

УДК 519.86:338.27

**ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ПРОГНОЗИРОВАНИЯ ПОКАЗАТЕЛЕЙ ПРОИЗВОДСТВА
МОНОКРИСТАЛЛОВ КРЕМНИЯ****К. А. Куделина, В. Р. Петренко**

Кременчугский национальный университет имени Михаила Остроградского
ул. Первомайская, 20, 39600, г. Кременчуг, Украина. E-mail: pvr@kdu.edu.ua

Предложена информационная технология прогнозирования коэффициентов извлечения материала и выхода в готовую продукцию в производстве монокристаллов кремния. Особенностью предложенной технологии является метод структурной идентификации ARIMA-моделей с использованием нейронной сети.

Ключевые слова: информационная технология, прогнозирование, оперативное управление производством, монокристалл, кластеризация, ARIMA-модели.

**INFORMATION TECHNOLOGY OF FORECASTING OF SILICON MONOCRYSTAL PRODUCTION
INDICATORS****К. А. Kudelina, V. R. Petrenko**

Kremenchuk Mykhailo Ostrohradskyi National University
vul. Pershotravneva, 20, 39600, Kremenchuk, Ukraine. E-mail: pvr@kdu.edu.ua

Proposed information technology prediction material recovery rates and output in the finished product in the production of single-crystal silicon. Feature of the proposed technology is, is a method of structural identification of ARIMA-models using a neural network.

Key words: information technology, forecasting, production operations management, monocrystals, clustering, ARIMA-models.

**ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПРОГНОЗУВАННЯ ПОКАЗНИКІВ ВИРОБНИЦТВА
МОНОКРИСТАЛІВ КРЕМНІЮ****К. О. Куделіна, В. Р. Петренко**

Кременчуцький національний університет імені Михайла Остроградського
вул. Першотравнева, 20, 39600, м. Кременчук, Україна. E-mail: pvr@kdu.edu.ua

Запропоновано інформаційну технологію прогнозування коефіцієнтів витягування матеріалу й виходу в готову продукцію у виробництві монокристалів кремнію. Особливістю запропонованої технології є метод структурної ідентифікації ARIMA-моделей із застосуванням нейронної мережі.

Ключові слова: інформаційна технологія, прогнозування, оперативне управління виробництвом, монокристал, кластеризація, ARIMA-моделі.

АКТУАЛЬНОСТЬ РАБОТЫ. Проблемы современного управления производством объемных монокристаллов кремния по методу Чохральского (ОМКЧ), вытекающие из необходимости повышения его эффективности, требуют оперативного принятия оптимальных решений при изменении условий производства.

Важными показателями состояния процесса производства ОМКЧ являются коэффициент извлечения материала и коэффициент выхода годной продукции. Прогнозирование значений этих показателей – необходимая компонента решения задач оперативно-календарного планирования и распределения производства марок кремния по технологическим установкам, оценки ожидаемых объемов оборотного материала, принятия решений при возникновении нештатных ситуаций [1]. Поэтому разработка информационной технологии контроля и прогнозирования этих производственных показателей с целью повышения эффективности оперативного управления ходом производства, безусловно, является актуальной задачей.

В последнее время усилия исследователей были направлены на разработку моделей и алгоритмов управления нижнего уровня, позволяющих решать конкретные задачи управления производством. Наибольшее внимание привлекло то направление в

приложении моделей, которое связано с разработкой и использованием эффективных алгоритмов краткосрочного прогнозирования. Именно предсказательная способность математических моделей определяет их ценность. Математические модели строятся не столько для описания того или иного явления, сколько для прогнозирования его развития.

В развитии методологии прогнозирования процессов различной природы большую роль сыграли научные разработки отечественных и зарубежных ученых: И.В. Бестужева-Лады, Бокса Дж., Г. Дженкинса, В.М. Глушкова, Ю.П. Лукашина, А.Г. Ивахненко, Е.В. Бодянского, М.З. Згуровского, В.М. Кунцевича, А.А. Павлова, Н.Д. Панкратовой, Ю.П. Зайченко и др.

В настоящее время по оценкам ученых насчитывается свыше сотни различных методов прогнозирования, однако число базовых значительно меньше.

Статистические методы прогнозирования представляют собой совокупность методов обработки количественной информации об объекте прогнозирования, объединенной по принципу выявления содержащихся в ней математических закономерностей изменения характеристик данного объекта с целью получения прогнозных моделей. Данные методы основаны на построении и анализе динамических рядов либо данных случайной выборки. В группу статисти-

ческих методов входят две подгруппы: экстраполяции и моделирования. К первой подгруппе относятся методы наименьших квадратов, экспоненциального сглаживания [2], скользящих средних [3], метод Бокса-Дженкинса [4] и др. Ко второй подгруппе относятся методы математического моделирования, регрессионного и корреляционного анализа и др.

Адекватным аппаратом для решения задач диагностики и прогнозирования могут служить искусственные нейронные сети [5], реализующие идеи предсказания и классификации при наличии обучающих последовательностей, причем как весьма перспективные следует отметить радиально-базисные структуры, отличающиеся высокой скоростью обучения и универсальными аппроксимирующими возможностями [5].

Генетические алгоритмы [6] позволяют решать задачи прогнозирования, классификации, поиска оптимальных вариантов и совершенно незаменимы в тех случаях, когда в обычных условиях решение задачи основано на интуиции или опыте, а не на строгом (в математическом смысле) ее описании.

Особое место в классификации методов прогнозирования временных рядов занимают комбинированные методы, основанные на использовании ряда частных прогнозов, полученных с помощью различных методов (нейронных сетей, нечетких систем, ARIMA-моделей и т.д.) с последующим их обобщением.

При этом, кроме собственно задачи прогнозирования, возникает необходимость в решении еще ряда сопутствующих задач, в частности таких, как кластеризация множества последовательностей значений технико-экономических показателей, построения частных моделей для каждого кластера и получение частных прогнозов, определение параметров комбинирования частных прогнозов и др. Все это и обуславливает целесообразность создания единой информационной технологии, реализация которой позволит автоматизировать решение упомянутой совокупности задач.

В связи с этим целью работы является разработка информационной технологии прогнозирования коэффициентов извлечения и выхода годной продукции в производстве монокристаллов кремния с целью повышения эффективности оперативного управления производством ОМКЧ.

МАТЕРИАЛ И РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ. Информационная технология – совокупность методов, производственных и программно-технологических средств, объединенных в технологическую цепочку, обеспечивающую сбор, хранение, обработку и вывод информации. Информационные технологии предназначены для снижения трудоемкости процессов использования информационных ресурсов.

Для реального производства технико-экономических показателей может быть немало число, особенно это относится к задачам оперативного управления, поскольку последовательности этих показателей зависят от типа оборудования и марки продукции, которых существует большое

число. Однако уменьшить трудоемкость процедуры прогнозирования можно, если из этого множества выделить подмножества последовательностей, которые очень близки по характеристикам, и для анализа и прогнозирования этих последовательностей использовать одну или несколько эталонных моделей.

Задача прогнозирования таких конкретных последовательностей сводится к задаче определения, к какому множеству (кластеру) она относится, и затем к использованию для прогнозирования эталонных моделей этого кластера. Исходя из этого, можно сформулировать основные требования к структуре информационной технологии контроля и прогнозирования технико-экономических показателей.

Основными операциями предлагаемой информационной технологии контроля и прогнозирования коэффициентов извлечения и выхода годной продукции в производстве Cz-Si монокристаллов являются: кластеризация последовательностей коэффициентов, синтез множества моделей для полученных кластеров, получение частных прогнозов по отдельным моделям, обобщение частных прогнозов. Общая схема информационной технологии показана на рис. 1.

Кластеризация последовательностей рассматриваемых коэффициентов нужна для определения групп марок и установок с идентичной динамикой коэффициентов.

Для полученных кластеров целесообразно синтезировать по несколько моделей, наиболее адекватно описывающих исследуемые последовательности. Для этого используется подход Бокса-Дженкинса [4].

С помощью полученных ARIMA-моделей определяются частные прогнозы, причем прогнозирование интересующего показателя выполняется параллельно по нескольким альтернативным моделям с последующей обработкой полученных результатов, заключающейся в построении обобщенного прогноза путем определенного объединения частных прогнозов.

На рис. 2 приведена детальная схема информационной технологии прогнозирования технико-экономических показателей производства монокристаллов кремния.

Рассмотрим подробнее назначение каждой операции информационной технологии.

Учет значений производственных показателей по маркам и установкам. Данная операция реализует учет значений коэффициентов выхода годной продукции и коэффициентов извлечения по каждой марке продукции для каждой технологической установки выращивания монокристаллов кремния в режиме реального времени, т.е. по мере поступления данных.

Коэффициенты рассчитываются, исходя из значений масс загруженного материала, извлеченного материала и годной продукции.

Данные значения масс получают из подсистемы «Обработка слитков» организационно-технологической системы управления производством Cz-Si монокристаллов [7].

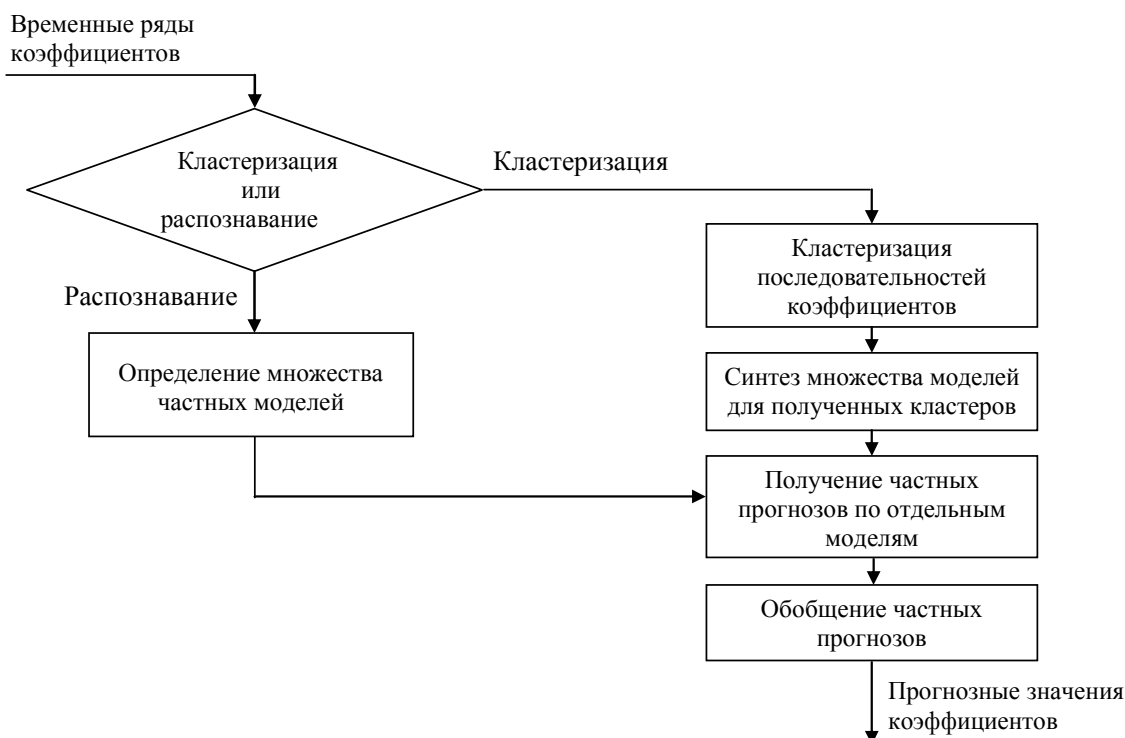


Рисунок 1 – Общая схема информационной технологии

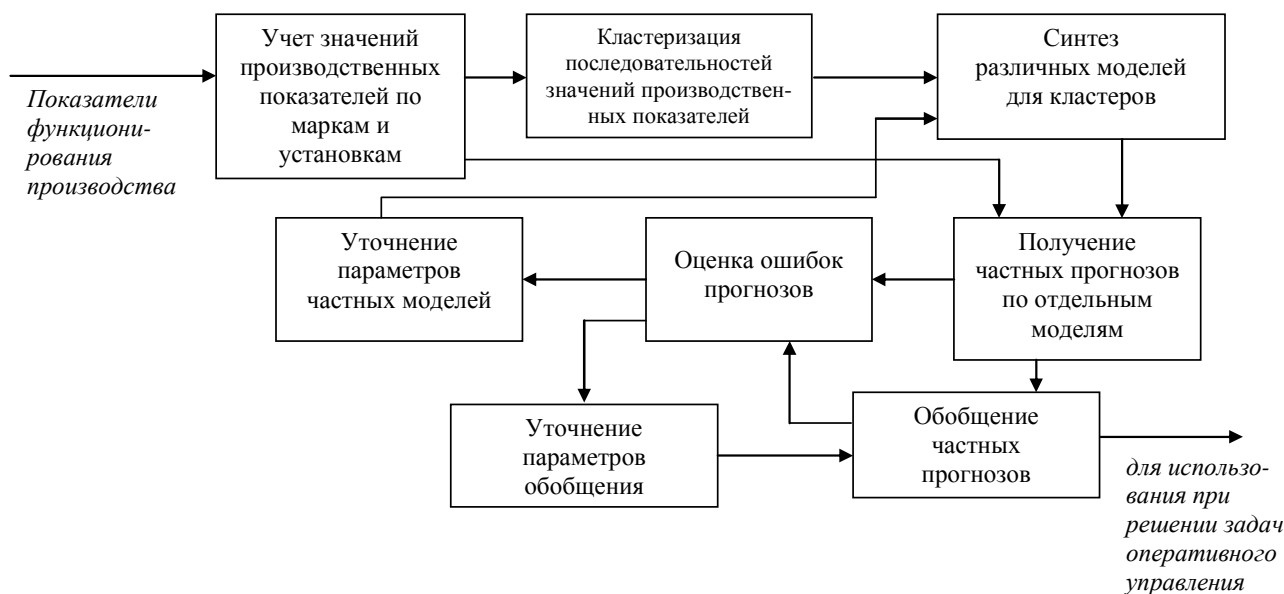


Рисунок 2 – Схема информационной технологии прогнозирования значений производственных показателей

Кластеризация последовательностей значений производственных показателей. Данная операция выполняется первый раз при внедрении системы, а последующие разы – при обнаружении разладки динамических рядов.

Полученные временные последовательности кластеризуются для определения групп марок и установок с идентичной динамикой коэффициента выхода в годную продукцию и коэффициента извлечения с целью сокращения мощности множества исследуемых последовательностей путем описания схожих по характеристикам последовательностей одной и той же моделью.

Кластеризацию предлагается выполнять по двум алгоритмам:

- 1) предложенному нами в [8] алгоритму кластеризации на основе метода взаимного поглощения;
- 2) алгоритму с использованием нейронной сети Кохонена, рассмотренному в [9].

В качестве вектора классификационных признаков используется вектор статистических оценок основных вероятностных характеристик последовательностей (среднее, дисперсия, выборочные статистически значимые автокорреляции).

В случае получения разных результатов окончательное решение о полученных кластерах принима-

ється відповідальним лицем (технологом).

Синтез різних моделей для кластерів. Данна операція виконується перший раз при впровадженні системи, а потім при кожному виявленні розладки динамічних рядів, сформованих значеннями контролюваних показників.

Для синтезу множини моделей для отриманих кластерів використовується підхід Бокса-Дженкінса [4], т.к. моделі ARIMA, які належать до класу лінійних моделей, можуть добре описувати достатньо широкий клас часових послідовностей, однорідних в сенсі стаціонарності кінцевих різниць.

Процедура синтезу ARIMA-моделей включає наступні основні етапи:

1. Ідентифікація моделі ARIMA (p,d,q).
2. Оцінювання ARIMA-моделей.
3. Діагностична перевірка адекватності моделі.

Детально кожен з цих етапів був розглянутий нами в [10].

Як відомо, важливе місце в процесі синтезу моделей часових рядів займає етап визначення структури моделі, який в багатьох випадках залежить від суб'єктивних особливостей дослідника. З метою спрощення процедури ідентифікації структури моделі і виключення впливу на неї особливостей дослідника, нами був запропонований [11] підхід до визначення структури моделі, оснований на використанні нейронної мережі.

Алгоритм даного підходу складається з наступних кроків:

- *формування навчальної вибірки* з значень автокореляцій і частиних автокореляцій для сгенерованих рядів деякої множини відомих моделей;

- *побудова і навчання нейронної мережі*, яка вирішує задачу класифікації, визначаючи, до якого з заданих класів належить даний вхідний набір;

- *тестування мережі* – розпізнавання структури деякої множини синтезованих моделей часових рядів за їхніми значеннями автокореляційної і частиних автокореляційних функцій.

Для попереднього оцінювання ARIMA-моделей використовується відомий метод, оснований на використанні рівнянь Юла-Уокера для оцінювання параметрів оператора авторегресії і алгоритма Ньютона-Рафсона для оцінювання параметрів оператора ковзаючого середнього [4]. Отримані таким чином оцінки потім уточнюються за методом Марквардта [12].

Отримання окремих прогнозів по окремим моделям. Данна операція виконується по мірі надходження нових даних про результати обробки кожного монокристалічного слитка.

З допомогою отриманих ARIMA-моделей будуються окремі прогнози, причому прогнозування цікавого показника виконується паралельно по декільким альтернативним моделям з наступною обробкою отриманих результатів. Використання набору моделей для прогнозуван-

ня одного і того ж показника є одним з шляхів підвищення достовірності прогнозів.

Обобщення окремих прогнозів. На основі отриманих результатів прогнозування по набору моделей будують комбінований прогноз, т.е. обобщенна оцінка, шляхом об'єднання окремих прогнозів. Обобщення прогнозів виконується суміщенням з оптимальними ваговими коефіцієнтами, т.е. такими, які забезпечують мінімальний середній квадрат помилки комбінованого прогнозу. Запропонований нами адаптивний алгоритм обобщення прогнозів детально розглянутий в роботі [13].

Оцінювання помилок прогнозів і уточнення параметрів. В якості характеристики поточної адекватності окремої моделі реальному процесу в даній технології використовується сигнал Тригга-Ліча [13], який обчислюється на основі помилок прогнозів. На основі значень сигналу Тригга-Ліча для окремих моделей прогнозування здійснюється уточнення параметрів обобщення.

Для уточнення параметрів окремих моделей використовується алгоритм, оснований на представленні вихідної ARIMA-моделі в просторі станів з наступним застосуванням рекуррентної фільтрації [14].

ВИВІДИ.

1. Розроблена інформаційна технологія оперативного контролю і прогнозування показників виробництва об'ємних монокристалів кремнію, базуючись на результатах урахування значень контролюваних показників, дозволяє виконувати структурно-параметричний синтез моделей їх динаміки, прогнозувати майбутні значення, виявляти зміни характеристик динамічних рядів, уточнювати їхні математичні моделі і адаптувати алгоритми прогнозування.

2. Особливостями запропонованої ІТ є:

- інтелектуальний метод структурної ідентифікації ARIMA-моделей на основі сукупності значень вибіркової автокореляції і частиних автокореляцій часової послідовності, який базується на використанні нейронної мережі і в відмінність від відомих методів аналогічного призначення не потребує наявності висококваліфікованих фахівців, і не використовує перебору структур моделі на заданому кінцевому множині структур;

- алгоритми кластеризації (на основі удосконаленого методу взаємного поглинання і з використанням нейронної мережі Кохонена), які дозволяють розбити вхідне множині послідовностей на деякі непересекаючі підмножини з близькими значеннями характеристик з метою зменшення трудомісткості їх моделювання. В якості вектора класифікаційних ознак запропоновано використовувати вектор статистичних оцінок основних ймовірнісних характеристик послідовностей (середнє, дисперсія, вибіркові статистично значимі автокореляції);

- адаптивний алгоритм побудови обобщеного прогнозу, базуючийся на адаптивному оцінюванні

параметров обобщения в зависимости от текущих значений меры точности частных прогнозов, позволяет повысить точность обобщенного прогноза.

3. Разработанная информационная технология была использована при проектировании подсистемы «Моделирование и прогнозирование» для автоматизированной организационно-технологической системы управления производством объемных Cz-Si-монокристаллов.

ЛИТЕРАТУРА

1. Оптимизация решений при возникновении нештатных ситуаций в процессе выращивания монокристаллов кремния / А.П. Оксанич, В.Р. Петренко, Л.Г. Шепель // Вісник Кременчуцького державного політехнічного університета. – 2007. – Вип. 3(44), ч. 2. – С. 164–168.

2. Brown R.G. Smoothing, Forecasting and Prediction of Discrete Time Series. – N.Y.: Prentice–Hall, 1963. – 468 p.

3. Теория статистики / Дж. Юл, М.Дж. Кендалл. – М.: Госстатиздат, 1960. – 245 с.

4. Анализ временных рядов. Прогноз и управление. / Дж. Бокс, Г. Дженкинс; пер. с англ. А.Л. Левшина. – М.: Мир, 1974. – Вып. 1. – 406 с.

5. Основы теории искусственных нейронных сетей / О.Г. Руденко, Е.В. Бодянский. – Х.: ТЕЛЕТЕХ, 2002. – 317 с.

6. Holland J. Adaptation In Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. – Cambridge: 1st MIT Press Ed. – 1992. – 211 p.

7. Петренко В.Р. Автоматизація управління технологічними процесами виробництва об'ємних Cz–Si–монокристалів: автореф. дис. на здобуття наук. ступеня д–ра техн. наук: спец. 05.13.07 Автоматизація процесів керування. – К., 2009. – 39 с.

8. Применение кластеризации в задаче моделирования стационарных временных последовательностей / В.Р. Петренко, К.А. Куделина, В.А. Самоляк // Новые технологии. – 2007. – № 4(18). – С. 90–94.

9. Кластеризація часових послідовностей з використанням нейронної мережі Кохонена / В.Р. Петренко, К.О. Куделіна // Системні технології. Регіональний міжвузівський збірник наукових праць. – 2008. – Т. 2, вип. 3(56). – С. 210–215.

10. Синтез ARIMA–моделей динамики технико–экономических показателей производства монокристаллического кремния / В.Р. Петренко, К.А. Куделина, Л.Г. Шепель // Новые технологии. – 2006. – № 2. – С. 189–196.

11. Нейромережевий підхід до структурної ідентифікації ARMA–моделей часових послідовностей / В.Р. Петренко, К.О. Куделіна // Складні системи і процеси. – 2008. – № 1(13). – С. 87–93.

12. Marquardt D.W. An Algorithm for least squares estimation of non–linear parameters // I. Int. Appl. Math. – 1963. – № 11. – P. 431–440.

13. Адаптивний алгоритм узагальнення часткових прогнозів / В.Р. Петренко, К.О. Куделіна, А.Г. Петров // Системні технології. – 2009. – № 2 (61). – С. 111–119.

14. Об одном алгоритме оценивания APCC–моделей стационарных процессов / В.Р. Петренко, И.Г. Кротюк // Системні технології. Регіональний міжвузівський зб. наукових праць. – 2001. – Вип. 4 (15). – С. 42–48.

REFERENCE

1. Decision-optimization at non-standard situations during Si-Cz monocrystals growing / A.P. Oksanich, V.R. Petrenko, L.G. Shepel // Transactions of Kremenchuk Mykhaylo Ostrogradskiy State University. – 2007. – V. 3(44), p. 2. – P. 164–168 [in Russian].

2. Brown R.G. Smoothing, Forecasting and Prediction of Discrete Time Series. – N.Y.: Prentice –Hall, 1963. – 468 p.

3. The Theory of Statistics / G. Yule, M. Kendall. – M.: Gosstatizdat, 1960. – 245 p. [in Russian].

4. Time series analysis: Forecasting and control / G. Box, G. Jenkins; translate from English by A.Levshin. – M.: World, 1974. – 406 p. [in Russian].

5. Basis of artificial neural network theory / O.G. Rudenko, E.V. Bodyanskiy. – Kharkov: Teletech, 2002. – 317 p. [in Russian].

6. Holland J. Adaptation In Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. – Cambridge: 1st MIT Press Ed. – 1992. – 211 p.

7. Petrenko V.R. Automation of controlling technological processes for solid Cz-Si-monocrystals production: The thesis on obtaining the Doctor's Degree of Engineering Sciences on speciality 05.13.07 Automation of controlling processes. – K., 2009. – 39 p. [in Ukrainian].

8. Clustering application in stationary time sequences modeling problem / V.R. Petrenko, C.O. Kudelina, V.A. Samolyak // New Technologies. – 2007. – № 4 (18). – P. 90–94 [in Russian].

9. Time sequences clustering with the Kohonen Neural Network usage / V.R. Petrenko, C.O. Kudelina // System Technologies. – 2008. – Vol. 2, № 3(56). – P. 210–215 [in Ukrainian].

10. ARIMA-models synthesis of monocrystalline silicon production technical and economic indices dynamics / V.R. Petrenko, K.O. Kudelina, L.H. Shepel // New Technologies. – 2006. – № 2 (12). – P. 189–196 [in Russian].

11. A neural network approach to structural identification of time sequences ARMA-models/ V.R. Petrenko, C.O. Kudelina // Complex systems and processes. – 2008. – № 1 (13). – P. 87–93 [in Ukrainian].

12. Marquardt D. W. An Algorithm for least squares estimation of non–linear parameters // I. Int. Appl. Math. – 1963. – № 11. – P. 431–440.

13. An Adaptive Algorithm of Particular Forecasts Combining/ V.R. Petrenko, C.O. Kudelina. A.G. Petrov. // System Technologies. – 2009. – № 2 (61). – P. 111–119 [in Ukrainian].

14. About the one of estimation algorithm of stationary processes ARMA–models / V.R. Petrenko, I.G. Krotiyuk // System Technologies. – 2001. – № 4(15). – P. 42–48 [in Russian].

Стаття надійшла 20.01.2011.

Рекомендована до друку д.т.н., проф. Оксаничем А.П.