

АВТОМАТИЗОВАНИЙ ВИБІР ПРОГНОЗУЮЧИХ МОДЕЛЕЙ У СИСТЕМАХ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ

Виконала студентка 4 курсу групи КА-44
Бусол Анна Євгеніївна

Науковий керівник: д.т.н., професор кафедри ММСА
Бідюк Петро Іванович

Актуальність:

- Комп'ютерні технології дозволяють обробляти великі обсяги даних за прийнятні витрати часу, будувати адекватні математичні моделі досліджуваних процесів, прогнозувати розвиток динамічних процесів та планувати управлінські дії.
- Прогнозування – один з вирішальних елементів організації управління та прийняття обґрунтованих рішень у різних сферах життєдіяльності.
- Системи підтримки прийняття рішень дозволяють швидко обрати кращий варіант розв'язання задач з множини існуючих за допомогою відповідних критеріїв якості.

Об'єкт і предмет дослідження:

Об'єкт: Лінійні та нелінійні нестационарні процеси в економіці та фінансах.

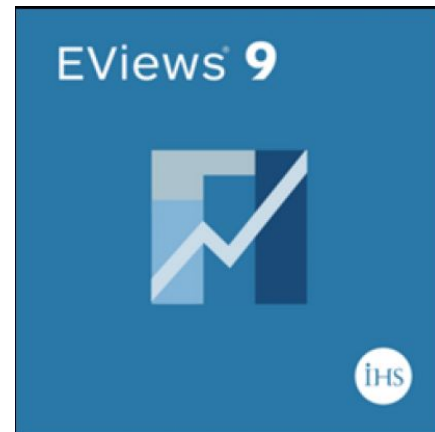
Предмет: Математичні моделі і методи формального опису процесів в економіці та фінансах.

Мета: Створення системи підтримки прийняття рішень для побудови математичних моделей на основі статистичних даних і автоматизованого вибору прогнозуючих моделей.

Постановка задачі:

- Виконати огляд відомих систем для побудови прогнозуючих моделей. Запропонувати комплексні критерії вибору моделі.
- Розробити архітектуру і функціональну схему системи для побудови та автоматизації вибору прогнозуючих моделей.
- Розробити ПЗ для розв'язання задач побудови математичної моделі вибраних фінансово-економічних процесів (ФЕП) з функцією автоматизованого вибору кращої моделі.
- Виконати обчислювальні експерименти за допомогою розробленого ПЗ з використанням фактичних статистичних даних.
- Проаналізувати отримані результати, визначити перспективи подальших досліджень.

СИСТЕМИ ДЛЯ СТАТИСТИЧНОЇ ОБРОБКИ ДАНИХ



Деякі типи регресійних моделей

- Авторегресія

$$y(k) = a_0 + a_1 y(k-1) + \dots + a_p y(k-p) = a_0 + \sum_{i=1}^p a_i y(k-i) + \varepsilon(k)$$

- Парна регресія

$$y(k) = a_0 + a_1 x(k) + \varepsilon(k)$$

- Множинна регресія

$$y(k) = a_0 + a_1 x_1(k) + a_2 x_2(k) + \dots + a_p x_p(k) + \varepsilon(k)$$

- Авторегресія з ковзним середнім

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

- Змішана регресія

$$y(k) = a_0 + \sum_{i=1}^l a_i y(k-i) + b_1 x_1(k) + b_2 x_2(k) + \dots + b_p x_p(k) + \varepsilon(k)$$

Критерії адекватності математичної моделі:

- Коефіцієнт множинної детермінації: $R^2 = \frac{\text{var}(\hat{y})}{\text{var}(y)}$

- Сума квадратів похибок:

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

- Статистика Дарбіна-Уотсона:

$$DW = 2 - 2 \frac{E[e(k)e(k-1)]}{\sigma_e^2}$$

- Інформаційний критерій Акайке: $AIC = N \cdot \ln(\sum_{k=1}^N e^2(k)) + 2n$

- Критерій Байєса-Шварца: $BSC = N \cdot \ln(\sum_{k=1}^N e^2(k)) + n \cdot \ln(N)$

- Статистика Стьюдента: $t_a = \frac{\hat{a} - a_0}{SE_{\hat{a}}}$

- Статистика Фішера: $F \sim \frac{R^2}{1-R^2}$

Критерії якості оцінок прогнозів:

- Середньоквадратична похибка:

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

- Середня похибка:

$$\text{СП} = \frac{1}{N} \sum_{k=1}^N [y(k) - \hat{y}(k)] = \frac{1}{N} \sum_{k=1}^N e(k)$$

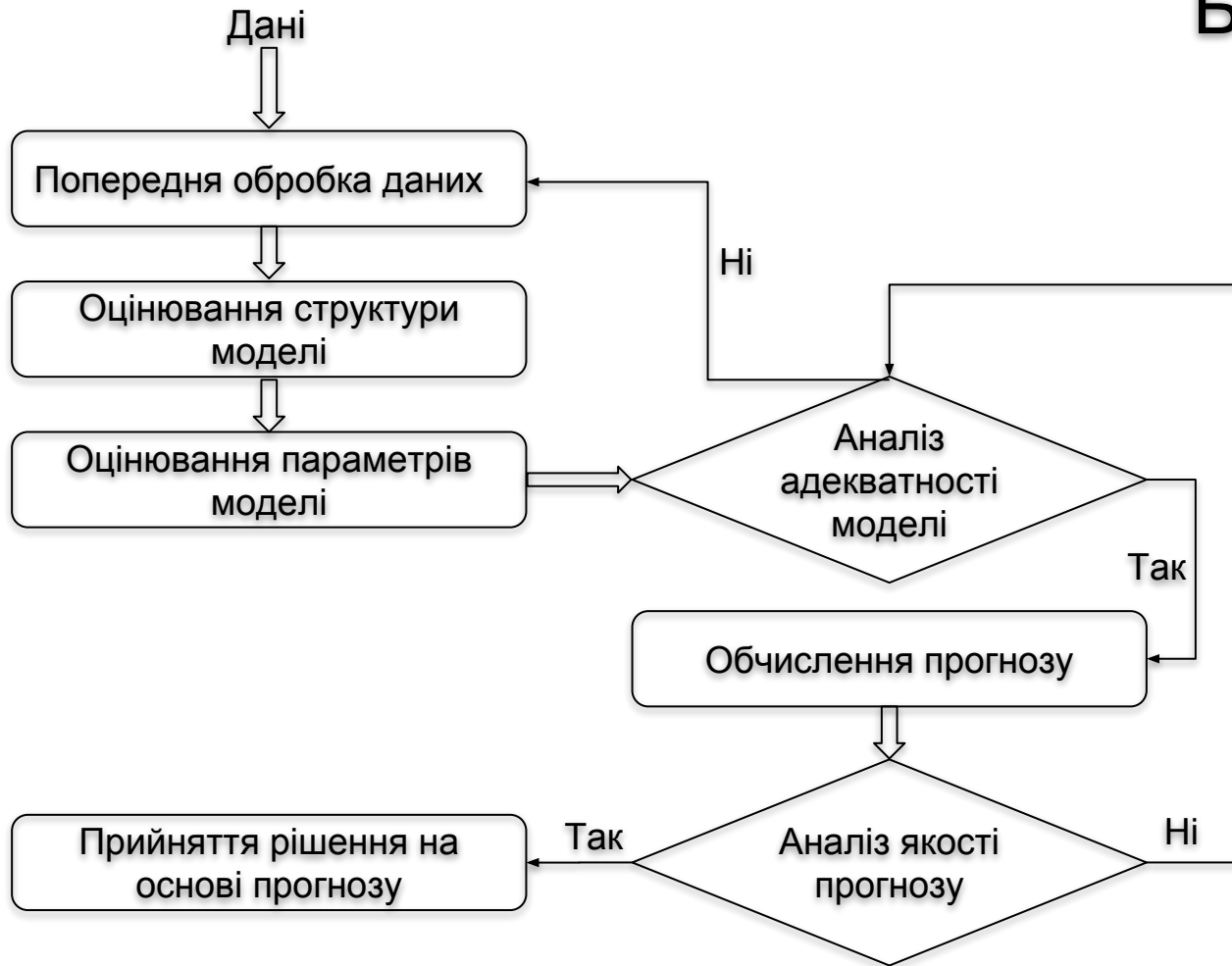
- Середня похибка у процентах:

$$\text{СПП} = \frac{1}{N} \sum_{k=1}^N \frac{[y(k) - \hat{y}(k)]}{y(k)} \times 100\%$$

- Середнє абсолютне значення похибки у процентах:

$$\text{САПП} = \frac{1}{N} \sum_{k=1}^N \frac{|y(k) - \hat{y}(k)|}{|y(k)|} \times 100\%$$

Блок-схема



Автоматизований вибір моделі за комплексним критерієм якості

Система розраховує m моделей, з яких вибирається краща за наведеним критерієм:

$$I = |1 - R^2| + 0,5 |2 - DW| \longrightarrow \min$$

$$I_1, \dots, I_m \longrightarrow I_{\min}$$

Приклади роботи програми

Decision Support System : Busol Ann, IASA-2018

Вхідний ряд

n	Y(n)
0	11348,13
1	11348,13
2	11348,13
3	11348,13
4	13580,25
5	13580,25
6	13580,25
7	13580,25
8	15790,43
9	15790,43
10	15790,43
11	15790,43
12	24516,29

Модель АРКС

Оберіть діапазони параметрів

Порядок АР: від 1 до 15 Порядок КС: від 0 до 15

Оберіть метод наближення

МНК

Оберіть критерій оптимальності моделі

Сумма квадратів похибок

Додатково

☒ Врахувати a0

Старт ...

Параметри моделі

Параметр	Значення
a0	898,12167
a1	0.8370267
a2	-1,3877787
a3	0.0031454
a4	0.6696389
a5	-0.5630880
a6	-6.8072283
a7	3.8429830
a8	-0.1673572
a9	0.0922206
a10	7.8825834
a11	0.0548990

Прогноз ...

Вихідний ряд

n	Y(n)
0	24317,8716396
1	24317,8716396
2	24358,7718490
3	31029,9795001
4	31330,9807180
5	31330,9807180
6	31366,3179893
7	35725,6821072
8	39293,6354541
9	39293,6354541
10	39374,0935880
11	48845,5950774
12	55219,6555336

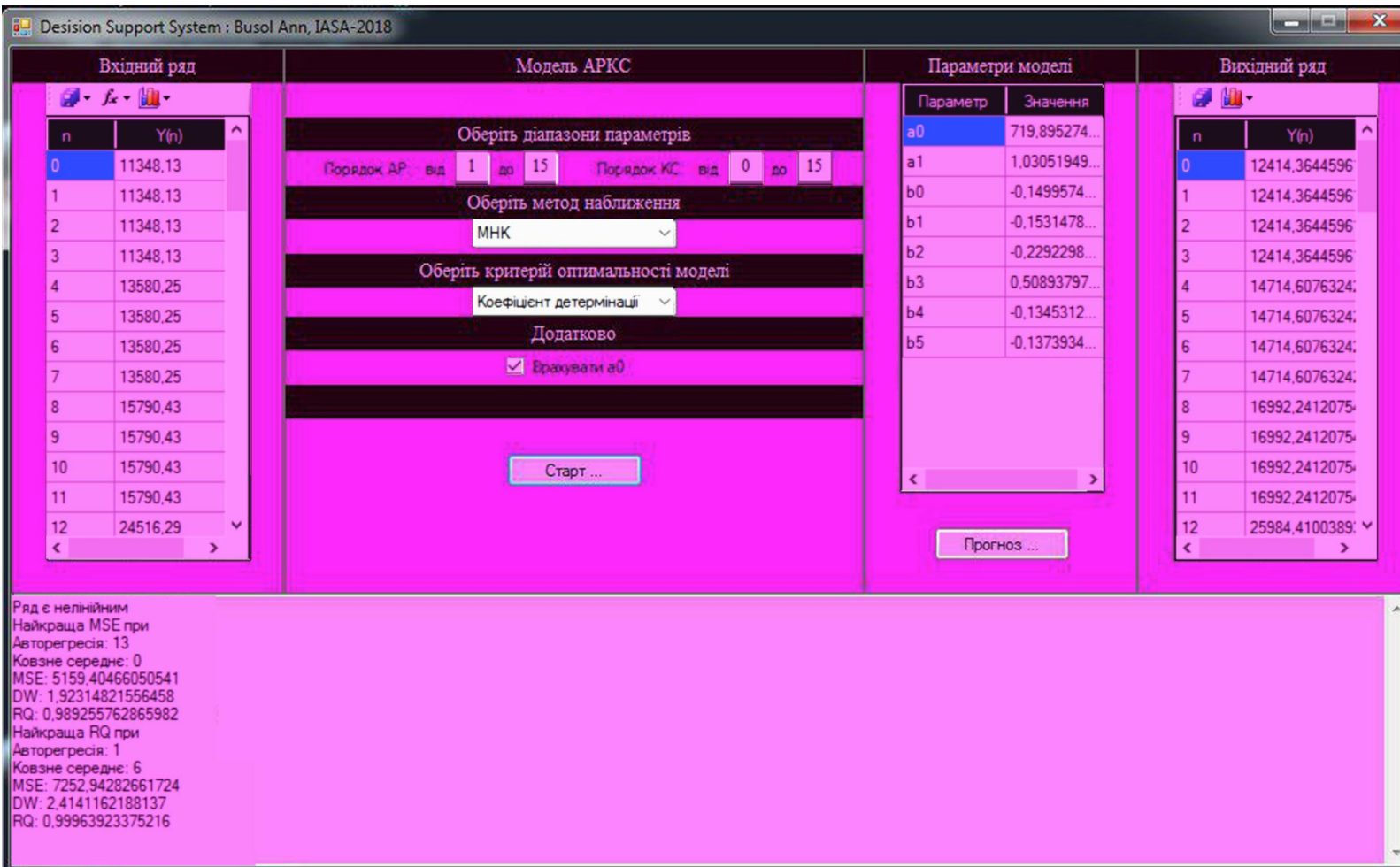
Завантажено ряд PODATOK
Кількість елементів: 67
Ряд є нелінійним
Найкраща MSE при
Авторегресія: 13
Ковзне середнє: 0
MSE: 5159,40466050541
DW: 1,92314821556458
RQ: 0,989255762865982

Побудована регресійна модель

$$y(k)=898,1217+0,837*y(k-1)-1,3878*y(k-2)+0,0031*y(k-3)+0,6696*y(k-4)-0,5631*y(k-5)-6,8072*y(k-6)+3,843*y(k-7)-0,1674*y(k-8)+0,0922*y(k-9)+7,8826*y(k-10)+0,006*y(k-11)+0,5367*y(k-12)-0,35*y(k-13)$$

$$R^2=0,9893; e^2=5159,405; DW=1,9231$$

$$СКП = 16932; СПП = 2,005\%; САПП = 2,005\%$$



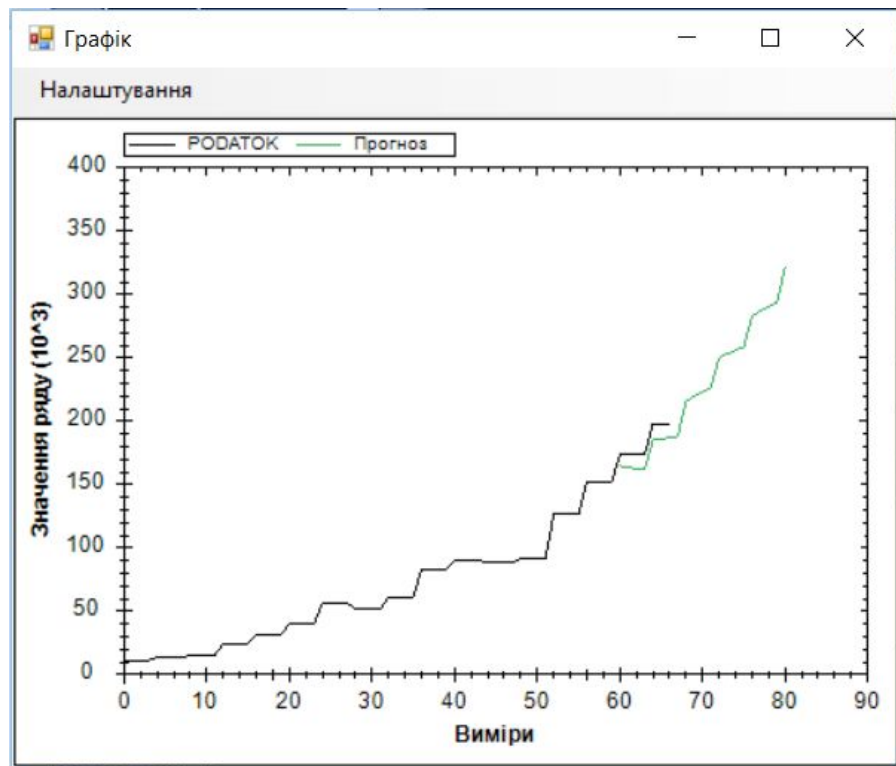
Побудована регресійна модель

$$\begin{aligned} y(k) &= 719,895 + 1,0305 * y(k-1) - 0,15 * v(k-1) - 0,1531 * v(k-2) - 0,2292 * v(k-3) + 0,5089 * v(k-4) - 0,1345 y(k) \\ &= 719,895 + 1,0305 * y(k-1) - 0,15 * v(k-1) - 0,1531 * v(k-2) - 0,2292 * v(k-3) + 0,5089 * v(k-4) - 0,1345 \end{aligned}$$

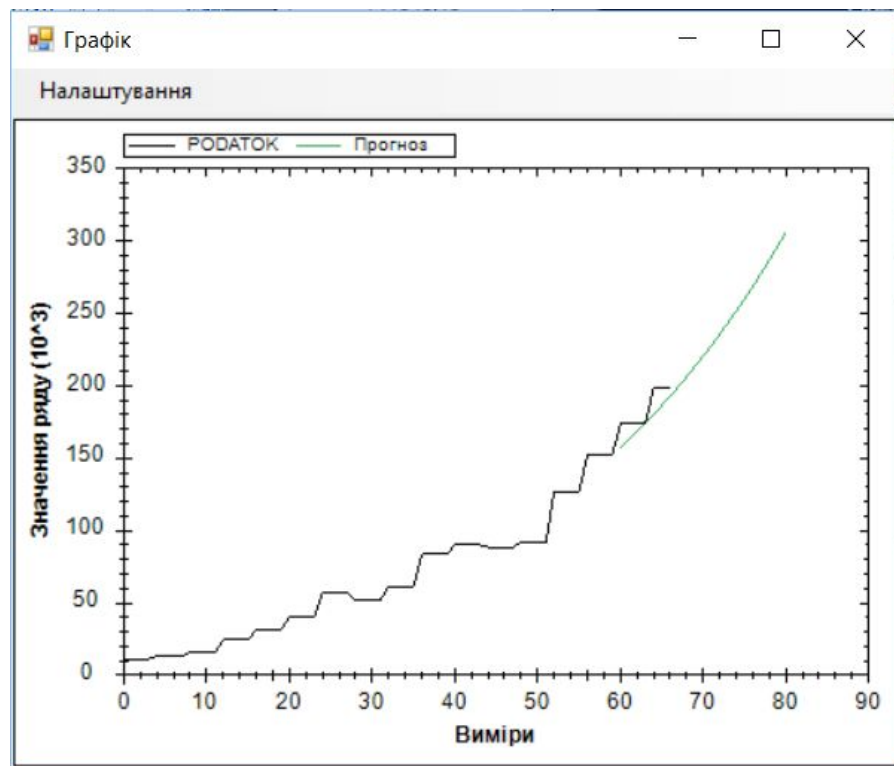
$$R^2 = 0,9996; e^2 = 7252,9; DW = 2,1414$$

$$СКП = 14613; СПП = 1,7186\%; САПП = 1,7186\%$$

Прогнози



САПП = 2,005%



САПП = 1,7186%

Аналіз результатів

Ряд	Тип моделі	Адекватність моделі			Характеристика прогнозу		
		R^2	$\sum e^2$	DW	СКП	СПП %	САПП %
PODATOK	АРКС(13, 0)	0,9893	5159	1,923 1	16932	2,005	2,005
	АРКС(8, 0)	0,9882	5603	2,017	2492	2,84	2,84
	АРКС(1, 6)	0,9996	7252	2,141 4	14613	1,7186	1,7186

Ряд	Тип моделі	Адекватність моделі			Характеристика прогнозу		
		R	$\sum e^2$	DW	СКП	СПП %	САПП %
GAZ	AP (1)	0,904	46,33	1	266	-3,39	3,39

Подальше дослідження

- Розширення СППР моделями інших типів, наприклад, моделями, які будуються за методом групового врахування аргументів: Байєсівська регресія, Байєсівська мережа та ін.
- Обчислення оцінок прогнозів шляхом комбінування прогнозів, отриманих за допомогою різних методів
- Подальше підвищення рівня автоматизації процесу моделювання і прогнозування

Висновки

- Виконано огляд моделей, які можуть бути використані для формального опису лінійних та нелінійних процесів в економіці та фінансах.
- Розроблена та програмно реалізована СППР для обробки статистичних даних, побудови моделей ФЕП та оцінювання прогнозів.
- Для вибраних ФЕП побудовані регресійні моделі, які характеризуються високим ступенем адекватності і забезпечують можливість обчислення високоякісних оцінок прогнозів.

Дякую за увагу!