УДК 681.51

ПОБУДОВА ТА ВИКОРИСТАННЯ МОДЕЛЕЙ ГЕТЕРОСКЕДАСТИЧНИХ ПРОЦЕСІВ ДЛЯ МОДЕЛЮВАННЯ І ПРОГНОЗУВАННЯ ФІНАНСОВИХ РИЗИКІВ

О.Б. ДЕМКІВСЬКИЙ

Розглядається задача моделювання та прогнозування фінансових ризиків на основі гетероскедастичних моделей. Запропонована узагальнена методика побудови моделей такого класу. На конкретному прикладі доведена ефективність використання даної методики. Побудовано прогноз поведінки дисперсії вартості акції компанії УКРНАФТА.

Зміцнення і розширення ринкових принципів господарювання в Україні посилює невизначеність підстав для прийняття економічних рішень. За таких умов необхідно впроваджувати нові інформаційні технології та приймати обгрунтовані рішення, науково прогнозувати тенденції розвитку економічних процесів, правильно оцінювати рівень ризику. Досі відомі підходи до оцінювання ризику, наприклад,підхід на основі функції Байєса, портфельний підхід, функція вигідності та інші. Однак на сьогодні не існує методики оцінювання фінансового ризику на основі використання часових рядів. Тому в роботі пропонується підхід до оцінювання фінансового ризику на основі моделей часових рядів.

Одним із поширених підходів до визначення міри ризику є використання стандартного відхилення величини прибутку [1]. У такому випадку для прогнозування ризику необхідно побудувати модель, яка описує дисперсію прибутку. Процеси з постійною дисперсією називаються гомоскедастичними, а із змінною — гетероскедастичними.

Для визначення наявності гетероскедастичності існують тести: Уайта, Бройша–Пагана/Годфрі, Голдфельда–Квандта.

Найчастіше на практиці застосовують спрощений тест на гетероскедастичність, який складається з таких кроків.

1. Оцінити авторегресію $y(k) = a_0 + a_0 y(k-1) + \varepsilon(k)$ або більш високого порядку, наприклад, другого чи третього.

2. Побудувати ряд $\{\varepsilon^2(k)\}$, скориставшись залишками від оцінювання попередньої моделі.

3. Оцінити регресію

$$\varepsilon^2(k) = \alpha_0 + \alpha_1\left[0,4\,\varepsilon(k-1) + 0,3\,\varepsilon(k-2) + 0,2\,\varepsilon(k-3) + 0,1\,\varepsilon(k-4)\right].$$

4. Якщо коефіцієнт α_1 відмінний від нуля в статистичному розумінні, тобто є значущим, то отримана модель для $\varepsilon^2(k)$ описує гетероскедастичний процес. Оскільки в цій моделі оцінюються тільки коефіцієнти α_1 і α_0, а всі інші відомі (0, 4; 0, 3; 0, 2; 0, 1), то для визначення відмінності від нуля кое-

фіцієнта α_1 можна застосувати теорію перевірки гіпотези, тобто t-статистику для цього коефіцієнта. Обгрунтування вибору коефіцієнтів $0,4$; $0,3$; $0,2$; $0,1$ наведено в роботі [1].

Привабливість такого тесту полягає в простих обчисленнях і можливості застосування тієї ж самої теорії перевірки гіпотез, яка використовується в аналізі лінійних моделей.

Нами пропонується узагальнена методика побудови моделі гетероскедастичного процесу, який складається з п'яти кроків.

Крок 1. У разі необхідності здійснити попередню обробку експериментальних даних (нормування, логарифмування, заповнення пропусків даних) і застосувати до них тести на гетероскедастичність. Якщо процес містить тренд, то перед побудовою моделі необхідно його видалити. Для підвищення надійності тестування слід застосувати не менше двох тестів. Зокрема, досвід побудови моделей свідчить, що необхідно користуватись спрощеним тестом, який є досить наочним і відносно простим. Доволі часто візуальний аналіз даних дозволяє отримати суттєву інформацію про присутність гетероскедастичності. Водночас із візуальним аналізом корисно розглядати параметри описової характеристики, які полегшують визначення структури моделі.

Крок 2. Користуючись АКФ та ЧАКФ для експериментальних даних побудувати модель АР(p) або АРКС(p,q) для процесу $\{y(k)\}$ та обчислити ряд із квадратів залишків $\{\hat{\varepsilon}^2(k)\}$, де $\hat{\varepsilon}(k)=e(k)$. Обчислити вибіркову дисперсію $\hat{\sigma}^2_{\hat{\varepsilon}}$ збурення $\hat{\varepsilon}(k)$

$$\hat{\sigma}^2_{\hat{\varepsilon}}=\frac{1}{N-1}\sum_{k=1}^{N}\hat{\varepsilon}^2(k),$$

де N — число залишків після побудови моделі АР чи АРКС.

Крок 3. Обчислити і побудувати графік вибіркової автокореляційної функції для квадратів залишків

$$\rho(s)=\frac{\sum\limits_{k=s+1}^{N}[\hat{\varepsilon}^2(k)-\hat{\sigma}^2_{\varepsilon}][\hat{\varepsilon}^2(k-s)-\hat{\sigma}^2_{\varepsilon}]}{\sum\limits_{k=1}^{N}[\hat{\varepsilon}^2(k)-\hat{\sigma}^2_{\varepsilon}]}. \quad (1)$$

Якщо існують такі значення $\rho(s)$, які відрізняються від нуля в статистичному розумінні, то це свідчить про наявність процесу АРУГ або УАРУГ. Для того щоб переконатись у присутності гетероскедастичності, використовують Q-статистику Люнга–Бокса [2], яка обчислюється за виразом

$$Q=N(N+2)\sum_{i=1}^{n}\rho(i)/(N-i),$$

де $n=N/4$ (емпірично встановлене значення). Якщо значення $\hat{\varepsilon}^2(k)$ некорельовані, то Q-статистика повинна мати розподіл χ^2 з n ступенями свободи.

Крок 4. Побудувати модель УАРУГ (або іншу модифікацію)

$$h(k) = a_0 + \sum_{i=1}^{q} \alpha_i \varepsilon^2(k-i) + \sum_{i=1}^{p} \beta_i\, h(k-i)\,, \qquad (2)$$

використовуючи ряд значень $\hat{\varepsilon}^2(k)$. Якщо в цій моделі хоча б один із коефіцієнтів α_i, $i \in 1,...,q$ є значущим, то процес дійсно гетероскедастичний. Оскільки модель (2) описує залишки моделі з деяким наближенням, то загалом доцільно продовжити процес уточнення моделей, які описують вихідний процес в цілому. Тобто, можна уточнити початкову модель АР(p) чи АРКС(p,q). Це робиться на наступному кроці.

Крок 5. Скористатися моделлю (2), для того щоб отримати дійсні значення залишків, які описуються цією моделлю, тобто згенерувати ряд $\{\hat{\varepsilon}_1(k)\}$. Згенерувати ще один ряд $\{y_1(k)\}$, де $y_1(k) = y(k) - \hat{\varepsilon}_1(k)$. За допомогою отриманого ряду побудувати уточнену модель процесу типу АР(p) чи АРКС(p,q). У разі потреби процес уточнення моделей можна продовжити.

Як приклад застосування наведеної методики розглянемо ряд $\varepsilon^2(k)$ для моделі перших різниць щоденної вартості акцій компанії УКРНАФТА за 2001 р. Динаміка ряду подана на рис. 1.

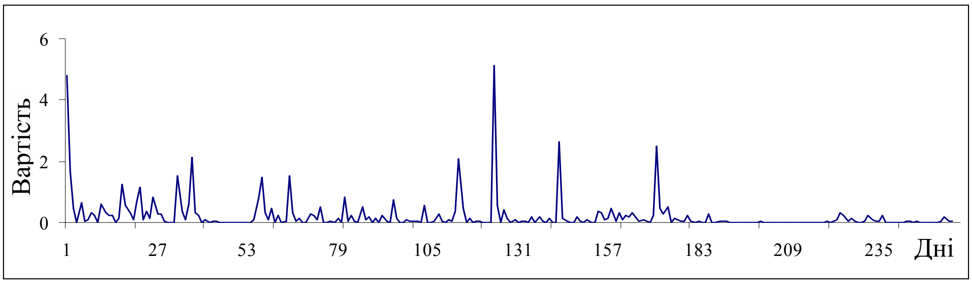

Рис. 1. Динаміка ряду $\varepsilon^2(k)$

Відповідно до методики обчислимо і побудуємо графіки вибіркової автокореляційної функції для квадратів залишків (рис. 2, 3).

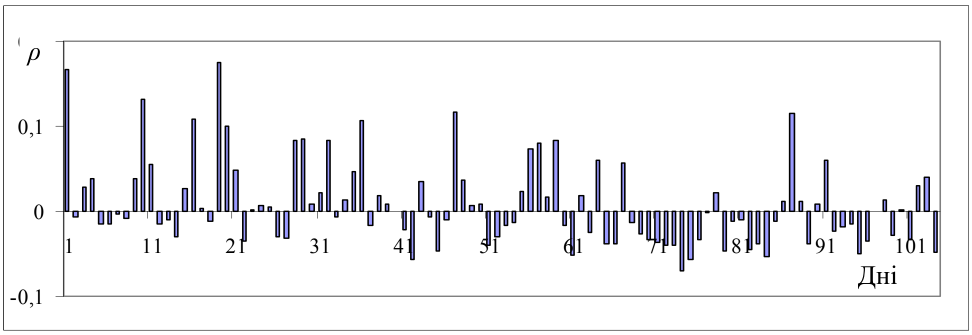

Рис. 2. Автокореляційна функція ряду (104 значення)

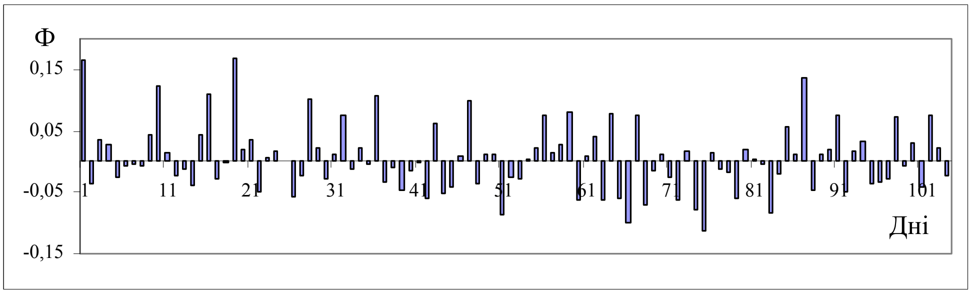

Рис. 3. Часткова автокореляційна функція ряду (104 значення)

У результаті дослідження АКФ і ЧАКФ робимо такі висновки.

1. Значення коефіцієнтів АКФ: $\rho_1 = 0,166; \rho_2 = -0,006; \rho_3 = 0,028$; $\rho_9 = 0,038$; $\rho_{10} = 0,132$.

2. Значення коефіцієнтів ЧАКФ: $\Phi_{1,1} = 0{,}166$; $\Phi_{2,2} = -0{,}035$; $\Phi_{3,3} = = 0{,}036$; $\Phi_{9,9} = 0{,}043$; $\Phi_{10,10} = 0{,}122$.

3. Для математичного опису процесу слід розглянути моделі АР(1), АР(2), АР(3), АР(4). Найвірогідніші номери запізнювань, що входять до складу моделі: 1, 2, 3, 9, 10. У табл. 1 наведені варіанти оцінювання декількох можливих структур регресійної моделі.

Таблиця 1. Варіанти оцінювання дисперсії щоденної вартості акцій компанії УКРНАФТА за 2001 р.

АР / a_i	(1)	(1, 2)	(1, 2, 3)	(1, 2, 3, 9)	(1, 2, 3, 9, 10)
a_0	0,1652 (5,1101)	0,1777 (5,2338)	0,1738 (4,8391)	0,1606 (4,1112)	0,1350 (3,3577)
a_1	0,1666 (3,1342)	0,1199 (1,9116)	0,1180 (1,8648)	0,1230 (1,9169)	0,1164 (1,8261)
a_2		-0,0263 (-0,4872)	-0,0495 (-0,7817)	-0,0458 (-0,7101)	-0,0429 (-0,6701)
a_3			0,0369 (0,6801)	0,0507 (0,7907)	0,0504 (0,7913)
a_4				0,0397 (0,7376)	0,0403 (0,6384)
a_5					0,1218 (2,2372)
RSS	58,5397	57,8552	57,6733	57,1763	55,8335
AIC	1,3820	1,3818	1,3909	1,4151	1,4037
BSC	1,4098	1,4235	1,4468	1,4862	1,4892
DW	2,0848	2,0029	1,9978	2,0121	2,0110
F	0,0389	0,0146	0,0164	0,0201	0,0440
R^2	0,0374	0,0144	0,0161	0,0197	0,0421

Аналіз отриманих результатів дозволяє зробити такі висновки.

З табл. 1 видно, що найкраща модель, яка описує дисперсію з обчислених варіантів є $\text{АР}(1,2)$, оскільки для неї $DW = 2{,}0029$; $R^2 =$ 0,0144;

$AIC = 1,4235$. Таким чином, модель, що адекватно описує процес, можемо записати як

$$\varepsilon_1^2(k) = 0,1777 + 0,1199\,\varepsilon_1^2(k-1) - 0,0263\,\varepsilon_1^2(k-2)\,.$$

Значення $RSS = 57,8552$ показує, що модель можна покращити за рахунок введення ковзного середнього.

Т а б л и ц я 2 . Варіанти оцінювання моделей з ковзним середнім

a_i \ АР	(1)	(1, 2)	(1, 2, 3)	(1, 2, 3, 9)	(1, 2, 3, 9, 10)
B_0	0,1652 (2,49E+15)	0,1777 (1,98E+14)	0,1738 (1,48E+15)	0,1606 (2,56E+14)	0,1350 (3,54E+15)
B_1	0,1666 (5,59E+14)	0,1199 (2,32E+14)	0,1180 (1,84E+14)	0,1230 (3,56E+14)	0,1164 (5,82E+14)
B_2		-0,0263 (-4,33E+13)	-0,0495 (-6,19E+14)	-0,0458 (-9,58E+13)	-0,0429 (-1,11E+15)
B_3			0,0369 (9,27E+14)	0,0507 (1,86E+14)	0,0504 (1,62E+15)
β_4				0,0397 (1,72E+14)	0,0403 (1,34E+15)
β_5					0,1218 (3,75E+15)
β_6	4,87E-15 (16,0641)	8,13E-14 (15,7368)	-8,31E-15 (-12,9103)	-5,35E-14 (-15,4494)	-2,78E-15 (-13,7294)
RSS	5,06E-29	3,63E-28	1,09E-29	7,65E-28	1,29E-29
AIC	-∞	-∞	-∞	-∞	-∞
BSC	-∞	-∞	-∞	-∞	-∞
DW	1,7425	1,6296	2,1976	1,8840	1,5241
F	∞	∞	∞	∞	∞
R^2	1,0000	1,0000	1,0000	1,0000	1,0000

Аналіз отриманих результатів дозволяє зробити такі висновки.

Адекватна модель (табл. 2) обов’язково повинна мати ковзне середнє. З обчислених варіантів моделей кращою є АРКС$((1,2,3),(1))$, оскільки для неї $DW = 2,1976$; $R^2 = 1$; $AIC = -\infty$. Таким чином, модель, що адекватно описує дисперсію, можемо записати як

$$\varepsilon_1^2(k) = 0,1738 + 0,1180\,\varepsilon_1^2(k-1) - 0,0495\,\varepsilon_1^2(k-2) + 0,0369\,\varepsilon_1^2(k-3) + \varepsilon_1^2(k)\,.$$

Отриману модель використаємо для побудови прогнозу поведінки дисперсії ряду на основі розв’язку різницевого рівняння.

1. Знайдемо однорідний розв’язок: $\alpha^3 - 0,118\alpha^2 + 0,0495\alpha - 0,0369 = 0$.

Оскільки дане рівняння має один дійсний корінь ($a_1 = -0,4311$) та два комплексні, повний розв’язок має вигляд

$$\varepsilon_1^2(k) = y^{pd} + A\alpha^k + \beta_1 r^k \cos(k\theta + \beta_2) + y^{ps}\,,$$

де $r = \sqrt{a_2'}$, а $\cos(\theta) = \dfrac{a_1'}{2\sqrt{a_2'}}$.

Значення a_1' та a_2' знаходяться так: з рівняння $\alpha^3 - a_1\alpha^2 - a_2\alpha - a_3 = 0$, знайшовши значення дійсного кореня α_1, можемо записати: $(\alpha - \alpha_1)(\alpha^2 - a_1'\alpha - a_2') = 0$. Розкривши дужки, запишемо a_1' та a_2' через значення a_1 та a_2.

$$\alpha^3 + a_1'\alpha^2 + a_2'\alpha - \alpha^2\alpha_1 - \alpha a_1'\alpha_1 - a_2'\alpha_1 = 0,$$

$$\alpha^3 + (a_1' - \alpha_1)\alpha^2 + (a_2' - a_1'\alpha_1) - \alpha_1 - a_2'\alpha_1 = 0,$$

$$a_1' = \alpha_1 + a_1 = -0,5491,$$

$$a_2' = a_1^2 + a_1 a_1 + a_2 = -0,2862.$$

Звідси маємо $r = 0,535$, $\theta = 2,11$.

2. Частковий розв'язок для детермінованої частини.

$$y^{pd} = 0,1738 - 0,1180 y^{pd} + 0,0495 y^{pd} - 0,0369 y^{pd}$$

або

$$y^{pd} = \frac{0,1738}{1 + 0,1180 - 0,495 + 0,0369} = 0,2634 .$$

3. Частковий розв'язок для стохастичної частини.

$$y^{ps} = \sum_{i=0}^{\infty} \beta_i \varepsilon(k-i), \text{ де } \beta_i = a_1^i,\ i = 0,1,2,\dots .$$

4. Повний розв'язок.

$$\varepsilon_1^2(k) = y^{pd} + A\alpha^k + y^{ps} .$$

Для того щоб знайти значення невідомої констант A, β_1, β_2, використаємо початкові умови $\varepsilon_1^2(0) = 4,809$, $\varepsilon_1^2(1) = 1,659$, $\varepsilon_1^2(2) = 0,450$. В результаті отримаємо

$$4,809 = 0,2634 + A(-0,4311)^0 + \beta_1 0,535^0 \cos(2,11\cdot 0 + \beta_2),$$

$$1,659 = 0,2634 + A(-0,4311)^1 + \beta_1 0,535^1 \cos(2,11\cdot 1 + \beta_2),$$

$$0,450 = 0,2634 + A(-0,4311)^2 + \beta_1 0,535^2 \cos(2,11\cdot 2 + \beta_2).$$

Знайшовши розв'язок системи наведених рівнянь,маємо $A = 6,6272$, $\beta_1 = -3,8771$, $\beta_2 = 1,0041$.

Отже, повний розв'язок має такий вигляд:

$$\varepsilon_1^2(k) = 0,2634 + 6,6272(-0,4311)^k - 3,8771\cdot 0,535^k \cos(k\cdot 2,11 + 1,0041) +$$

$$+ \sum_{i=0}^{\infty} (-0,8513)^i \varepsilon(k-i).$$

Отриманий розв'язок свідчить про присутність гармонічного коливального процесу, який відповідає реальним коливанням цін акцій компанії УКРНАФТА. Загалом розв'язок має збіжний характер, оскільки $\alpha = -0,4311$, $r = 0,535$, тобто, обидва значення є меншими від одиниці за модулем.

Таким чином, прогнозоване значення математичного сподівання дисперсії на s періодів дискретизації на основі отриманого розв'язку можна записати як

$$\varepsilon_1^2(k+s) = 0,2634 + \left[y(k) + 0,28\right]/0,1858\,(-0,4311)^k -$$
$$-\left[y(k-1) + 6,0227\right]/0,5348 \cdot 0,535^s \cos\left[s \cdot 2,11 -\right.$$
$$\left.- \arccos\left(y(k-2) - 6,8906\right)/3,8771\right. + \sum_{i=0}^{s-1}(-0,8513)^i\,\varepsilon(k+s-i)\,.$$

Для визначення точності прогнозу існує багато методів, які можна застосувати для оцінки прогнозуючих моделей. При цьому найуживаніший критерій — середньоквадратична похибка або її квадратний корінь. Цей критерій являє собою суму квадратичних відхилень прогнозованих значень від досліджуваних.

Корінь з середньоквадратичної похибки, який використовує N прогнозованих значень, має форму

$$RMSE\,(\hat{\theta}_{i\bullet}^2) = \sqrt{\sum_{t=1}^{N}\left(\hat{\theta}_{i,t}^2 - \theta_{i,t}^2\right)^2}\,.$$

Інший спосіб вимірювання працездатності прогнозуючої моделі базується на модулі відносного відхилення від істинного значення.

Усереднений модуль відносної похибки обчислюється за формулою

$$MAPE\,(\hat{\theta}_{i\bullet}^2) = \sum_{t=1}^{N}\frac{\left|\hat{\theta}_{i,t}^2 - \theta_{i,t}^2\right|}{\theta_{i,t}^2}\,.$$

Міра стійкості до відхилень від звичайного стану виражена за допомогою медіани квадратичної похибки

$$MedSE\,(\hat{\theta}_{i\bullet}^2) = Median\left(\hat{\theta}_{i,t}^2 - \theta_{i,t}^2\right)^2.$$

Ці три критерії порівнювалися між собою з використанням індексу працездатності, так званого Savage–Niehans правила (Саваджа–Ньюханса), який був узятий з теорії прийняття рішень

$$Perf_i = \sum_{j=1}^{n}\frac{EC_i - \min_i(EC_i)}{\min_i(EC_i)}\,,$$

де EC — один з критеріїв, описаних вище; n означає кількість акцій, для яких були отримані прогнози. Він може інтерпретуватися як відносна втрата точності однієї з моделей порівняно з тією моделлю, яка постфактум виявилася кращою для акції j.

Врешті-решт, зручним методом є регресування досліджуваних дисперсій за прогнозованими, тобто запуск регресії

$$\theta_{i,t}^2 = \upsilon + \omega \hat{\theta}_{i,t}^2 + \phi_{i,t} \,.$$

А перевірка, чи є константа υ рівною нулю та чи наближення коефіцієнта ω дорівнює 1, гарантує неупередженість прогнозу.

Простими та наочними критеріями якості прогнозу є максимальне та мінімальне відхилення від істинного значення ряду як абсолютне, так і в процентах (табл. 3).

На рис. 4 наведено результати прогнозування поведінки ряду, який описує дисперсію акцій компанії УКРНАФТА.

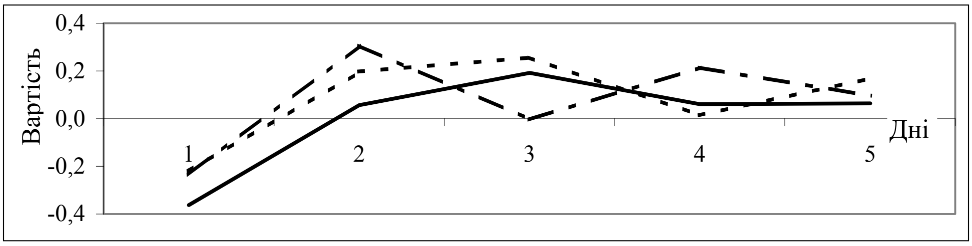

Рис. 4. Прогнозування поведінки дисперсій акцій компанії УКРНАФТА на 5 кроків:
—— — істинні значення;
········ — прогноз на основі прогнозуючої функції;
— · - — прогноз на основі повного розв'язку

Т а б л и ц я 3 . Статистичні характеристики наведеного прогнозу

Метод прогнозування	Максимальне відхилення		Мінімальне відхилення		Сума квадратів похибок
	Абсолютне	В, %	Абсолютне	В, %	
На основі прогнозуючої функції	0,1418	39,19	0,0465	6,95	0,1068
На основі повного розв'язку	0,2494	140,24	0,0316	9,47	0,1682

Таким чином, запропонована методика ітеративної побудови моделей гетероскедастичних процесів на основі часових рядів дає можливість отримати адекватну модель для опису динаміки поведінки дисперсії та побудувати на її основі прогноз поведінки ряду. Ця методика та моделі, описані на її основі, використані при проектуванні та реалізації систем підтримки прийняття рішень для аналізу, моделювання та прогнозування динаміки фінансово-економічних процесів [3]. Відкрита архітектура системи дозволяє досить просто розширювати її функціональні можливості шляхом введення нових моделей та алгоритмів прогнозування.

ЛІТЕРАТУРА

1. *Bollerslav T.* Generalised Autoregresiv Conditional Heteroscedasticity // J. of Econometrics. — 1986. — **31**, № 2. — P. 307–327.
2. *Enders W.* Applied econometric time series. — New York, Wiley & Sons, 1994. — 433 p.
3. *Демківський О.Б.* Побудова системи підтримки рішення при аналізі та прогнозуванні динаміки економічних змін // Зб. наук. пр. Ін-т проблем моделювання в енергетиці НАНУ. — Вип. 14. — Київ. — 2002. — С. 66–71.

Надійшла 10.12.2002