

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»
НАВЧАЛЬНО-НАУКОВИЙ КОМПЛЕКС «ІНСТИТУТ ПРИКЛАДНОГО
СИСТЕМНОГО АНАЛІЗУ»
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ СИСТЕМНОГО АНАЛІЗУ**

Дипломна робота на тему:

**Автоматизований вибір прогнозуючих моделей у системах підтримки
прийняття рішень**

**Виконав студент IV курсу
групи КА-23**

Грицаюк Максим

Керівник: проф. Бідюк П.І.

МЕТА РОБОТИ

Розробка системи підтримки прийняття рішень для автоматизації вибору прогнозуючих регресійних моделей з використанням фактичних статистичних даних.

ОБ'ЄКТ ТА ПРЕДМЕТ ДОСЛІДЖЕННЯ

Об'єкт дослідження: процеси різної природи у вигляді часових рядів.

Предмет дослідження: математичні моделі і методи опису обраних процесів, критерії адекватності моделей і якості оцінок прогнозів.

АКТУАЛЬНІСТЬ РОБОТИ

- ❖ сьогодні існує необхідність побудови адекватних математичних і статистичних моделей для оцінювання коротко- і довгострокових прогнозів з метою підвищення якості рішень, які приймаються на їх основі.
- ❖ оскільки у таких поширених програмних додатках, як Eviews та STATISTICA, відсутній автоматизований вибір кращої моделі, то тема є надзвичайно актуальною.

ПОСТАНОВКА ЗАДАЧІ

1. Виконати огляд сучасних математичних моделей для опису динаміки часових рядів в економіці та фінансах.
2. Розробити комбінований критерій для автоматизованого вибору кращої регресійної моделі, побудованої на основі статистичних даних.
3. Розробити програмне забезпечення для виконання необхідних обчислювальних експериментів (мова програмування C#).

ПОСТАНОВКА ЗАДАЧІ

4. Зібрати статистичні дані для виконання обчислювальних експериментів.
5. Побудувати регресійні моделі на основі зібраних даних та обчислити статистичні оцінки адекватності моделей та якості прогнозів.
6. Вибрати кращі моделі для короткострокового прогнозування та виконати порівняльний аналіз отриманих результатів.

МОДЕЛІ, ВИКОРИСТАНІ В СППР

Модель авторегресії з ковзним середнім АРКС(p, q):

$$y(k) = a_0 + a_1y(k - 1) + \dots + a_p y(k - p) + b_1v(k - 1) + \dots + b_q v(k - q)$$

Способи підрахунку ковзного середнього:

1. Просте ковзне середнє

$$MA(k) = \frac{y(k-n) + \dots + y(k-1) + y(k)}{n}, \text{ де } n - \text{ вікно ковзного середнього}$$

2. Експоненційне ковзне середнє

$$EMA(k) = \frac{w_1y(k) + w_2y(k - 1) + \dots + w_ny(k - n)}{\sum_{i=1}^n w_i}$$

МОДЕЛІ, ВИКОРИСТАНІ В СППР

Вагові коефіцієнти експоненційного ковзного середнього:

ваги ЕКС зменшуються з віддаленістю від поточного періоду

$$\diamond w_1 = 1 - a; w_2 = (1 - a)^2; \dots; w_n = (1 - a)^n, \text{ де } a = \frac{2}{n+1}$$

$$\diamond w_1 = e^n; w_2 = e^{n-1}; \dots; w_n = e$$

3. Білий шум, що накладений на вибірку

$\xi(k) \sim N(0; 0,01), \forall k \in N$ (на нормованій вибірці)

$$КС(k) = y(k - 1) + \xi(k)$$

МЕТОДИ ОЦІНЮВАННЯ ПАРАМЕТРІВ МОДЕЛЕЙ, ВИКОРИСТАНІ В СППР

1. МНК:

$$\hat{\theta} = [X^T X]^{-1} X^T y$$

2. РМНК:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \gamma(k)[y(k) - \hat{\theta}^T(k-1)\Psi(k)]$$

$$\gamma(k) = \frac{P(k-1)\Psi(k)}{a_k^{-1} + \Psi^T(k)P(k-1)\Psi(k)}$$

$$P(k) = P(k-1) - \frac{P(k-1)\Psi(k)\Psi^T(k)P(k-1)}{a_k^{-1} + \Psi^T(k)P(k-1)\Psi(k)}$$

де $\Psi(k)$ - вектор вимірів;

$$\Psi^T(k) = [-y(k-1); -y(k-2); \dots; -y(k-p); v(k-1); v(k-2); \dots; v(k-q)]$$

КРИТЕРІЇ АДЕКВАТНОСТІ МОДЕЛІ, ВИКОРИСТАНІ В СППР

1. Сума квадратів похибок моделі:

$SSE = \sum_{k=1}^N (y(k) - \hat{y}(k))^2$, де N – потужність вибірки, $y(k)$ – реальне значення ряду в момент часу k , $\hat{y}(k)$ – оцінене значення.

2. Коефіцієнт множинної детермінації R^2 :

$R^2 = \frac{var(\hat{y})}{var(y)}$, де var – вибіркова дисперсія.

3. Статистика Дарбіна-Уотсона:

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

КРИТЕРІЇ АДЕКВАТНОСТІ МОДЕЛІ, ВИКОРИСТАНІ В СППР

4. Інформаційний критерій Акайке:

$$AIC = N \ln\left(\sum_{k=1}^N e^2(k)\right) + 2n, \text{ де } n = p + q + 1$$

5. Критерій Байєса-Шварца:

$$BSC = N \ln\left(\sum_{k=1}^N e^2(k)\right) + n \ln(N)$$

КРИТЕРІЇ ЯКОСТІ ОЦІНОК ПРОГНОЗІВ, ВИКОРИСТАНІ В СППР

1. Середньо-квадратична похибка:

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

2. Середня абсолютна похибка в процентах:

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

3. Коефіцієнт нерівності Тейла:

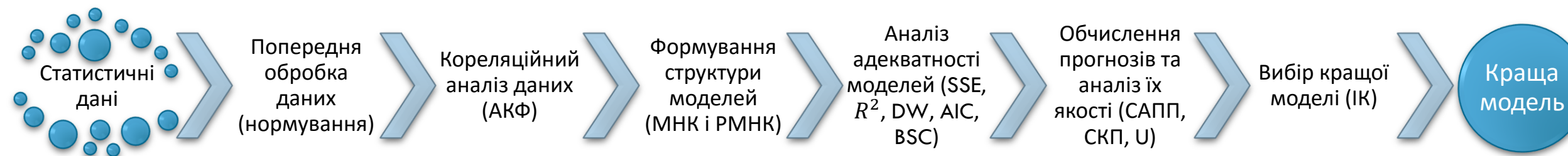
$$U = \frac{\sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - \hat{y}(k)]^2}}{\sqrt{\frac{1}{N} \sum_{k=1}^N y^2(k) + \frac{1}{N} \sum_{k=1}^N \hat{y}^2(k)}}$$

ЗАПРОПОНОВАНИЙ КОМБІНОВАНИЙ КРИТЕРІЙ ВИБОРУ КРАЩОЇ МОДЕЛІ В СППР

Інтегральний критерій якості:

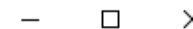
$$IK = e^{1-R^2} + \frac{SSE}{N} + \begin{cases} \ln(AIC + BSC), & AIC + BSC > 0 \\ e^{AIC+BSC}, & AIC + BSC \leq 0 \end{cases} +$$
$$+ e^{2-DW} + \ln(СКП) + \ln(САПП) + e^U$$

ФУНКЦІОНАЛЬНА СХЕМА



ІНТЕРФЕЙС СППР

СППР вибору кращої моделі



Введення даних

Файл з рядом

Вид ковзного середнього
Експоненційне

Вікно ковзного середнього:

Прогноз кроків:

Нормувати дані

Краща модель

Результати для всіх моделей

Модель	ІК	BSC	САПП, %	DW	U (коэф. Тейла)	Σe^2	AIC	R^2
--------	----	-----	---------	----	-----------------	--------------	-----	-------

ПРИКЛАДИ РОБОТИ СППР (ІНДЕКС СПОЖИВЧИХ ЦІН УКРАЇНИ У 1991 – 2009 РР.)

Нормовані дані, експоненційне ковзне середнє з вікном 3, прогноз на 1 крок.

СППР вибору кращої моделі

Введення даних

Файл з рядом
C:\Users\max\Desktop\індекс спож | Обзор

Вид ковзного середнього
Експоненційне

Вікно ковзного середнього: 3

Прогноз кроків: 1

Нормувати дані | Зчитати

Краща модель

Графік

АРКС(1,5)
 $y(k) = 0,00053233252539858 + -0,0150324476162293 \times y(k-1) + 1,75046274542699 \times x(k-1) + -0,847630051561242 \times x(k-2) + -0,0125463562588824 \times x(k-3) + 0,215901210717078 \times x(k-4) + -0,104403552731362 \times x(k-5)$

Результати для всіх моделей

Пуск! | АКФ

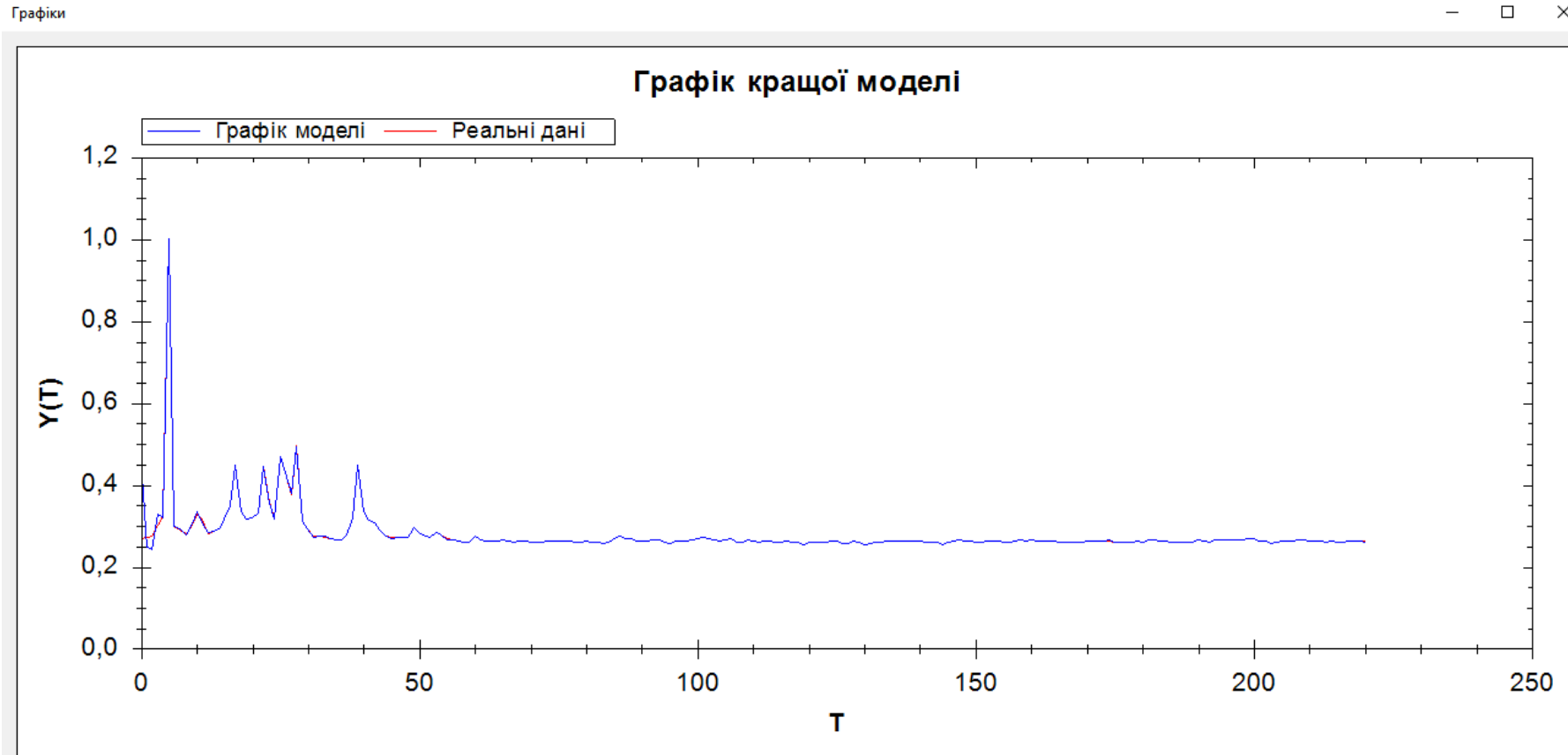
▶	Модель	ІК	BSC	САПП, %	DW	U (коэф. Тейла)	Σe^2	AIC	R ²
▶	АРКС(1,5)	-8,06264292941392	-650,599093975079	0,919100820280261	1,29438525971694	0,00461672025828326	0,043766612196216	-674,354486799545	0,951091961132958
	АРКС(1,6)	-5,85521339076987	-438,321567546669	2,35708968780747	1,72941932669541	0,0119260017173612	0,112084450180044	-465,470587917488	0,869957087981076
	АРКС(1,2)	-1,36391247806459	-638,713067941244	9,61136719621284	1,50180936416925	0,0504828836400029	0,0497220305583653	-652,287578126653	0,961637101062541
	АРКС(1,4)	-0,570004187143976	-652,323537311413	8,20624540369852	0,930504003799432	0,0394140213603446	0,0445026775987595	-672,685302589528	0,950637294688016
	АРКС(1,3)	0,551255291100063	-460,996334299599	19,3755724777952	1,74565104849168	0,107269945397686	0,108824332265833	-477,96447203136	0,864184374716404
	АРКС(1,1)	5,52900325249369	-633,212091923731	64,5544633027925	1,04514803567303	0,244011998500693	0,052246297702681	-643,392974562788	0,959627732063113

Автокореляційна функція

	АКФ	АКФ Y	АКФ KC
▶	r(1)	0,310609943996...	0,721898788677...
	r(2)	0,246921154649...	0,516809592660...
	r(3)	0,210538688404...	0,365981055478...
	r(4)	0,174163217413...	0,328045652362...
	r(5)	0,205414285795...	0,319521539312...
	r(6)	0,201429353180...	0,301673041822...
	r(7)	0,132526864618...	0,258686208278...
	r(8)	0,140834773854...	0,245346533470...
	r(9)	0,139572410521...	0,259837677370...
	r(10)	0,171825720220...	0,317898076553...
*			

ПРИКЛАДИ РОБОТИ СППР (ІНДЕКС СПОЖИВЧИХ ЦІН УКРАЇНИ У 1991 – 2009 РР.)

Нормовані дані, експоненційне ковзне середнє з вікном 3, прогноз на 1 крок.



ПРИКЛАДИ РОБОТИ СППР (ВВП УКРАЇНИ У 1996 – 2005 РР.)

Нормовані дані, експоненційне ковзне середнє з вікном 3, прогноз на 1 крок.

СППР вибору кращої моделі

Введення даних

Файл з рядом
C:\Users\max\Desktop\ввп по місяц. Обзор

Вид ковзного середнього
Експоненційне

Вікно ковзного середнього: 3

Прогноз кроків: 1

Нормувати дані Зчитати

Краща модель
Графік

АРКС(1,3)

$$y(k) = -0.000476848966164546 + -0.546738798585468 y(k-1) + 1.75787675273392 x(k-1) + 0.0637376239833722 x(k-2) + -0.457457012972843 x(k-3)$$

Результати для всіх моделей

Пуск! АКФ

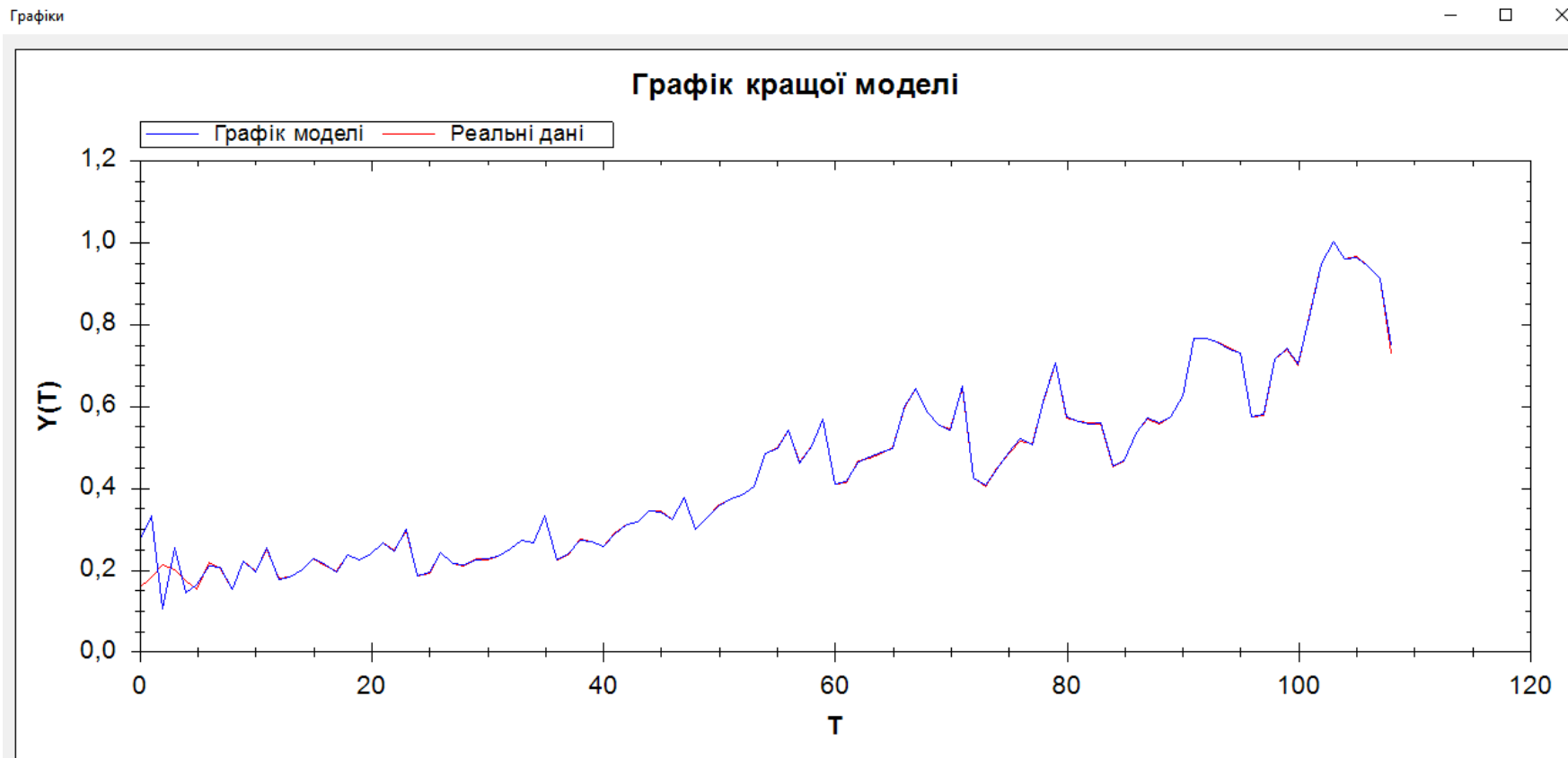
	Модель	IK	BSC	САПП, %	DW	U (коэф. Тейла)	Σe^2	AIC	R ²
▶	АРКС(1,3)	-3.84683536018211	-296.512126625286	2.69190732161432	2.00138275341115	0.0132807834174802	0.0517025596756041	-309.922782760907	0.988881039790974
	АРКС(1,2)	-2.26893208011035	-426.237180732351	3.67722210681823	1.50637343675279	0.0187304914196914	0.0162436562667932	-436.965705640848	0.991512480835499
	АРКС(1,6)	3.0250483067192	-297.344202979077	24.4493898104765	1.86117733821414	0.10893052474378	0.04504867982748	-318.801252796071	0.988029124996336
	АРКС(1,5)	3.0564434545358	-418.527583111658	20.6962175130171	1.46719227980486	0.0937769470915216	0.0153179969959567	-437.302501701528	0.994358449192565
	АРКС(1,9)	3.30670155371513	-283.441398739539	26.6667055134969	1.85128867537633	0.117647210043863	0.0449888259696225	-312.944842237906	0.987987433216111
	АРКС(1,8)	4.34716748685234	-410.042038211083	24.5572042768662	1.09941777332963	0.109358345264169	0.0145492445427996	-436.863350482325	0.991385233658891
	АРКС(1,10)	5.12085270875416	-385.298074280452	25.4897968489993	0.86351012579821	0.113041907905344	0.0167759063987818	-417.483649005943	0.988462679393756
	АРКС(1,7)	5.21235879249827	-395.110789702985	26.3777576546261	0.868484999651634	0.116520977714027	0.0174465939194052	-419.249970747103	0.988135284803449
	АРКС(2,10)	5.27660263692478	-304.433127990972	27.7473240552389	0.900712893054402	0.121833809333843	0.0339653249525752	-339.300833943587	0.983623786504026
	АРКС(2,9)	5.27661083674159	-309.115503726773	27.7473392990942	0.900710674882759	0.121833868112305	0.0339652480562116	-341.301078452264	0.983623956716281
	АРКС(2,8)	5.27661145835975	-313.797641246865	27.7473415159954	0.900710546348026	0.121833876660407	0.0339652460771172	-343.301084745231	0.983623970068339
	АРКС(2,7)	5.27661233599368	-318.479790984615	27.747343079594	0.90071030718825	0.121833882689454	0.0339652402556551	-345.301103255857	0.983623985947835
	АРКС(2,5)	5.27661244359053	-327.844056299691	27.7473432580176	0.900710277431368	0.121833883377433	0.033965239355945	-349.301106116685	0.983623987712666
	АРКС(2,6)	5.27661244446065	-323.161924940776	27.7473433031115	0.900710278659575	0.12183388355131	0.0339652393973924	-347.301105984894	0.983623988219339
	АРКС(2,4)	5.2766124537286	-332.526187750193	27.7473432720044	0.900710274505694	0.121833883431365	0.0339652392856943	-351.301106340063	0.983623987926522
	АРКС(2,3)	5.27661245387388	-337.2083189858	27.7473432724749	0.90071027447301	0.121833883433179	0.0339652392830267	-353.301106348545	0.983623987932169

Автокореляційна функція

	АКФ	АКФ Y	АКФ КС
▶	r(1)	0.933818400818...	0.953017166931...
	r(2)	0.872122061119...	0.895736769990...
	r(3)	0.824858290900...	0.836982756758...
	r(4)	0.771955023588...	0.777791411838...
	r(5)	0.701492375780...	0.717661035853...
	r(6)	0.643066318945...	0.671967529710...
	r(7)	0.629596602405...	0.651661044466...
	r(8)	0.632166355124...	0.644698029047...
	r(9)	0.613610889946...	0.634344161886...
	r(10)	0.601825799597...	0.629734952394...
*			

ПРИКЛАДИ РОБОТИ СППР (ВВП УКРАЇНИ У 1996 – 2005 РР.)

Нормовані дані, експоненційне ковзне середнє з вікном 3, прогноз на 1 крок.



ПРИКЛАДИ РОБОТИ СППР (ГРОШОВИЙ АГРЕГАТ МЗ УКРАЇНИ У 1993 – 2007 РР.)

Нормовані дані, експоненційне ковзне середнє з вікном 3, прогноз на 1 крок.

СППР вибору кращої моделі

Автокореляційна функція

Введення даних

Файл з рядом
C:\Users\vmx\Desktop\M3.txt

Вид ковзного середнього
Експоненційне

Вікно ковзного середнього: 3

Прогноз кроків: 1

Нормувати дані

Краща модель

Графік

АРКС(1,4)
 $y(k) = -3,89673978698317E-07 + -0,184549323452863 \cdot y(k-1) + 1,75293162026426 \cdot x(k-1) + -0,554415721334067 \cdot x(k-2) + -0,164678584201347 \cdot x(k-3) + 0,222549608202264 \cdot x(k-4)$

Результати для всіх моделей

Пуск! АКФ

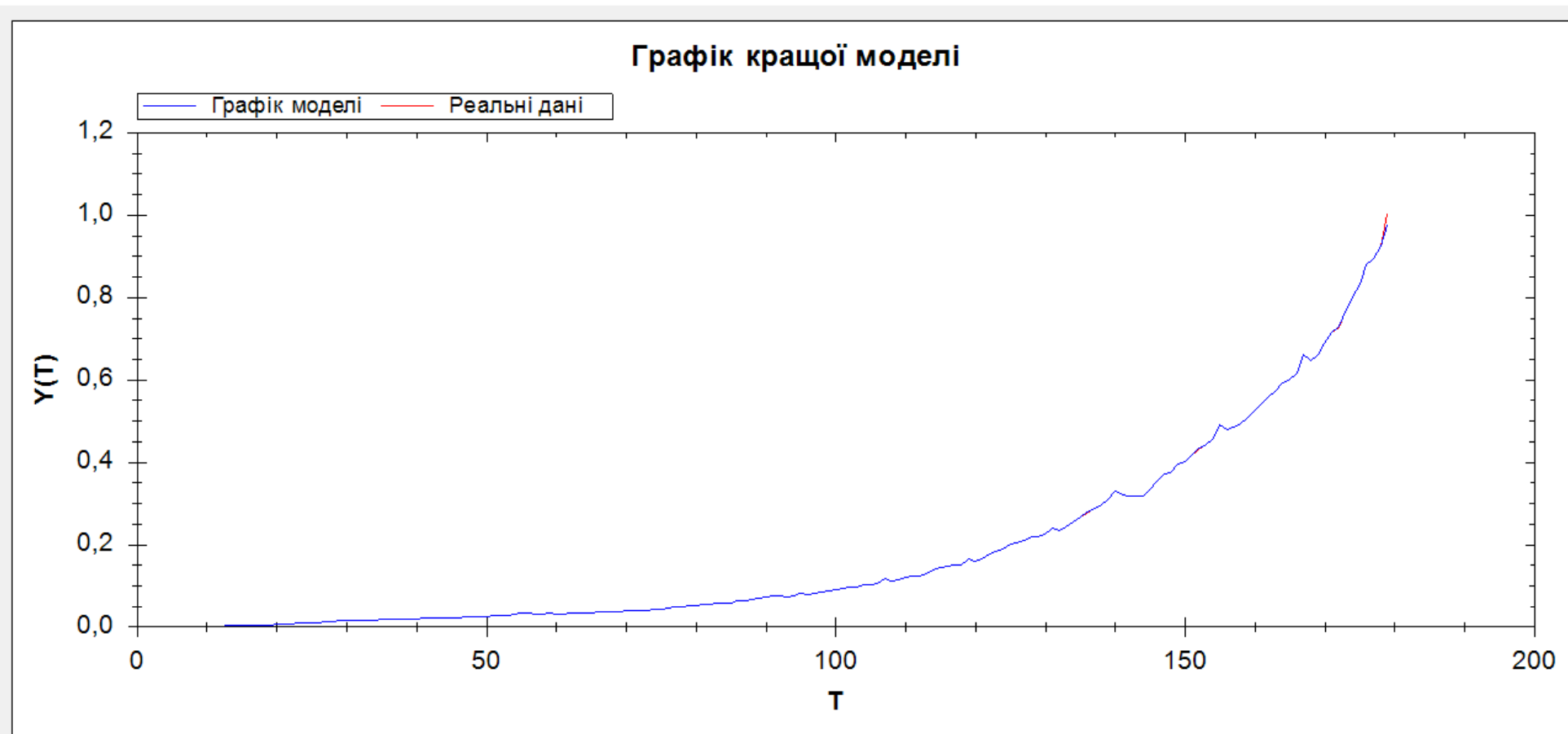
	Модель	IK	BSC	САПП, %	DW	U (коэф. Тейла)	Σe^2	AIC	R ²
▶	АРКС(1,4)	-3,89591065681952	-2518,77407176903	2,60925249383479	2,8581834232471	0,0132187173249004	6,50683568199134E-07	-2537,89838660407	0,999999585086
	АРКС(1,8)	-0,381520512552551	-3244,557646655334	7,56890254437273	2,33329798682095	0,0393330529444081	1,00487412251159E-08	-3276,43150461175	0,9999978663334
	АРКС(1,7)	-0,198173910280377	-3192,33786100125	7,58380417009663	2,11214093774935	0,0394135438411886	1,3848235585121E-08	-3221,02433325381	0,999997192929
	АРКС(1,5)	-0,0395146007048173	-2596,40283049444	9,60168732246474	3,09876267699345	0,0504294769603682	4,09674699272797E-07	-2618,71453113533	0,999996666100
	АРКС(1,9)	0,457160695866365	-3137,0597784647	9,57067381943041	2,17454159672307	0,0502584030064533	1,77967225787494E-08	-3172,12102232895	0,999994471812
	АРКС(1,6)	0,626053953842591	-2883,19439826821	11,0864503868254	2,58200705085612	0,0586853108711703	8,01737448185979E-08	-2908,69348471494	0,999993932899
	АРКС(1,10)	2,14421225601556	-3354,86831680909	8,10166441576189	0,889717647855572	0,0422185236317773	5,12037167812364E-09	-3393,11694647918	0,999993617790
	АРКС(2,6)	2,55233934940312	-3252,11392652705	9,53140861593684	0,919014660583382	0,0500418916666297	9,9166318242636E-09	-3280,80039877962	0,999995451358
	АРКС(2,2)	2,55233974290812	-3272,86348751719	9,53140933450335	0,919014608968639	0,0500418956280436	9,91663083695603E-09	-3288,800416601	0,999995440325
	АРКС(2,1)	2,55233974615903	-3278,05087340866	9,53140934356176	0,919014608875716	0,0500418956779821	9,91663083523736E-09	-3290,80041663202	0,999995440251
	АРКС(2,3)	2,55233974658039	-3267,67610186518	9,53140934130979	0,919014608456808	0,0500418956655671	9,91663083145897E-09	-3286,80041670022	0,999995440343
	АРКС(2,4)	2,55233975761849	-3262,4887157804	9,5314093914938	0,919014610330897	0,0500418959422285	9,91663084691194E-09	-3284,80041642129	0,999995439867
	АРКС(2,5)	2,5523402111054	-3257,30131482885	9,53141044957456	0,919014576275865	0,0500419017753643	9,91663168598716E-09	-3282,80040127557	0,999995425922
	АРКС(2,7)	2,55234115246924	-3246,92649651906	9,5314049563079	0,919013639615242	0,0500418714913157	9,91663427307074E-09	-3278,80035457747	0,999995484606
	АРКС(2,8)	2,55255698818439	-3241,72556436425	9,53144637266037	0,918945092653765	0,0500420998172175	9,91738477179506E-09	-3276,7868082285	0,999994979617
	АРКС(2,9)	2,55614052169224	-3236,29460339123	9,53125532614066	0,917708580852861	0,0500410465898806	9,93088908989033E-09	-3274,54323306132	0,999997435341

	АКФ	АКФ Y	АКФ КС
▶	r(1)	0,965500088781...	0,965594176086...
	r(2)	0,932148974827...	0,932058331108...
	r(3)	0,898705737543...	0,899409424409...
	r(4)	0,867918627054...	0,869019711462...
	r(5)	0,838517295369...	0,840243128453...
	r(6)	0,811038227318...	0,813186451737...
	r(7)	0,785540564925...	0,787559795761...
	r(8)	0,760290490675...	0,762580222123...
	r(9)	0,736244280745...	0,738680819925...
	r(10)	0,713358529128...	0,715299998818...
*			

ПРИКЛАДИ РОБОТИ СППР (ГРОШОВИЙ АГРЕГАТ МЗ УКРАЇНИ У 1993 – 2007 РР.)

Графіки

— □ ×



РЕЗУЛЬТАТИ

Процес	Модель	ІК	Адекватність моделі			Якість прогнозу	
			R^2	$\sum e^2$	DW	САПП, %	U
ІСЦ	АРКС(1,5)	-8,062	0,951	0,043	1,294	0,919	0,004
	АРКС(1,6)	-5,855	0,869	0,112	1,729	2,357	0,011
	АРКС(1,2)	-1,363	0,961	0,049	1,501	9,611	0,05
ВВП	АРКС(1,3)	-3,846	0,988	0,051	2,001	2,691	0,013
	АРКС(1,2)	-2,268	0,991	0,016	1,506	3,677	0,018
	АРКС(1,6)	3,025	0,988	0,045	1,861	24,449	0,108
Грошовий агрегат М3	АРКС(1,4)	-3,895	0,999	6,506E-07	2,858	2,609	0,013
	АРКС(1,8)	-0,381	0,999	1,004E-08	2,333	7,568	0,039
	АРКС(1,7)	-0,198	0,999	1,384E-08	2,112	7,583	0,039

НАУКОВА НОВИЗНА

1. Запропонований і апробований практично інтегральний критерій адекватності моделі та якості прогнозів, який забезпечує вибір кращої моделі.
2. Побудовано нові регресійні моделі для таких процесів: ВВП, ІСЦ та грошовий агрегат МЗ (загальна грошова маса), які дають можливість обчислювати високоякісні короткострокові прогнози.

ПРАКТИЧНА ЦІННІСТЬ

Практична цінність виконаної роботи полягає у можливості розширення функцій програми до комерційної системи підтримки прийняття рішень для побудови прогнозуючих математичних моделей на основі статистичних даних.

ПУБЛІКАЦІЯ

Грицаюк Максим. Автоматизований вибір прогнозуючих моделей в СППР. – Київ: Збірник ІПСА НТУУ «КПІ» Системні науки і кібернетика, 2016. – С. 19-34



ДЯКУЮ ЗА УВАГУ!