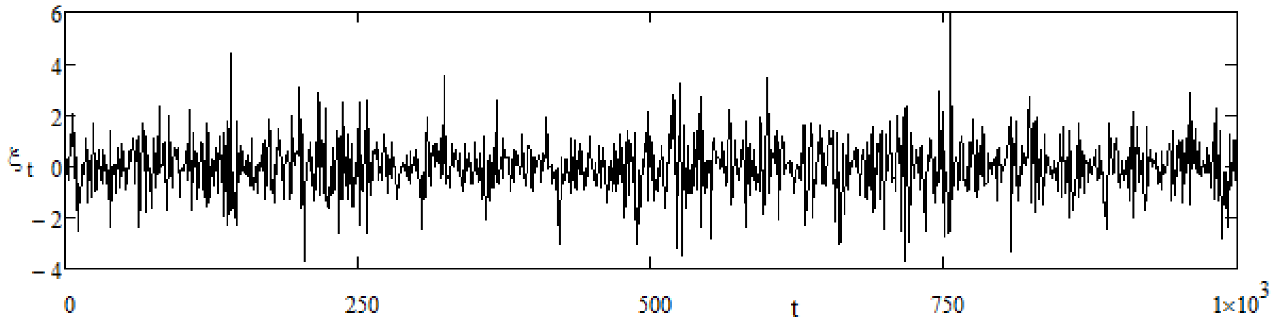

Рис. 1

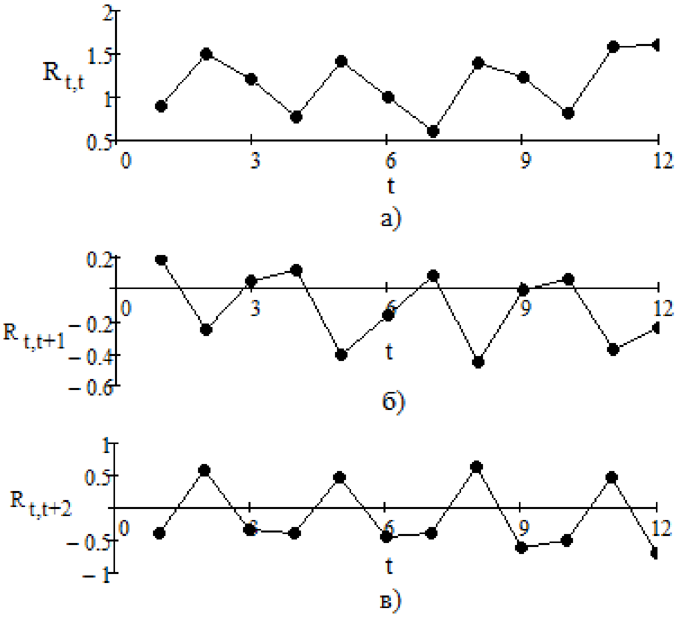

Рис. 2

Далі обчислено вибіркове середнє та вибіркову дисперсію оцінок кожного параметра (усередненням за множиною 100 реалізацій). Наприклад, якщо $a_{1,1}$ − параметр, $\hat{a}_{1,1}(k)$ − його оцінка, обчислена на основі k -ої згенерованої реалізації, $k = \overline{1,100}$, то у відповідному рядку табл. 1 наведено:

Середнє $= \overline{a}_{1,1} = \dfrac{1}{100}\sum_{k=1}^{100}\hat{a}_{1,1}(k)$ і

Дисперсія $= \dfrac{1}{99}\sum_{k=1}^{100}\left(\hat{a}_{1,1}(k) - \overline{a}_{1,1}\right)^2$. Результати, наведені у табл. 1, підтверджують слушність використовуваних оцінок.

Прогнозування. Лінійна модель авторегресії з постійними коефіцієнтами є однією з найбільш поширених у задачах прогнозування навантажень енергосистем [14, 15]. Зокрема, в Інституті електродинаміки НАН України автори [14] розробили методику оперативного та короткострокового прогнозування електроспоживання обласної електроенергетичної системи, а також енергоємних промислових підприємств на основі авторегресійної моделі Бокса-Дженкінса. Однак для електроспоживання окремих домогосподарств, малих та середніх підприємств характерною є «значна нерегулярність» [11], суттєві коливання дисперсії, а також добова циклічність. Модель періодичної авторегресії з випадковими коефіцієнтами, для якої характерна умовна гетероскедастичність (випадкова умовна дисперсія) [8] дає змогу врахувати такі особливості процесу електроспоживання в задачах його оперативного прогнозування. А саме, оптимальний у середньоквадратичному розумінні прогноз $\hat{\xi}_t$ послідовності авторегресії з випадковими коефіцієнтами на основі спостереження попередніх її відліків $\mathbf{x}'_{t-1} = (\xi_{t-1}, \xi_{t-2}, ..., \xi_{t-p})$ має вигляд [8]

$$\hat{\xi}_t = \mathbf{M}\left(\xi_t \middle| \mathbf{x}'_{t-1}\right) = \sum_{k=1}^{p} a_{k,t}\xi_{t-k} = \mathbf{x}'_{t-1}\mathbf{a}_t, \tag{13}$$

тобто, в даному випадку маємо лінійний прогноз, як і для моделі авторегресії з невипадковими періодичними коефіцієнтами. Базовою характеристикою точності прогнозування для обох моделей є дисперсія (безумовна) похибки $\varepsilon_t = \xi_t - \hat{\xi}_t$. Для послідовності авторегресії з невипадковими коефіцієнтами $\mathbf{D}(\varepsilon_t) = \mathbf{D}(\varepsilon_t | \mathbf{x}'_{t-1})$. Однак для моделі RCA умовна дисперсія $\mathbf{D}(\varepsilon_t | \mathbf{x}'_{t-1}) = \mathbf{x}'_{t-1}\mathbf{R}_t\mathbf{x}_{t-1} + \sigma_t^2$, тобто умовну дисперсію похибки прогнозу (13) так само можна оцінювати на основі спостереження попередніх відліків $\mathbf{x}'_{t-1} = (\xi_{t-1}, \xi_{t-2}, ..., \xi_{t-p})$ прогнозованої послідовності.

На рис. 3 показано приклад реалізації погодинного електроспоживання (кВт·год) середнього підприємства.

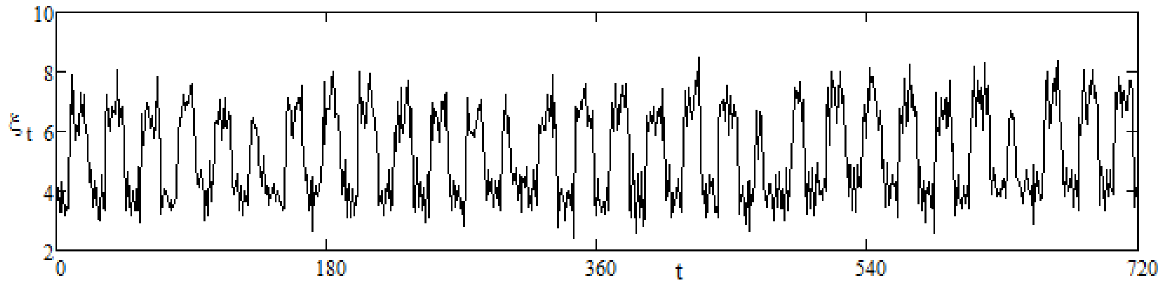

Рис. 3

Здійснено статистичне оцінювання характеристик моделі RCPAR десятого порядку ($p = 10$), період $L = 24$ год, кількість циклів у вибірці $q = 60$ (загальний обсяг вибірки $q \cdot 24 = 1440$).

На рис. 4, *а* представлено порівняння реальних даних електроспоживання ξ_t та його прогнозу (на один крок вперед) $\hat{\xi}_t$ протягом трьох діб. На рис. 4, *б* d_t − оцінка безумовної дисперсії похибки прогнозу та $d1_t$ − умовна дисперсія похибки прогнозу протягом доби.

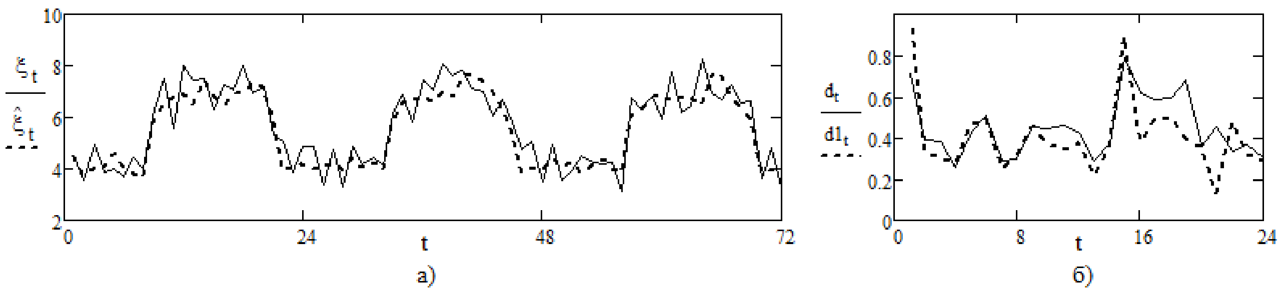

Рис. 4

У табл. 2 наведено характеристики точності прогнозування (для різних інтервалів упередження), які використовуються у багатьох роботах щодо прогнозування електроспоживання [14–16] (у таблиці ΔP, δP – максимальні абсолютна та відносна похибки відповідно, $\Delta P_{ск}$ – середньо квадратична похибка, MAPE – Mean Absolute Percentage Error (середня абсолютна відсоткова похибка)).

Таблиця 2

Величина інтервалу уперед-ження, год	ΔP, кВт·год	δP, %	$\Delta P_{ск}$, кВт·год	MAPE, %
1	1,14	22,55	0,51	8,91
6	1,59	33,05	0,71	11,30
11	1,8	41,38	0,82	12,68
16	1,95	32,58	0,80	11,84
21	1,87	34,75	0,68	9,80

Зауважимо, що у нашому випадку MAPE є більшою за відповідні характеристики прогнозування електроспоживання обласної енергосистеми та промислових енергоємних підприємств, наведені в [14] (MAPE складає від 2,9 до 8,4 відсотків), але меншою, ніж величини MAPE прогнозування електроспоживання окремих домогосподарств, наведені, наприклад, у [16] (MAPE для прогнозу на одну годину вперед складає від 10 до 23,5 відсотків). Це можна пояснити наявністю «ефекту агрегації електроспоживання» [15]. Тобто, чим вищим є сумарний рівень електроспоживання великої кількості споживачів, тим меншою є його варіабельність, і тим вищої точності (меншої MAPE) його прогнозування можна досягти. Отже, для коректного порівняння точності прогнозування різними методами потрібно буде провести відповідний аналіз їхнього функціонування для даних одного й того самого енергооб’єкта. Однак на даному етапі скористаємося нещодавно опублікованими результатами авторів [15], які описали ефект агрегації такою закономірністю: $\overline{\mathrm{MAPE}}(W)=\sqrt{\frac{\beta_0}{W^p}+\beta_1}$, де $\overline{\mathrm{MAPE}}(\cdot)$ – математичне сподівання MAPE, W – середній рівень електроспоживання, p, β_0, β_1 – параметри, які можна оцінити емпірично. У роботі [15] отримано оцінки параметрів p, β_0, β_1 (і відповідно оцінки залежностей $\overline{\mathrm{MAPE}}(W)$) при здійсненні прогнозування на одну годину вперед із використанням моделі сезонної авторегресії ковзної суми (SARMA), штучної нейронної мережі (ШНМ) та опорно-векторної регресії (Support Vector Regression (SVR)). Для підприємства, реалізацію електроспоживання якого показано на рис. 3, $W=5{,}26$ кВт$\cdot$год, при цьому оцінки $\overline{\mathrm{MAPE}}(W)$ для методів прогнозування, проаналізованих у [15], становлять від 15,52 до 30,52 відсотків, що перевищує відповідні значення, наведені в табл. 2.

Висновки. У роботі обгрунтовано метод оцінювання параметрів послідовності L-періодичної авторегресії з випадковими коефіцієнтами на основі її зображення у вигляді сукупності L стаціонарних та стаціонарно зв’язаних підпослідовностей та оцінювання характеристик кожної з них методом найменших квадратів. Також наведено результати комп’ютерного імітаційного моделювання такої послідовності другого порядку та оцінювання її періодичних характеристик. Результати моделювання, зокрема, показують, що реалізації вибіркових дисперсій досліджуваних оцінок зменшуються із зростанням обсягу вибірки, що є аргументом, який підтверджує їхню слушність. Розглянуті математичні моделі та методи можуть бути застосовані в задачах розробки та використання інформаційних технологій аналізу та оперативного прогнозування електроспоживання підприємств. У порівнянні з вказаними вище моделями SARMA, ШНМ, SVR, які використовуються для прогнозування електроспоживання, застосування моделі RCPAR дає змогу досягнути менших значень MAPE, крім того, дає можливість оцінювати умовну дисперсію похибки прогнозування.

1. Бабак С.В., Мыслович М.В., Сысак Р.М. Статистическая диагностика электротехнического оборудования: монография. К.: Институт электродинамики НАН Украины, 2015. 456 с.

2. Зварич В.М. Системи діагностики енергетичного обладнання на базі лінійних моделей авторегресії: автореф. дис. ... докт. техн. наук: 05.13.05 / Інститут електродинаміки НАН України. Київ. 2013. 35 с.

3. Марченко Б.Г., Щербак Л.Н. Линейные случайные процессы и их приложения. К.: Наукова думка, 1975. 143 с.

4. Млинко Б.Б., Фриз М.Є., Щербак Л.М. Методологія математичного моделювання стохастичних сигналів на основі умовних лінійних випадкових процесів. *Моделювання та інформаційні технології. Збірник наукових праць*. К.: Інститут проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України, 2016. Вип. 77. С. 20–25.

5. Фриз М.Є. Властивості умовних лінійних процесів та їх застосування в прикладних задачах математичного моделювання стохастичних сигналів. *Математичне та комп’ютерне моделювання. Серія: Технічні науки: збірник наукових праць*. 2012. Вип. 6. С. 228–238.

6. Pierre P.A. Central limit theorems for conditionally linear random processes. *SIAM Journal of Applied Mathematics*. 1971. Volume 20. Issue 3. Pp. 449–461. DOI: https://doi.org/10.1137/0120048

7. Gardner W.A., Napolitano A., Paura L. Cyclostationarity: Half a century of research. *Signal Processing*. Elsevier, 2006. No 86 (4). Pp. 639–697. DOI: https://doi.org/10.1016/j.sigpro.2005.06.016

8. Nicholls D.F., Quinn B.G. Random Coefficient Autoregressive Models: an Introduction. Lecture Notes in Statistics, 11. New York: Springer Verlag, 1983. 154 p.

9. Aknouche A., Guerbyenne H. Periodic stationarity of random coefficient periodic autoregressions. *Statistics and Probability Letters*. 2009. Volume 79. Issue 7. Pp. 990–996. DOI: https://doi.org/10.1016/j.spl.2008.12.012

10. Franses P.H., Paap R. Random-coefficient periodic autoregressions. *Statistica Neerlandica*. 2011. Volume 65. No 1. Pp. 101–115. DOI: 10.1111/j.1467-9574.2010.00477.x

11. Humeau S., Wijaya T.K., Vasirani M., Aberer K. Electricity load forecasting for residential customers: Exploiting aggregation and correlation between households. *Sustainable Internet and ICT for Sustainability (SustainIT)*. Palermo, Italy, 30-31 October 2013. Pp. 1–6. DOI: 10.1109/SustainIT.2013.6685208

12. Фриз М.Є. Умовні лінійні випадкові послідовності. *Збірник наукових праць Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України*. 2011. Вип. 60. С. 41–45.

13. Magnus J.R., Neudecker H. The elimination matrix: Some lemmas and applications. *SIAM Journal on Algebraic and Discrete Methods*. 1980. Volume 1. Issue 4. Pp. 422–449. DOI: https://doi.org/10.1137/0601049

14. Черненко П.О., Мартинюк О.В., Мірошник В.О. Моделювання та короткострокове прогнозування технологічної складової електричного навантаження обласної енергосистеми. *Технічна електродинаміка*. 2016. № 4. С. 68–70. DOI: https://doi.org/10.15407/techned2016.04.068

15. Sevlian R., Rajagopal R. A scaling law for short term load forecasting on varying levels of aggregation. *Electrical Power and Energy Systems*. 2018. Volume 98. Pp. 350–361. DOI: https://doi.org/10.1016/j.ijepes.2017.10.032

16. Rodrigues F., Cardeira C., Calado J.M.F. The Daily and Hourly Energy Consumption and Load Forecasting Using Artificial Neural Network Method: A Case Study Using a Set of 93 Households in Portugal. *Energy Procedia*. 2014. Volume 62. Pp. 220–229. DOI: https://doi.org/10.1016/j.egypro.2014.12.383

УДК 519.246.8

СТАТИСТИЧЕСКИЙ АНАЛИЗ ПЕРИОДИЧЕСКОЙ АВТОРЕГРЕССИИ СО СЛУЧАЙНЫМИ КОЭФФИЦИЕНТАМИ В ЗАДАЧАХ ОПЕРАТИВНОГО ПРОГНОЗИРОВАНИЯ ЭЛЕКТРОПОТРЕБЛЕНИЯ ПРЕДПРИЯТИЙ

М.Е. Фриз[1], канд.техн.наук, **Л.Н. Щербак**[2], докт.техн.наук
[1 –] **Тернопольский национальный технический университет имени Ивана Пулюя,**
ул. Руська, 56, Тернополь, 46001, Украина. E-mail: mykh.fryz@gmail.com
[2 –] **Киевский международный университет,**
ул. Львовская, 49, Киев, 03179, Украина. E-mail: prof_scherbak@ukr.net

Охарактеризованы свойства и использование периодических условных линейных случайных процессов (УЛСП) в задачах математического моделирования, анализа и оперативного прогнозирования электропотребления. Обоснована возможность осуществления статистического анализа УЛСП с дискретным временем с использованием модели авторегрессии со случайными коэффициентами. Такая модель является частным случаем УЛСП. Предложен метод оценивания параметров последовательности периодической авторегрессии со случайными коэффициентами, суть которого состоит в представлении исследуемой последовательности в виде совокупности L (где L – период) стационарных и стационарно связанных подпоследовательностей и применении к каждой из них двухэтапного метода наименьших квадратов для нахождения соответствующих оценок. Приведены результаты компьютерного имитационного моделирования, подтверждающие состоятельность предложенных оценок. Рассмотрен пример оперативного прогнозирования электропотребления организации, относящейся к классу средних и малых предприятий с использованием модели периодической авторегрессии со случайными коэффициентами. Библ. 16, рис. 4, табл. 2.

Ключевые слова: математическая модель, условный линейный случайный процесс, период, авторегрессия со случайными коэффициентами, оценивание параметров, имитационная модель, прогнозирование, электропотребление.

STATISTICAL ANALYSIS OF RANDOM COEFFICIENT PERIODIC AUTOREGRESSION AND ITS APPLICATION FOR SHORT-TERM ELECTRICITY CONSUMPTION FORECASTING

M. Fryz [1], **L. Scherbak** [2]
[1 –] **Ternopil Ivan Puluj National Technical University,**
str. Ruska, 56, Ternopil, 46001, Ukraine. E-mail: mykh.fryz@gmail.com
[2 –] **Kyiv International University,**
str. Lvivska, 49, Kyiv, 03179, Ukraine. E-mail: prof_scherbak@ukr.net

A conditional linear random process (CLRP) has been defined as the stochastic integral of a random function with respect to a process with independent increments. When the process with independent increments is Poisson then CLRP *represents the signal as a sum of a large amount of stochastically dependent impulses whose times of occurrence are the times of a Poisson process. For example, the electricity loads of the electrical power systems, also the processes of gas and water consumption, electrophysiological signals et al. can be modelled using CLRP. Moreover, the stochastic periodicity of the signals can be taken into account. A random coefficient autoregressive model has been shown to be a member of the class of discrete-time CLRP and suitable for estimation purposes. The main goal of the paper is to develop the procedure for the parameter estimation of random coefficient periodic autoregressive (RCPAR) model. The model has periodic parameters and consequently periodic probability distribution. The estimations have been obtained as a result of applying the least squares method to the set of L (where L is a period) stationary and jointly stationary subsequences of RCPAR model. The simulation results have been presented which confirm the consistency of the developed estimations, that is, the precision of the estimates increases with the increase in the sample size. The results of short-term electricity consumption forecasting of the enterprise (which belongs to the class of small and medium-sized) have been presented and analyzed using RCPAR model.* References 16, tables 2, figures 4.

Key words: mathematical model, conditional linear random process, period, random coefficient autoregression, parameter estimation, computer simulation, forecasting, electricity consumption.

1. Babak S.V., Myslovich M.V., Sysak R.M. Electrotechnical equipment statistical diagnostics: monograph. Kyiv: Institut elektrodinamiki NAN Ukrainy, 2015. 456 p. (Rus)

2. Zvarych V.M. Diagnostics systems of power engineering equipment on the linear autoregressive models basis: author's abstract of Dr. tech. sci. diss.: 05.13.05. Instytut elektrodynamiky NAN Ukrainy. Kyiv. 2013. 35 p. (Ukr)

3. Marchenko B.G., Scherbak L.N. Linear random processes and their applications. Kyiv: Naukova dumka, 1975. 143 p. (Rus)

4. Mlynko B.B., Fryz M.Ye., Scherbak L.M. Methodology of stochastic signal mathematical modelling using conditional linear random processes. *Modeliuvannia ta informatsiini tekhnologii. Zbirnyk naukovykh prats.* Kyiv: Instytut problem modeliuvannia v enerhetytsi im. G.Ye. Pukhova NAN Ukrainy, 2016. Issue 77. Pp. 20–25. (Ukr)

5. Fryz M.Ye. Properties of conditional linear processes and their application in the problems of stochastic signal mathematical modelling. *Matematychne ta kompiuterne modeliuvannia. Seria: Tekhnichni nauky: zbirnyk naukovykh prats*. 2012. Issue 6. Pp. 228–238. (Ukr)

6. Pierre P.A. Central limit theorems for conditionally linear random processes. *SIAM Journal of Applied Mathematics*. 1971. Volume 20. Issue 3. Pp. 449–461. DOI: https://doi.org/10.1137/0120048

7. Gardner W.A., Napolitano A., Paura L. Cyclostationarity: Half a century of research. *Signal Processing*. Elsevier, 2006. No 86 (4). Pp. 639–697. DOI: https://doi.org/10.1016/j.sigpro.2005.06.016

8. Nicholls D.F., Quinn B.G. Random Coefficient Autoregressive Models: an Introduction. Lecture Notes in Statistics, 11. New York: Springer Verlag, 1983. 154 p.

9. Aknouche A., Guerbyenne H. Periodic stationarity of random coefficient periodic autoregressions. *Statistics and Probability Letters*. 2009. Volume 79. Issue 7. Pp. 990–996. DOI: https://doi.org/10.1016/j.spl.2008.12.012

10. Franses P.H., Paap R. Random-coefficient periodic autoregressions. *Statistica Neerlandica*. 2011. Volume 65. No 1. Pp. 101–115. DOI: 10.1111/j.1467-9574.2010.00477.x

11. Humeau S., Wijaya T.K., Vasirani M., Aberer K. Electricity load forecasting for residential customers: Exploiting aggregation and correlation between households. *Sustainable Internet and ICT for Sustainability (SustainIT)*. Palermo, Italy, 30-31 October 2013. Pp. 1–6. DOI: 10.1109/SustainIT.2013.6685208

12. Fryz M.Ye. Conditional linear random sequences. *Zbirnyk naukovykh prats Instytutu problem modeliuvannia v enerhetytsi im. G.Ye. Pukhova NAN Ukrainy*. 2011. Issue 60. Pp. 41–45. (Ukr)

13. Magnus J.R., Neudecker H. The elimination matrix: Some lemmas and applications. *SIAM Journal on Algebraic and Discrete Methods*. 1980. Volume 1. Issue 4. Pp. 422–449. DOI: https://doi.org/10.1137/0601049

14. Chernenko P.O., Martyniuk O.V., Miroshnyk V.O. Modeling and short-term forecasting of technology component of electrical load of the regional electric power system. *Tekhnichna Elektrodynamika*. 2016. No 4. Pp. 68–70. (Ukr) DOI: https://doi.org/10.15407/techned2016.04.068

15. Sevlian R., Rajagopal R. A scaling law for short term load forecasting on varying levels of aggregation. *Electrical Power and Energy Systems.* 2018. Volume 98. Pp. 350 – 361.

DOI: https://doi.org/10.1016/j.ijepes.2017.10.032

16. Rodrigues F., Cardeira C., Calado J.M.F. The Daily and Hourly Energy Consumption and Load Forecasting Using Artificial Neural Network Method: A Case Study Using a Set of 93 Households in Portugal. *Energy Procedia*. 2014. Volume 62. Pp. 220–229. DOI: https://doi.org/10.1016/j.egypro.2014.12.383

Надійшла 03.04.2017
Остаточний варіант 28.01.2019