



УДК 519.766.4

ПРОГНОЗУВАННЯ ОБ'ЄМІВ ІМПОРТУ ПРИРОДНОГО ГАЗУ В УКРАЇНУ З ВИКОРИСТАННЯМ ФІЛЬТРА КАЛМАНА

Караюз І.В., Бідюк П.І.

Національний технічний університет України «Київський політехнічний інститут»

Робота присвячена розв'язанню задачі короткострокового прогнозування процесу імпорту природного газу в Україну на основі регресійних моделей і адаптивного фільтра Калмана. Мета роботи: побудова адекватної моделі для прогнозування процесу імпорту природного газу і застосування адаптивного фільтра Калмана для отримання оптимальних оцінок стану досліджуваного процесу. Для оцінювання стану процесу імпорту природного газу застосовано різні варіанти оптимального адаптивного фільтра Калмана, які передбачають повторне обчислення коваріації похибок вимірів за допомогою рухомих вікон даних. Основні результати роботи: регресійні моделі процесу імпорту природного газу та оцінки короткострокових прогнозів, отримані на основі адаптивного фільтра Калмана з ковзним вікном для обчислення коваріації похибок вимірів. Таким чином, на основі статистичних даних побудовано адекватну модель досліджуваного процесу із застосуванням адаптивного оптимального фільтра Калмана, який забезпечив обчислення високоякісних оцінок прогнозів для ковзного вікна розмірністю три виміри. Ключові слова: регресійна модель, адаптивний фільтр Калмана, імпорт природного газу, короткострокове прогнозування.

Вступ. Прогнозування розвитку макроекономічних процесів – одна із самих поширених задач економетричного аналізу [1, 2]. Для розв'язання цієї задачі розроблено множини методів, які забезпечують отримання високоякісних коротко- та середньострокових прогнозів для стаціонарних і нелінійних нестационарних процесів [3, 4]. Задачі прогнозування особливо актуальні для економік, що розвиваються, оскільки макроекономічні процеси у таких економіках, як правило, протікають в умовах наявності багатьох збурюючих впливів, мають високу динаміку, тренди, нелінійності різних типів, скачкоподібні переходи і т. ін. Це зумовлює необхідність застосування спеціальних методів оцінювання станів досліджуваних процесів з метою врахування випадкових впливів.

Робота присвячена актуальній задачі: короткостроковому прогнозуванню макроекономічного процесу – об'єму імпорту природного газу в Україну. Для розв'язання цієї задачі використано дані Державної служби статистики України (Держстат) [5], а саме форми державного статистичного спостереження «Експорт – імпорт окремих видів товарів за країнами світу», форма містить статистичні похибки (похибки вимірів). Отже використання оптимальних фільтрів є обґрунтованим та необхідним. Виконано порівняльний аналіз результатів роботи фільтра Калмана (ФК), адаптивного фільтра Калмана (АФК) [6] та адаптивного фільтра Калмана із вікном (АФКВ).

Мета. Побудова адекватної моделі для прогнозування процесу імпорту природного газу в Україну і застосування адаптивного фільтра Калмана для отримання оптимальних оцінок стану досліджуваного процесу.

Постановка задачі:

1. Побудувати математичну модель для опису об'ємів імпорту природного газу в Україну на основі наявних статистичних даних з метою обчислення оцінок короткострокових прогнозів. Перевірити гіпотезу щодо нормального закону розподілу похибок моделі.

2. Застосувати ФК, АФК та АФКВ для оптимального оцінювання стану – об'ємів імпорту природного газу і короткострокового прогнозування.

3. Виконати порівняльний аналіз оцінок прогнозів отриманих за побудованою моделлю та вказаних фільтрів.

Модель процесу. Досить часто прийнятну якість оцінок прогнозів можна отримати за допомогою моделей авторегресії, авторегресії з ковзним середнім або



авторегресії з інтегрованим ковзним середнім. Розглянемо можливість побудови моделі у формі розширеної авторегресії такого вигляду:

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=1}^q b_j x(k-j) + \varepsilon(k),$$

де p – порядок авторегресії; q – порядок регресії; $y(k)$ – вимір у момент k ; $x(k)$ – вимір регресора в момент k ; a_i – i -й параметр авторегресійної частини моделі; b_j – j -й параметр регресійної частини моделі; $\varepsilon(k)$ – випадкова нормально розподілена величина. Такі моделі дають можливість будувати функції прогнозування на довільну кількість кроків з використанням умовного математичного сподівання.

Алгоритм фільтрації лінійної системи. Враховуючи наявність впливу на макроекономічні процеси перехідних економік великої кількості збурень, необхідно застосувати адаптивний алгоритм оптимальної фільтрації, наприклад, фільтр Калмана. Послідовність виконання цієї процедури для скалярного випадку наведена нижче [7].

Крок 1. Математична модель лінійної динамічної системи:

$$x(k) = F(k)x(k-1) + C(k)u(k-1) + w(k-1);$$

де $x(k) = y(k)$ – вектор (змінна) стану процесу; $F(k) = \alpha_1$ – коефіцієнт динаміки; $C(k)u(k-1) = \alpha_0 + \beta_1 x(k-1)$ – врахування керуючого впливу (регресора); $w(k) \sim N(0, Q(k))$ – випадкові збурення стану; $z(k) = H(k)x(k) + v(k)$ – рівняння вимірів; $H(k) = 1$ – коефіцієнт чутливості вимірів; $v(k) \sim N(0, R(k))$ – шум датчика; початкові умови: $E\{x_0\} = \hat{x}_0$, $E\{\hat{x}_0, \hat{x}_0^T\} = P_0 = P'_0$; $P(k)$ – апостеріорна коваріаційна матриця похибок оцінок вектора стану; $P'(k)$ – апріорна коваріаційна матриця похибок оцінок вектора стану; $K(k)$ – коефіцієнт фільтра Калмана; початкові коваріації:

$$x(k) = F(k)x(k-1) + C(k)u(k-1) + w(k-1);$$

$$E\{w(k), v^T(k)\} = 0, \quad E\{w(k), x_0^T(0)\} = 0, \quad E\{v(k), x_0^T\} = 0.$$

Крок 2. Екстраполяція оцінки стану:

$$\hat{x}(k) = F(k)\hat{x}(k-1) + C(k)u(k-1) = \alpha_1 y(k-1) + \alpha_0 + \beta_1 x(k-1).$$

Крок 3. Екстраполяція коваріації похибок оцінок:

$$P'(k) = F(k)P(k-1)F^T(k) + Q(k-1) = \alpha_1^2 P(k-1) + Q(k-1).$$

Крок 4. Скалярний коефіцієнт фільтра: для адаптивних фільтрів будемо перераховувати коваріацію похибок вимірів (шуму) як середньоквадратичне відхилення похибок фільтру: $R(k) = \sqrt{E\{\hat{x}_0, \hat{x}_0^T\}}$. Для АФКВ використовуються виміри з вікна даних вибраного розміру.

$$K(k) = P'(k)H^T(k)\{H(k)P'(k)H^T(k) + R(k)\}^{-1} = \frac{P'(k)}{P'(k) + R(k)}.$$

Крок 5. Перерахунок оцінки стану з урахуванням виміру $z(k)$:

$$\hat{x}(k) = \hat{x}(k) + K(k)[z(k) - H(k)\hat{x}(k)].$$



Крок 6. Перерахунок коваріації похибок оцінок:

$$P(k) = [I - K(k)H(k)]P'(k).$$

Крок 7. Перейти на крок 2.

Критерії вибору кращих моделей. Для вибору кращої моделі з множини оцінених кандидатів використано такі критерії якості:

- середньоквадратична похибка:

$$e(i) = \frac{1}{N} \sum_{i=1}^N [y(i) - \hat{y}(i)]^2;$$

- критерій Дарбіна-Уотсона:

$$DW = 2 - 2\rho;$$

де ρ – коефіцієнт корельованості похибок моделі між собою.

Статистики для перевірки якості оцінок прогнозів:

- середньоквадратична похибка (СКП):

$$\text{СКП} = \frac{1}{N} \sum_{i=1}^N [y(i) - \hat{y}(i)]^2;$$

- середня абсолютна похибка прогнозу (САПП):

$$\text{САПП} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y(i) - \hat{y}(i)}{y(i)} \right| * 100;$$

- коефіцієнт Тейла:

$$U = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (y(i) - \hat{y}(i))^2}}{\sqrt{\frac{1}{N} \sum_{i=1}^N y(i)^2 + \frac{1}{N} \sum_{i=1}^N \hat{y}(i)^2}}.$$

Статистичні дані стосовно імпорту газу в Україну (млн. м³), використані для побудови моделі, наведено на рис. 1.

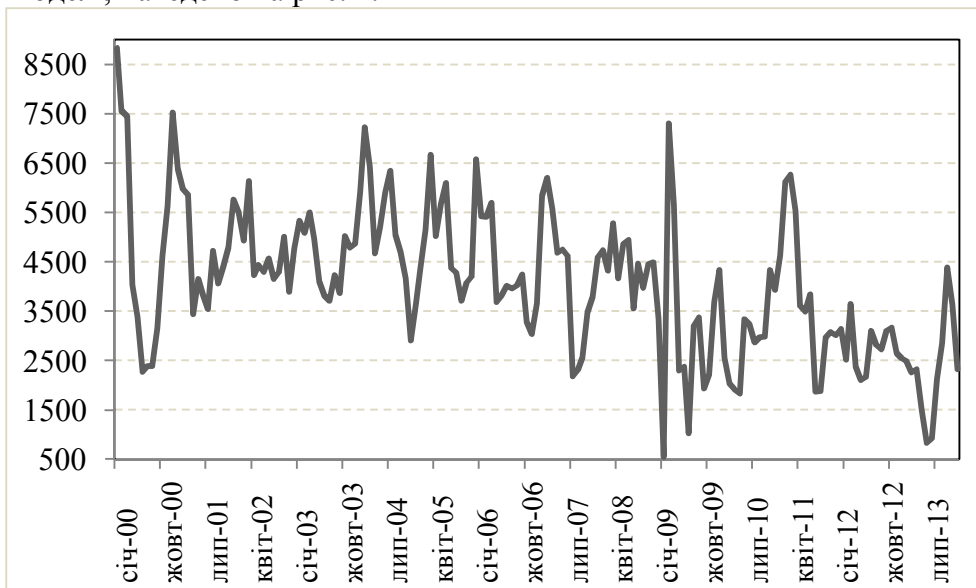


Рисунок 1 – Імпорт газу (по осі ординат у млн. м³)



Побудуємо модель, яка складається з авторегресії 1-го порядку та регресії від споживання:

$$y(k) = a_0 + a_1 y(k-1) + b_1 x(k-1) + \varepsilon(k),$$

де $y(k)$ – об’єм імпорту газу; $x(k)$ – об’єм споживання газу. Після оцінювання параметрів моделі отримаємо таке рівняння:

$$y(k) = 1211,1 + 0,6 \cdot y(k-1) + 0,1 \cdot x(k-1).$$

Перевіримо закон розподілу похибки. Будуємо гістограму (рис. 2) та, використовуючи критерій згоди, перевіряємо гіпотезу, що розподіл є нормальним:

$$\sum_{k=1}^r \frac{(n_k - np_k)^2}{np_k} = 5,94 < 7,8 = \chi_{\alpha,3}^2 = \chi_{0.05,3}^2 - \text{гіпотеза приймається.}$$

У наведеному виразі: n – кількість вимірів; r – кількість інтервалів; n_k – кількість вимірів, що потрапили у відповідний інтервал; p_k – ймовірність потрапляння у відповідний інтервал за нормальним законом; α – рівень значущості.

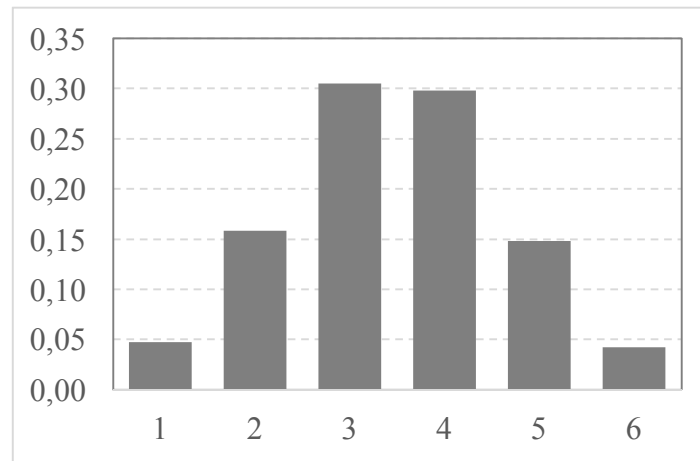


Рисунок 2 – Гістограма похибки моделі імпорту газу

Якість обраної моделі характеризується такими статистиками: $R^2 = 0,51$; $e(i) = 994916$; $DW = 1,94$. Якість однокрокового прогнозу характеризують статистики, наведені у табл. 1. Мінімальне значення середньої абсолютної похибки у процентах склало 8,95 % при використанні адаптивного фільтра з вікном даних для обчислення коваріацій похибок вимірів шириною три значення. Найбільше значення САПП отримано для адаптивного фільтра, описаного в [6].

Таблиця 1 – Якість прогнозу

	ФК	АФК	АФКВ 20	АФКВ 10	АФКВ 5	АФКВ 3
СКП	254053	326009	278887	256746	213547	180396
САПП	13,10	15,47	13,74	12,52	10,15	8,95
U	0,08	0,10	0,09	0,08	0,08	0,07
K	0,54	0,46	0,52	0,56	0,65	0,72

На рис. 3 наведено залежності значень САПП та коефіцієнта АФК від розміру вікна даних, використаних для адаптації алгоритму оптимальної фільтрації. Зростання коефіцієнта фільтра при зменшенні розміру вікна свідчить про зменшення значення відповідної коваріації, яка розраховується за даними.

В табл. 2 наведено оцінки прогнозів на 1–3 кроки з використанням різних алгоритмів оптимального оцінювання стану. Очевидно, що прогноз на три кроки незадовільний, за винятком оцінки, обчисленої за АФКВ 3. Для підвищення якості оцінок у модель необхідно ввести складову, яка описує тренд процесу.

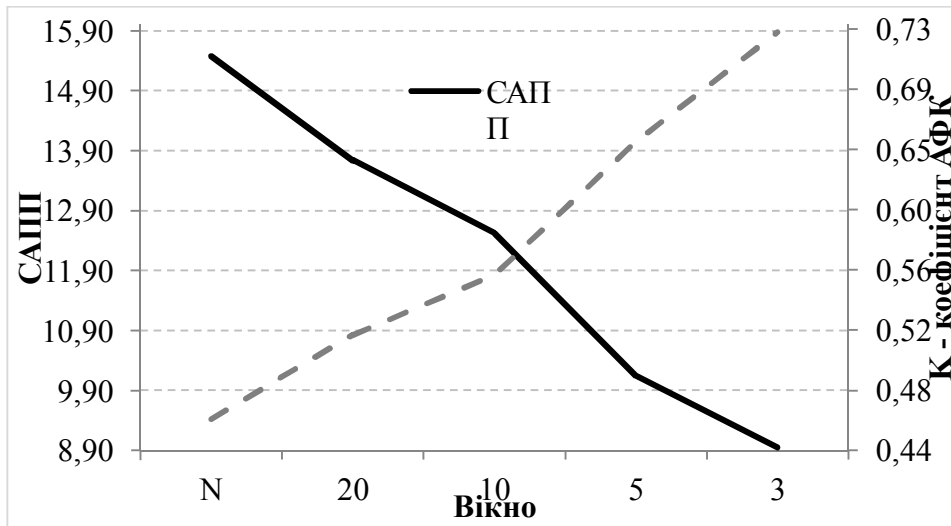


Рисунок 3 – Залежність САІП та коефіцієнту АФК від розміру вікна

Таблиця 2 – Прогноз на 3 кроки

	Грудень 2013		Січень 2014		Лютий 2014	
	прогноз	%	прогноз	%	прогноз	%
Дані	2303		2471		1649	
ФК	2787	-21,0	2898	-17,3	2510	-52,2
АФК	2901	-25,9	3009	-21,8	2696	-63,5
АФКВ 3	2725	-18,3	2760	-11,7	1895	-14,9
АФКВ 5	2525	-9,6	2695	-9,1	2158	-30,9
АФКВ 10	2722	-18,2	2838	-14,9	2410	-46,1
АФКВ 20	2754	-19,6	2853	-15,5	2419	-46,7

Для підвищення якості оцінок прогнозів у модель необхідно ввести складову, яка описує тренд процесу:

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=1}^q b_j x(k-j) + \sum_{l=1}^m c_l k^l + \varepsilon(k).$$

Після оцінювання параметрів моделі отримаємо таке рівняння:

$$y(k) = 1540,9 + 0,4 \cdot y(k-1) + 0,1 \cdot x(k-1) + 18,4 \cdot k - 0,3 \cdot k^2 + 0,001 \cdot k^3.$$

Перевіримо закон розподілу похибки. Будуємо гістограму (рис. 4) та, використовуючи критерій згоди, перевіряємо гіпотезу, що розподіл нормальний:

$$\sum_{k=1}^r \frac{(n_k - np_k)^2}{np_k} = 6,1 < 7,8 = \chi_{\alpha,3}^2 = \chi_{0,05,3}^2 - \text{гіпотеза приймається.}$$

Тут: n – кількість вимірів; r – кількість інтервалів; n_k – кількість вимірів, що потрапили в відповідний інтервал; p_k – ймовірність потрапляння в відповідний інтервал за нормальним законом; α – рівень значущості.

Якість обраної моделі характеризується такими статистиками: $R^2 = 0,59$; $e(i) = 823443$; $DW = 1,9$. Якість однокрокового прогнозу характеризують статистики, наведені у табл. 3. Мінімальне значення середньої абсолютної похибки у процентах склало 6,83 % при використанні адаптивного фільтра з вікном даних для обчислення коваріацій похибок вимірів шириною три значення. Найбільше значення САІП отримано для адаптивного фільтра, описаного в [6].

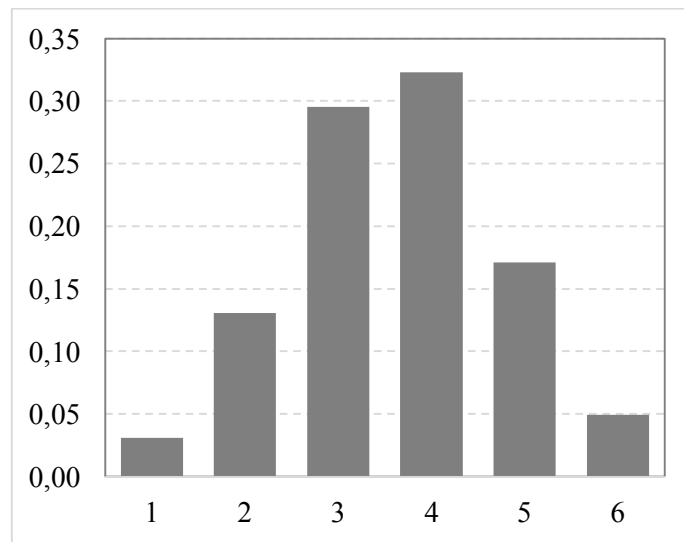


Рисунок 4 – Гістограма похибки моделі імпорту газу

Таблиця 3 – Якість прогнозу

	<i>ФК</i>	<i>АФК</i>	<i>АФКВ 20</i>	<i>АФКВ 10</i>	<i>АФКВ 5</i>	<i>АФКВ 3</i>
СКП	221372	240682	232166	211167	156820	128024
САПП	11,96	13,00	12,41	11,32	8,72	6,83
U	0,08	0,08	0,08	0,08	0,07	0,06
K	0,52	0,48	0,51	0,56	0,67	0,77

У табл. 3 наведено залежності значень САПП та коефіцієнта АФК від розміру вікна даних, використаних для адаптації алгоритму оптимальної фільтрації. Зростання коефіцієнта фільтра при зменшенні розміру вікна свідчить про зменшення значення відповідної коваріації, яка розраховується за даними.

В табл. 4 наведено оцінки прогнозів на 1–3 кроки з використанням різних алгоритмів оптимального оцінювання стану і другої моделі, яка враховує тренд.

Таблиця 4 – Прогноз на 3 кроки

	Грудень 2013		Січень 2014		Лютий 2014	
	прогноз	%	прогноз	%	прогноз	%
Дані	2303		2471		1649	
ФК	2585	-12,3	2699	-9,2	2224	-34,8
АФК	2619	-13,7	2729	-10,5	2288	-38,7
АФКВ 20	2576	-11,9	2689	-8,8	2202	-33,6
АФКВ 10	2563	-11,3	2677	-8,4	2178	-32,1
АФКВ 5	2478	-7,6	2624	-6,2	2077	-25,9
АФКВ 3	2384	-3,5	2522	-2,1	1689	-2,4

Висновки. На основі статистичних даних щодо імпорту природного газу в Україну побудовано регресійні моделі для короткострокового прогнозування об'єму імпорту. Для оцінювання стану процесу імпорту газу застосовано різні варіанти оптимального фільтра Калмана. Кращі результати прогнозування отримано в результаті застосування адаптивного фільтра з ковзним вікном розміром три значення для розрахунку коваріацій, а гірші за допомогою відомого алгоритму АФК для обох побудованих моделей. Гірші результати застосування відомого АФК можна пояснити наявністю застарілої інформації при перерахуванні коваріації шуму вимірів.

У подальших дослідженнях для прогнозування макроекономічних процесів необхідно застосувати множину взаємодоповнюючих методів прогнозування і процедури комбінування оцінок, отриманих за допомогою різних методів. Також необхідно врахувати зміни у зовнішніх впливах на макроекономіку держави.



СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Хачатрян С. Р. Методы и модели решения экономических задач / С. Р. Хачатрян, М. В. Пинегина, В. П. Буянов. – М. : Экзамен, 2005. – 384 с.
2. Бидюк П. И. Анализ и моделирование процессов переходной экономики / П. И. Бидюк, О. В. Половцев. – К. : ПЛАБ-75, 1999. – 210 с.
3. Довгий С. О. Методи прогнозування у системах підтримки прийняття рішень / С. О. Довгий, П. І. Бідюк, О. М. Трофимчук. – К. : Азимут-Україна, 2011. – 607 с.
4. Pole A. Applied Bayesian forecasting and time series analysis / A. Pole, M. West, J. Harrison. – New York : Chapman & Hall/CRC, 1994. – 409 p.
5. Державна служба статистики України [Електронний ресурс]. – Режим доступу : <http://www.ukrstat.gov.ua/>, вільний. – Заголовок з екрану. – Мова укр., рос., англ.
6. Nummenmaa A. Recursive Noise Adaptive Kalman Filtering by Variational Bayesian Approximations / S. Sarkka, A. Nummenmaa // IEEE Transactions on Automatic Control. – 2009. – Vol. 54 (3). – P. 596-600.
7. Grewal M. S. Kalman filtering: theory and practice using MATLAB / M. S. Grewal, A. P. Andrews. – [3rd ed.]. – Hoboken, New Jersey : A John Wiley & Sons, Inc., 2008. – 575 p.

Караюз И.В., Бидюк П.И. ПРОГНОЗИРОВАНИЕ ОБЪЕМОВ ИМПОРТА ПРИРОДНОГО ГАЗА В УКРАИНУ С ИСПОЛЬЗОВАНИЕМ ФИЛЬТРА КАЛМАНА

В работе решается задача краткосрочного прогнозирования процесса импорта природного газа в Украину на основе регрессионных моделей и адаптивного фильтра Калмана. Цель работы: построение адекватной модели для прогнозирования процесса импорта природного газа и использования адаптивного фильтра Калмана для получения оптимальных оценок состояний исследуемого процесса. Для оценки состояния процесса импорта природного газа использованы разные варианты оптимального адаптивного фильтра Калмана, которые предусматривают повторное вычисление ковариации погрешностей измерений при помощи скользящих окон данных. Основные результаты работы: регрессионные модели процесса импорта природного газа и оценки краткосрочных прогнозов, полученные на основе адаптивного фильтра Калмана со скользящим окном для вычисления ковариации погрешностей измерений. Таким образом, на основе статистических данных построена адекватная модель исследуемого процесса с использованием адаптивного оптимального фильтра Калмана, который обеспечил вычисления высококачественных оценок прогнозов для скользящего окна размерности три измерения.

Ключевые слова: регрессионная модель, адаптивный фильтр Калмана, импорт природного газа, краткосрочное прогнозирование.

Karayuz I.V., Bidiuk P.I. FORECASTING IMPORT OF NATURAL GAS INTO UKRAINE USING KALMAN FILTER

The paper considers solution of the problem of short-term forecasting for import of natural gas into Ukraine. The main purpose of the study is constructing of high quality model for forecasting the process of import of natural gas and application of Kalman filter for computing optimal estimates of states for the process under investigation. Different algorithms of optimal adaptive Kalman filter were applied for computing the state of the import process that suppose repetitive computing of measurement errors covariances using moving data windows. The basic results of the study are as follows: regression models for the process of import of natural gas and estimates of short-term forecasts based on adaptive Kalman filter with moving window for computing measurement errors covariances. Thus, on the basis of statistical data an adequate model was constructed for the process under study with Kalman filter application that provided estimation of high quality forecasts for window size of three samples.

Keywords: regression model, adaptive Kalman filter, import of natural gas, short-term forecasting.

Статтю прийнято
до редакції 2.06.14.