

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

Факультет комп'ютерних наук та кібернетики
Кафедра прикладної статистики

**Кваліфікаційна робота
на здобуття ступеня бакалавра**
за спеціальністю 124 Системний аналіз

на тему:

**ПРОГНОЗУВАННЯ МАКРОЕКОНОМІЧНИХ ПОКАЗНИКІВ НА
ОСНОВІ МОДИФІКОВАНИХ АДАПТИВНИХ МОДЕЛЕЙ**

Виконала студентка 4-го курсу
Соколова Тетяна Володимирівна



(підпис)

Науковий керівник:
асистент, кандидат фіз.-мат. наук
Макушенко Ігор Анатолійович



(підпис)

Засвідчую, що в цій роботі немає
запозичень з праць інших авторів без
відповідних посилань.

Студент




(підпис)

Роботу розглянуто й допущено до захисту
на засіданні кафедри прикладної
статистики

« 05 » _____ червня _____ 2023 р.,
протокол № 11

Завідувач кафедри
І. В. Розора



(підпис)

РЕФЕРАТ

Обсяг роботи 42 сторінки, 18 ілюстрацій, 1 таблиця, 17 джерел.

ПРОГНОЗУВАННЯ МАКРОЕКОНОМІЧНИХ ПОКАЗНИКІВ, АДАПТИВНІ МОДЕЛІ. МОДИФІКОВАНІ АДАПТИВНІ МОДЕЛІ, ЕКСПОНЕНЦІЙНЕ ЗГЛАДЖУВАННЯ, ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ, АДАПТИВНІ ПАРАМЕТРИ.

Об'єктом роботи є прогнозування макроекономічних показників на основі модифікованих адаптивних моделей.

Мета роботи: дослідження та аналіз модифікованих адаптивних моделей; реалізація таких моделей.

Результати роботи: розглянуто два типи модифікованих адаптивних моделей: моделі з адаптивними параметрами адаптації та адаптивні комбіновані моделі, виконано порівняння цих моделей як на штучно змодельованих часових рядах з різними характеристиками, так і на реальних даних.

Ці модифіковані адаптивні моделі стали хорошою основою для подальшого розвитку авторегресійних моделей, моделей на основі машинного навчання, гібридних моделей тощо.

Перспективи розвитку модифікованих адаптивних моделей для прогнозування макроекономічних показників полягають у поліпшенні їхньої точності і здатності до адаптації до нових умов. Це може включати в себе використання більш складних архітектур нейронних мереж, врахування більш широкого спектру факторів, таких як соціальні та політичні чинники, і вдосконалення методів тренування і оцінки моделей.

Зміст

| | |
|---|----|
| ВСТУП | 4 |
| 1. УЗАГАЛЬНЕНА МОДЕЛЬ БРАУНА ТА ЗГЛАДЖУВАННЯ | 6 |
| 2. МОДЕЛІ З АДАПТИВНИМИ ПАРАМЕТРАМИ АДАПТАЦІЇ | 9 |
| 2.1 Швидкість реакції як функція слідкуючого контрольного сигналу (Модель Трігга-Ліча) | 9 |
| 2.2 Регулювання параметра адаптації за змінами спектральних характеристик. Модель Рао та Шапіро (спектральна модель) | 15 |
| 2.3 Адаптація параметра методом еволюції | 22 |
| 3. АДАПТИВНІ КОМБІНОВАНІ МОДЕЛІ | 29 |
| 3.1 Адаптивна селективна модель | 29 |
| 3.2 Адаптивна гібридна модель | 32 |
| 4. ПОБУДОВА ТА АНАЛІЗ МОДЕЛЕЙ | 35 |
| ВИСНОВКИ | 43 |
| ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ | 46 |

ВСТУП

Економічне прогнозування характеризує майбутній розвиток, виходячи з гіпотези, що основні фактори і тенденції минулого періоду збержуться на період прогнозу або можна обґрунтувати і врахувати напрямок їхніх змін в перспективі, що розглядається. Надії тут покладаються на інерційність економічних систем. Між тим в більшості випадків мінливість економічних явищ зростає.

Найбільшу інерційність мають макроекономічні характеристики. Для параметрів, що описують процеси, які відбуваються на рівні галузей, підприємств, цехів, характерна велика залежність від місцевих умов. У сучасних умовах дослідник часто має справу з новими економічними явищами з короткими статистичними рядами або зі старими явищами, в яких відбуваються фундаментальні зміни. Застарілі дані під час моделювання часто виявляються безкорисними або навіть шкідливими. До того ж статистичний опис процесу рідко може задовольнити потреби, тому що необхідно знати, не як розвивається процес в середньому, а як буде розвиватись його тенденція, що спостерігається на даний момент часу. Це означає, що потрібно будувати процеси, спираючись в основному на малу кількість найсвіжіших даних. У цьому випадку альтернативою статистичному обґрунтуванню моделі може стати надання їй адаптивних властивостей.

Напрявлений розвиток економічної системи прокладає собі шлях через одиничні явища і умовах зіткнення протиборчих сил і тому вважається стохастичним процесом. Для прогнозування стаціонарних процесів математичний апарат працює доволі успішно [1]. Але економічні процеси, як правило, є нестаціонарними. Чим більший період прогнозу, тим більше можливостей для зміни тенденцій економічного розвитку, особливо в сучасних реаліях. Крім того, в періоді, який розглядають, можуть статися взагалі непередбачувані події, що значно деформуватимуть процес.

У роботі наведений аналіз та порівняння модифікованих адаптивних моделей двох типів: моделі з адаптивними параметрами адаптації та адаптивні комбіновані моделі.

Адаптивні методи можуть застосовуватись для прогнозування показників фондового ринку, грошових потоків, сезонних тенденцій тощо [2]. Інструментом прогнозу в таких методах є модель. Корируючи параметри моделі з кожним кроком на основі нових даних, відбувається адаптація моделі до мінливих умов розвитку досліджуваного явища. Таким чином модель постійно пристосовується до нової інформації.

У багатьох випадках економіко-математичні моделі прогнозування можуть бути корисним інструментом для дослідження. При цьому для збільшення точності прогнозів в мінливих умовах варто використовувати адаптивні методи прогнозування. Вони вирізняються тим, що відображують поточні властивості ряду і здатні неперервно враховувати еволюцію динамічних характеристик процесів. Мета адаптивних методів – побудова економіко-математичних моделей, що можуть самостійно «налаштовуватись», враховуючи умови, що з часом змінюються. Саме тому такі моделі передбачаються насамперед для прогнозування на короткі періоди часу.

1. УЗАГАЛЬНЕНА МОДЕЛЬ БРАУНА ТА ЗГЛАДЖУВАННЯ

Браун [3] розглядав таку модель:

$$x_t = \xi_t + \varepsilon_t,$$

$$\xi_t = a_1 f_1(t) + a_2 f_2(t) + \dots + a_n f_n(t) = \sum_{i=1}^n a_i f_i(t),$$

де ε_t – випадкова неавтокорельована величина з середнім значенням 0 та дисперсією σ_ε^2 [4];

$f_i(t)$ – відомі функції часу, математичні або емпіричні;

a_i – коефіцієнти, які будуть переглядатись та оцінюватись.

Тоді прогнозом буде

$$\begin{aligned} \hat{x}(t + \tau) &= \hat{a}_1(t) f_1(t + \tau) + \hat{a}_2(t) f_2(t + \tau) + \dots + \hat{a}_n(t) f_n(t + \tau) = \\ &= \sum_{i=1}^n \hat{a}_i(t) f_i(t + \tau). \end{aligned}$$

Нехай $x = (x_1, x_2, \dots, x_T)$ вектор рядок з T доступних для обробки даних,

$$f(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}$$

– вектор стовпчик n функцій моделі.

Також \mathcal{F} – матриця з елементів $f_i(t)$ (значення i -тої функції в момент t) розміру $n \times T$. Похибка прогнозу $e(t) = x(t) - a' f(t)$, де

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

вектор коефіцієнтів, що будуть визначатись після обробки даних. Критерієм для визначення є мінімізація зважених квадратів похибок [4]

$$\sum_{i=1}^T \omega_i^2 e^2,$$

де $\omega_{T-j}^2 = \beta^j$, $0 < \beta < 1$, $\beta = 1 - \alpha$ – фактор дисконтування.

Ці ваги – діагональні елементи матриці W розміру $T \times T$, в якій всі інші елементи дорівнюють 0.

Браун визначив, що коефіцієнти, що задовольняють потрібний критерій

$$a' = xW'WFF^{-1},$$

де F – матриця функцій $(FW)(FW)'$ розміру $n \times n$ в стабільному стані.

$$F_{ik}(T) = \sum_{j=0}^{T-1} \beta^j f_i(T-j)f_k(T-j).$$

Він позначив $g(T) = xW'WF$ – вектор, що зберігає всі попередні дані.

Тоді можна оцінити коефіцієнти

$$\hat{a}' = \hat{a}'(T) = g(T)F^{-1}(T),$$

тобто шляхом множення вектору даних на обернену матрицю функцій моделі.

Отже, прогноз можна записати у вигляді

$$x(T + \tau) = \hat{a}'(T)f(T + \tau) = \sum_{i=1}^n \hat{a}_i(T)f_i(T + \tau).$$

Також Браун запропонував матрицю переходу L , для якої справедливе співвідношення $f(t + 1) = Lf(t)$. Проте така матриця існує лише для наборів поліноміальних, експоненційних та тригонометричних функцій.

В такому випадку можна оцінити коефіцієнти моделі за допомогою адаптивного згладжування, запропонованого Р. Ф. Мейером [5], використовуючи відносний час

$$\hat{x}(t + \tau) = f'(\tau)\hat{a}(t).$$

Матриця функцій $F(t)$ складається з елементів

$$F_{ik}(t) = \sum_{j=0}^t \beta_j f_i(-j) f_k(-j),$$

а вектор $g(t)$ має n компонент

$$g_i(t) = \sum_{j=0}^t \beta^j f_i(-j) x(t-j).$$

Якщо доступно принаймні n спостережень, то матриця $F(t)$ додатньо визначена та має обернену. Тоді коефіцієнти у рівнянні прогнозу можуть бути оцінені так:

$$\hat{a}(t) = F^{-1}(t)g(t),$$

з кожним новим спостереженням $g(t)$ та $F(t)$ мають переобчислюватись.

Для визначення оцінок коефіцієнтів моделі Браун вивів n -компонентний вектор

$$h = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} = F^{-1}f(0),$$

за допомогою якого можна оновлювати коефіцієнти за формулою:

$$\hat{a}(t) = L' \hat{a}(t-1) + h e_1(t-1),$$

де $e_1(t-1)$ – похибка прогнозу, зробленого в момент $(t-1)$ на 1 крок вперед.

Далі розглянемо модифікації такого прогнозування.

2. МОДЕЛІ З АДАПТИВНИМИ ПАРАМЕТРАМИ АДАПТАЦІЇ

2.1 Швидкість реакції як функція слідкуючого контрольного сигналу (Модель Трігга-Ліча)

Д. Трігг та А. Ліч [6] запропонували модифікувати прогнозуючі системи, що використовують експоненційне згладжування, шляхом зміни швидкості реакції залежно від величини контрольного сигналу, що визначається співвідношенням [7]:

$$K_t = \frac{\hat{e}_t}{\tilde{e}_t},$$

де $\hat{e}_t = (1 - \gamma)\hat{e}_{t-1} + \gamma e_t$ – згладжена похибка,

$\tilde{e}_t = (1 - \gamma)\tilde{e}_{t-1} + \gamma |e_t|$ – згладжене абсолютне значення похибки,

$0 < \gamma < 1$ – стала згладжування.

В найпростішій моделі це еквівалентно регулюванню параметра згладжування α . Найбільш очевидний спосіб примусити систему автоматично реагувати на розходження прогнозів та фактичних даних – це збільшення α , щоб надати свіжим даним більшу вагу, таким чином забезпечуючи більш швидке пристосування моделі до нової ситуації. Як тільки система адаптувалась, необхідно зменшити величину α для фільтрації шуму.

Простий спосіб досягнення такої адаптивної швидкості реакції полягає в виборі

$$\alpha_t = |K_t|.$$

На рисунку 2.1 показано випробування поліноміальної моделі нульового порядку зі змінним параметром α при прогнозуванні штучного ряду.

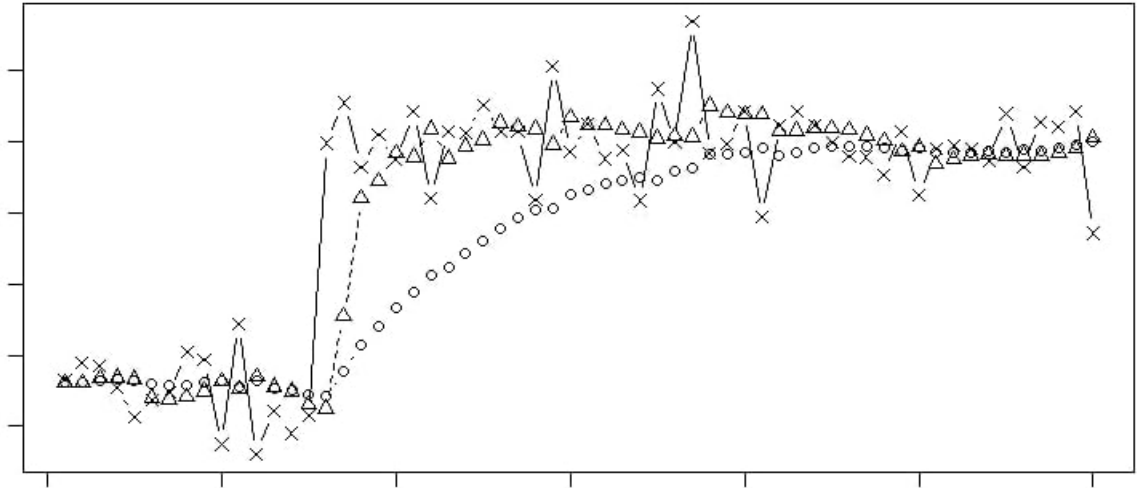


Рисунок 2.1 – Реакції моделей на ступінчасту зміну

Хрестики на рисунку відповідають значенням членів часового ряду, в якому спостерігається зміна ступінчастого типу. Ряд був штучно згенерований за моделлю:

$$x_t = a'_1 + \varepsilon_t, \quad t < t_1;$$

$$x_t = a''_1 + \varepsilon_t, \quad t > t_1;$$

$$a'_1 = \text{const};$$

$$a''_1 = \text{const};$$

$$a'_1 \neq a''_1,$$

де ε_t – неавтокорельовані випадкові нормальні відхилення з нульовим математичним сподіванням і дисперсією σ^2 [8]. Реакція найпростішої моделі експоненційного типу з постійним коефіцієнтом згладжування $\alpha = 0,1$ відмічена кружечками. Пунктирна лінія характеризує реакцію подібної системи, але зі змінним α_t . Можна побачити, що система з адаптивним α пристосовується до ступінчастих змін набагато швидше, а після опрацювання сходинки розмах її

коливань не більший, ніж у звичайної системи, оскільки контрольний сигнал, побудований за принципом згладженої похибки, залишається великим, як правило, тільки поки прогнозуюча система знаходиться в перехідному режимі. Аналогічна модифікація може бути здійснена і для більш складних моделей. Розглянемо частинний випадок узагальненої моделі Р. Брауна – модель лінійного зростання ($n = 2$)

$$\hat{x}_\tau(t) = \hat{a}_{1,t} + \hat{a}_{2,t} \cdot \tau,$$

для якої рівняння оновлення коефіцієнтів будуть:

$$\hat{a}_{1,t} = \hat{a}_{1,t-1} + \hat{a}_{2,t-1} + (1 - \beta^2)e_1(t - 1);$$

$$\hat{a}_{2,t} = \hat{a}_{2,t-1} + (1 - \beta)^2 e_1(t - 1).$$

Із рівнянь видно, що оцінка середнього рівня процесу \hat{a}_1 реагує на похибку прогнозу зі швидкістю $(1 - \beta^2)$. В моделях Брауна з n параметрами швидкість реакції визначається величиною $(1 - \beta^n)$, яку називають еквівалентною сталою згладжування.



Рисунок 2.2 – Порівняння реакції моделей на лінійне зростання

В багатопараметричній моделі Брауна можна цілком природно прирівняти еквівалентну сталу згладжування до модуля контрольного сигналу. У лінійній моделі ми могли б покласти

$$(1 - \beta_t^2) = |K_t|,$$

звідки

$$\beta_t = \sqrt{1 - |K_t|}.$$

Це б означало, що кожний елемент вектора h як функції від β щоразу зазнає відповідних змін. Однак залежність всіх елементів h_t від контрольного сигналу погіршує прогноз та робить його нестійким. Але якщо обмежитись модифікацією тільки першої компоненти h_1 вектора h , то така система призводить до більш стабільних результатів.

На рисунку 2.2 показано ряд з лінійною тенденцією зростання, на який накладені ті самі випадкові дані, що і на рисунку 2.1. В цьому прикладі середньоквадратичне відхилення взято пропорційним до середнього рівню ряду. Реакція, що відповідає прогнозу на один крок вперед звичайної моделі лінійного зростання з еквівалентною сталою $(1 - \beta^2) = 0,1$, відмічена кружечками. Прогнози аналогічної моделі, але з $h_1 = |K_t|$ показані пунктиром.

На рисунку 2.3 лінія, позначена хрестиками, показує ті самі випадкові значення, але з синусоїдальним коливанням з періодом 52 (у році 52 тижні).

Через 15 інтервалів виникає ступінчаста зміна і відмічена кружечками лінія показує реакцію моделі

$$\hat{x}_\tau(t) = \hat{a}_{1,t} + \hat{a}_{2,t}\tau + \hat{a}_{3,t} \sin \frac{2\pi\tau}{52} + \hat{a}_{4,t} \cos \frac{2\pi\tau}{52}$$

при $(1 - \beta^4) = 0,1$, $\tau = 1$.

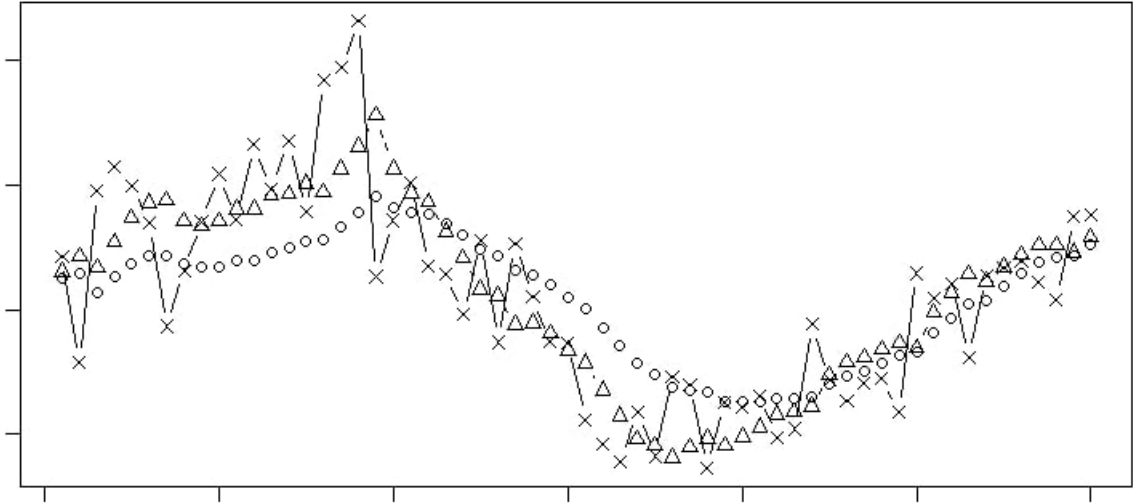


Рисунок 2.3 – Порівняння реакції моделей на синусоїдальні коливання зі стрибком

Пунктирна лінія позначає реакцію тої самої моделі, коли $h_1 = |K_t|$.

Практичне моделювання [9] дало результати, схожі на представлені на рис. 2.1-2.3, тобто прирівнювання h_1 до модуля слідкуючого контрольного сигналу дає значну більшу швидкість пристосування до нового характеру ряду.

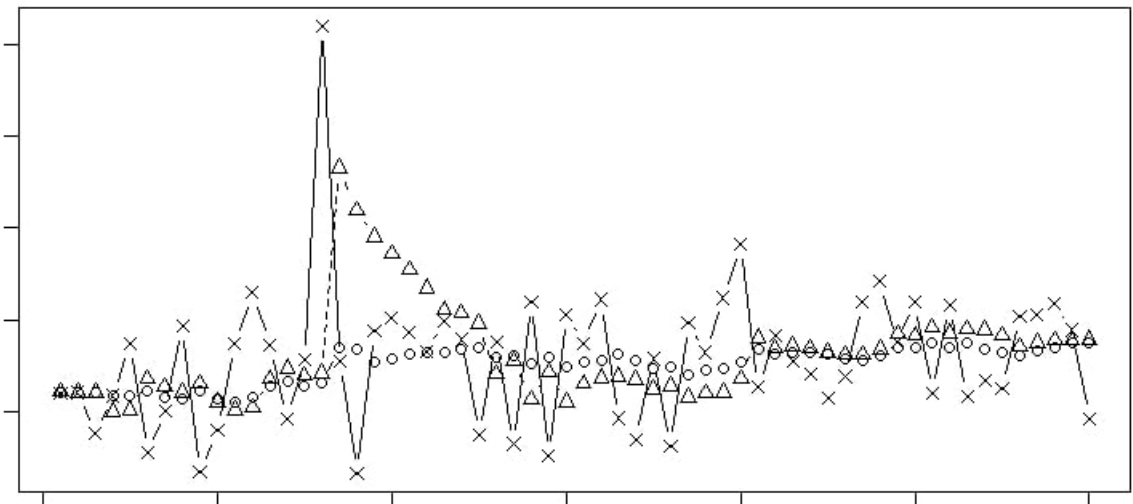


Рисунок 2.4 – Порівняння реакції моделей на одиничний імпульс

Ясно, що не у всіх реальних ситуаціях моделі з адаптивним параметром α дають кращі прогнози, порівняно зі звичайними адаптивними системами. На рисунку 2.4 показані порівняльні реакції цих моделей на великий одиничний імпульс. Система з адаптивною швидкістю показує більше збурення від цього випадкового впливу. М. Л. Шоун [10] пропонує вирішувати цю проблему надмірної реакції на великий одиничний імпульс шляхом затримки використання контрольного сигналу на один крок. Якщо взяти за перший елемент вектора h для даного моменту t модуль контрольного сигналу, підрахований в попередній момент $t - 1$, то вказаний недолік можна подолати. Реакція на сходинку або на лінійний тренд при цьому стане повільнішою.

Метод прирівнювання $h_t = |K_t|$ простий і ефективний. Якщо динаміка ряду не змінюється, то шум фільтрується так само ефективно, як і у звичайному методі зі сталим параметром α . Коли такі збурення все ж виникають, то адаптація проходить швидше.

Цей метод особливо корисний для моделювання рядів з короткою історією або на початковому етапі побудови моделі. Щоб ним скористатись, достатньо провести грубу оцінку коефіцієнтів прогнозуючої моделі. Якщо похибки прогнозів по цій моделі великі, то у контрольного сигналу проявляється тенденція до швидкого абсолютного зростання і система пристосовується з відповідною швидкістю. Цей метод дозволяє обійти проблему визначення оптимального значення α . Натомість виникає задача вибору найкращого значення γ для обчислення контрольного сигналу.

У наведених вище прикладах моделювання величина $\gamma = 0,1$. Якщо необхідна більша обережність по відношенню до швидкості процесу адаптації, то величину γ можна зменшити і взяти $\gamma = 0,05$. Після того, як γ задана, модель працює автоматично. Іноді все ж необхідне втручання людини для коригування

прогнозів, що дає система з адаптивною швидкістю реакції, оскільки не можна очікувати, що система сама впорається з будь-якими змінами в структурі ряду. Крім того, дослідник може знати додаткову цінну інформацію.

Моделювання реакції найпростішої моделі на ступінчасті зміни показало, що хоча модель і досить швидко встановлюється на новому середньому рівні, контрольний сигнал ще деякий час залишається великим. Це заважає системі фільтрувати випадковий шум, що особливо проявляється, коли константа γ , яка використовувалась для отримання згладженої похибки і середнього абсолютного відхилення, досить мала, наприклад, 0,05. Велика γ збільшує дисперсію експоненційної середньої похибки та при цьому погіршує певною мірою можливість досягнення мети, задля якої використовується адаптивна швидкість реакції.

2.2 Регулювання параметра адаптації за змінами спектральних характеристик. Модель Рао та Шапіро (спектральна модель)

Зміни в структурі рядів (гострий пік або ступінчаста зміна) будуть відображатись в їхніх спектральних характеристиках. Основна ідея методу [11] – використати зміни в спектрі для керування величиною α в процедурі обчислення експоненційного середнього. Спектр оцінюється спочатку для відрізка часового ряду з членів $1, 2, \dots, N$, потім для відрізка $2, \dots, N + 1$ і так далі. Таким чином, спектр послідовно оцінюється для відрізків часового ряду постійної довжини, що частково перетинаються. Довжина цих відрізків N має бути достатньо великою, щоб отримувати стійкі оцінки деякої кількості компонент спектра, але не надто великою, інакше момент фундаментальної зміни ряду буде втрачено під час усереднення.

Спектральна щільність випадкового процесу характеризує розподіл всієї енергії стаціонарного ряду як функцію частоти. Математично це можна виразити у вигляді перетворення Фур'є автоковаріаційної функції ряду [4].

Оцінки спектральної щільності як функції частоти отримують так:

$$\hat{S}_x(h\Delta\nu) = \left[\hat{R}_{xx}(0) + 2 \sum_{p=1}^{m-1} \hat{R}_{xx}(p) \cos \frac{\pi p h}{m} + \hat{R}_{xx}(m) \cos \pi h \right],$$

де $h = 0, 1, \dots, m$; $\Delta\nu$ – частота першої гармоніки;

$$\hat{R}_{xx}(p) = \frac{1}{N-p} \sum_{i=1}^{N-p} x_i x_{i+p} - \frac{1}{(N-p)^2} \sum_{i=1}^{N-p} x_i \sum_{i=1}^{N-p} x_{i+p},$$

де m – максимальна величина лагу, для якої обчислюється автоковаріаційна функція;

N – довжина відрізка часу.

Таблиця 2.1. Послідовні спектри

| Відрізок ряду | Спектральні характеристики | | | |
|----------------------|----------------------------|----------------|-----|----------------|
| | $k = 0$ | $k = 1$ | ... | $k = 6$ |
| x_1, \dots, x_{36} | \hat{F}_{10} | \hat{F}_{11} | ... | \hat{F}_{16} |
| x_2, \dots, x_{37} | \hat{F}_{20} | \hat{F}_{21} | ... | \hat{F}_{26} |
| x_3, \dots, x_{38} | \hat{F}_{30} | \hat{F}_{31} | ... | \hat{F}_{36} |
| ... | ... | ... | ... | ... |

Первинні спектральні оцінки потім вирівнюються методом Тьюкі-Хеннінга [1]:

$$\hat{f}_x(h\Delta\nu) = \frac{1}{4} \hat{S}_x[(h-1)\Delta\nu] + \frac{1}{2} \hat{S}_x(h\Delta\nu) + \frac{1}{4} \hat{S}_x[(h+1)\Delta\nu] \text{ для } h = 1, 2, \dots, m-1;$$

$$\hat{f}_x(0) = \frac{1}{2}\hat{S}_x(0) + \frac{1}{2}\hat{S}_x(\Delta\nu);$$

$$\hat{f}_x(m\Delta\nu) = \frac{1}{4}\hat{S}_x[(m-1)\Delta\nu] + \frac{1}{2}\hat{S}_x(m\Delta\nu).$$

В експериментах [9] було обрано довжину відрізка часу $N = 36$ точок, при $m = 6$, тобто всього 7 компонент спектра. Таким чином спочатку оцінюється спектр відрізка від точки x_1 до точки x_{36} , потім від x_2 до x_{37} і так далі.

$\hat{F}_{ik} = \ln \hat{f}_{ik}$ в таблиці 2.1 означає натуральний логарифм оцінки спектральної щільності для k -тої гармоніки на i -ому відрізку ряду. Очевидно, що зміни в структурі ряду будуть представлені змінами в послідовних спектрах, тобто векторами $(\hat{F}_{ik} - \hat{F}_{i+1,k})$.

Для порівняння цих змін потрібні елементи \hat{F}_{ik} , які перед обчисленням різниць бажано вирівнювати. Для вирівнювання використовується ковзне середнє для трьох значень F і відхилення від цього ковзного середнього обчислюється за формулою:

$$\delta_{ik} = \frac{1}{3}(\hat{F}_{i-2,k} + \hat{F}_{i-1,k} + \hat{F}_{ik}) - \hat{F}_{ik}.$$

Позначимо як λ_i максимальну за модулем зміну спектра, що вирівняли за три періоди:

$$\lambda_i = \max|\delta_{ik}|, k = 0, \dots, 6.$$

Якщо величина λ_i мала відносно стандартного відхилення, то це означає, що ніяких значних змін в структурі часового рядку не відбулось. Таким чином, α варто залишати малою. Якщо ж значення λ_i велике, то величину α варто збільшувати до максимуму, тобто до 1.

Для визначення величини α значення λ_i можуть бути використані порізно. Вважалось, що нове значення α має залежати від $(\lambda_i)^2$, щоб зробити модель більш чутливою до змін в структурі рядів.

Вводили величину

$$d_i = b + c \left(\frac{\lambda_i}{\sigma} \right)^2,$$

де σ – стандартне відхилення δ_{ik} ;

b і c – константи.

Величина α визначалась за формулою:

$$\alpha_i = \max[0,1; \min(\exp d_i - 1; 1)].$$

Константи b і c можуть бути визначені з умов:

$$b + r_1^2 c = 0,67; \quad (e^{0,67} - 1) = 0,95;$$

$$b + r_2^2 c = 0,095; \quad (e^{0,095} - 1) = 0,1,$$

де r_1 – значення $\frac{\lambda_i}{\sigma}$, за якого α досягає значення 0,95; r_2 – значення, нижче якого α лишається рівним 0,1. Оскільки $\frac{\delta_{i0}^2}{\sigma^2}, \frac{\delta_{i1}^2}{\sigma^2}, \dots, \frac{\delta_{im}^2}{\sigma^2}$ можна вважати випадковими величинами, асимптотично розподіленими як χ^2 з одним ступенем свободи (тому що, за припущенням, δ_{ik} розподілені асимптотично нормально), автори методу пропонують наступну інтерпретацію r_1 та r_2 .

Для того, щоб знайти таке значення A , при якому $P(\lambda_i < A) = 0,99$, необхідно, щоб $n \cdot P(|\delta_{ik}| > A) = 0,01$, де n – кількість частот в розкладі (в нашому випадку $n = 7$). Звідси $P(|\delta_{ik}| > A) = 0,0014$, або $P\left(\frac{\delta_{ik}^2}{\sigma^2} > \frac{A^2}{\sigma^2}\right) = 0,0014$.

Але $\frac{\delta_{ik}^2}{\sigma^2}$ асимптотично розподілені як χ^2 з одним ступенем свободи. Звідси

$$\frac{A^2}{\sigma^2} = 10,2, \text{ або } \frac{A}{\sigma} = 3,1.$$

Відповідно, $r_1 = 3,1$ означає, що α_i прирівнюється до 0,95, коли λ_i приймає таке значення, ймовірність перевищення якого за умови, що докорінних змін ряду не відбулось, дорівнює 0,01, тобто отримання настільки великих λ_i за відсутності змін в динаміці ряду малоімовірне. Тому можна припустити, що ряд зазнав суттєвих змін і потрібно забезпечити швидку адаптацію моделі.

Інтерпретацію r_2 можна отримати аналогічним чином. Лишається визначити стандартне відхилення, необхідне для отримання d_i .

Автори цього методу отримали наступну наближену оцінку дисперсії:

$$D(\delta_{ik}) = \sigma^2 = \frac{Km}{N} \left(\frac{2}{3} - \frac{2}{9} R_p^2 - \frac{4}{9} R_{2p}^2 \right) (1 + q_k),$$

де K – константа, що залежить від специфічного фільтру;

$$q_k = \begin{cases} 1, & k = 0; \\ 0, & k \neq 0; \end{cases}$$

R_p^2 – когерентність.

Когерентність на частоті $k\Delta\nu$ є показником кореляції між відповідними частотними складовими двох відрізків ряду $\{x_j, \dots, x_{j+N-1}\}$ і $\{x_l, \dots, x_{l+N-1}\}$ по N членів у кожному.

Для стаціонарних процесів [8] когерентність практично не залежить від частоти, а от j та l поєднує тільки рівняння $p = |j - l|$. Для R_p^2 автори вивели наступне просте співвідношення:

$$R_p^2 = \left(1 - \frac{p}{N}\right)^2.$$

Для розглянутого ними прикладу $N = 36$, $m = 6$, $K = 0,75$, $p = 1$. Звідси

$$D(\delta_{ik}) = \sigma^2 = 0,0075(1 + q_k).$$

Таким чином, алгоритм обчислень повністю визначений.

Проте застосування спектральної моделі пов'язане ще з однією складністю, яку можна проілюструвати наступним прикладом. Нехай x_1, \dots, x_{100} – часовий ряд продажів. Припустимо, що цей ряд характеризується додатнім ступінчастим стрибком від x_{60} до x_{100} , тобто $M(x_1) = \dots = M(x_{49}) < M(x_{50}) = \dots = M(x_{100})$, і що оцінка спектра S_i обчислюється для послідовних відрізків ряду

$$x_{i-35}, \dots, x_i, i = 36, \dots, 100$$

(S_i – функція частоти). Ніяких значних змін в спектрі не має бути. Утім S_{49} та S_{50} повинні суттєво відрізнитись, оскільки перше значення не містить ступінчасту зміну, а друге містить. Деякі відмінності можуть також існувати для кількох наступних значень спектра. Далі спектр практично не змінюється. Хоча між S_{84} і S_{85} знову буде суттєва різниця, бо S_{84} включає ступінчастий стрибок. Таким чином, значення α_{85} буде великим, навіть якщо структура ряду в точці 85 не змінилась.

Може здатись, що найпростіший спосіб уникнути вказаного недоліку полягає в тому, щоб в програмі передбачити автоматичне встановлення $\alpha = 0,1$ через 36 точок (в даному прикладі) після того, як виявлено зміну структури ряду. Однак цей шлях може бути небезпечним. По-перше, в даній точці може статись справжня зміна, яка залишиться непоміченою. По-друге, може статись серія проміжних стрибків у рівні продажів між першою і 36 точками.

Ця проблема була вирішена оцінкою двох спектрів: для 36-точкових та 18-точкових відрізків часового ряду. Для кожного з цих спектрів визначалось

відповідне α_i . Менше з цих значень використовувалось в якості поточного значення коефіцієнту згладжування.

Авторами методу пропонується також інше, простіше рішення даної проблеми. Перед підрахунком спектральних оцінок початок 36-точкового відрізка перетворюється шляхом заміни x_t , $t = 1, \dots, 6$, на $x'_t = x_t \sin \frac{2\pi(t-1)}{24}$.

Автори спектральної моделі бачать обмеженість свого методу і вважають, що він може бути застосований, ймовірно, для невеликої кількості рядів. Вони рекомендують цей метод тільки для процесів із суттєвим коливанням. Проте для таких рядів ефект затримки реакції моделі буде особливо впливати на точність прогнозування. У такому випадку бажано використовувати «наївну» модель $x_t(t) = x_t$.

Рао і Шапіро вважають, що застосування спектральної моделі є особливо ефективним для визначення типу змін у структурі ряду. Нехай $k\Delta\nu$ позначає частоту, при якій $\lambda_i = |\delta_{ik}|$. Експерименти показують, що коли зміни в структурі ряду мають ступінчастий характер, то k приймає малі значення, тобто максимальні зміни спектру відбуваються в низьких частотах.

Коли зміни мають форму піку, значення k відповідає високим частотам. Лінійне зростання викликає максимальні відхилення в середині частотного діапазону.

Ці особливості дають можливість використовувати зміни у спектрі для введення суттєвих змін у структуру моделі, а не тільки для регулювання параметра згладжування.

2.3 Адаптація параметра методом еволюції

Інший спосіб модифікації однопараметричної експоненційної моделі з корекцією коефіцієнта лінійного тренду запропонований У. Чоу [12]. Він так само відмовляється від пошуку одного оптимального значення параметра згладжування і пропонує процедуру, що забезпечує адаптацію цього параметра до змін динаміки ряду. Метод комбінує експоненційне згладжування [3] з методом еволюційного планування.

Припускається, що використовуються три оцінки наступного члену ряду, які отримують за допомогою трьох різних значень параметра α : високого, нормального, низького. Ці значення спершу вибираються випадково, а потім модифікуються протягом руху моделі за часом. Оцінка, отримана при нормальному значенні параметра, вважається прогнозом, інші дві оцінки – контрольні. Після отримання нової фактичної точки можна зробити висновок про те, яке з трьох значень параметра дає найменшу похибку. Тепер це значення параметра розглядається вже як нормальне. Прогноз на наступний період робиться саме на основі даного найкращого значення. Одночасно розглядаються контрольні оцінки майбутнього спостереження по більш високому та більш низькому (в порівнянні з нормальним) значенням параметра.

У своїх експериментах Чоу використовував значення параметрів

$$0,05 \leq \alpha - h, \alpha, \alpha + h \leq 0,95,$$

де $h = 0,05$ – крок зміни (еволюції) параметра;

α – нормальне значення параметра;

$(\alpha - h)$ – низьке значення параметра;

$(\alpha + h)$ – високе значення.

В якості початкового значення він брав $\alpha = 0,1$ (найпоширеніше на практиці). Всього було проаналізовано 60 рядів місячних даних про угоди на різні види продукції. Дані представляли собою різноманітні приклади поведінки економічних часових рядів, зокрема циклічний рух. У 59 випадках цей метод показав переваги над стандартною процедурою і в одному випадку результати були майже однакові. Узагальнюючи, Чоу робить висновок, що точність його схеми (за відносною похибкою прогнозу) більш ніж на 10% вища, ніж у стандартного методу. Це особливо проявляється при моделюванні нестационарних рядів зі змінними автокореляційними властивостями.

Ідея Чоу була розвинута в роботі С. Робертса та Р. Ріда [13], в якій розглянуто ще і двопараметричні та трипараметричні моделі. Адаптаційна оптимізація моделі тут також проводиться методом еволюційного планування. В даній схемі реалізується безліч дослідів при різних комбінаціях значень параметрів згладжування, заданих із розрахунку по три значення на кожний параметр. Ці досліді утворюють так званий повний факторний експеримент з додаванням центральної точки. Всього при k керованих параметрах у кожний момент часу обчислюється $2k + 1$ оцінок майбутнього спостереження, і одна з них (та, що відповідає центральній точці плану) вважається прогнозом.

Після отримання фактичної точки аналізується точність пробних оцінок: аналізуються похибки, які є функціями розглянутих параметрів. В результаті можна зробити висновок про перенесення центральної точки в точку з мінімальною похибкою. Прогноз у наступний момент часу здійснюється при значеннях параметрів згладжування, які визначаються центральною точкою. Внаслідок перенесення центральної точки змінюється набір високих і низьких рівнів параметрів. При цьому для стабільності прогностичної системи накладають обмеження, щоб значення параметрів знаходились в межах від 0,05 до 0,95.

Можна провести аналогію цієї схеми з біологічною еволюцією. Незначні зміни незалежних параметрів можна вважати мутаціями. Відбір кращих значень параметрів є аналогією до процесу природного відбору. Звідси і назва методу – еволюційне планування.

У якості базової моделі автори схеми взяли трипараметричну модель Уінтерса [14]. Це модель з адитивним коефіцієнтом зростання та мультиплікативною сезонністю; більш прості моделі було отримано шляхом накладення на неї відповідних обмежень.

Розглядали три випадки:

- 1) Один параметр α . Відсутність тенденції зростання та сезонності.
- 2) Два параметри α_1, α_2 .
 - а) Присутня тенденція зростання та відсутня сезонність;
 - б) Присутня сезонність та відсутня тенденція зростання.
- 3) Три параметри $\alpha_1, \alpha_2, \alpha_3$. Присутня і сезонність, і тенденція зростання. Повна модель Уінтерса.

Однопараметрична система. Оцінку майбутнього члена часового ряду отримують за трьома значеннями єдиного параметра α .

Двопараметрична система. Розглянемо випадок 2.а. Результати у випадку 2.б аналогічні. Система має два параметри α_1, α_2 . Оцінку майбутнього члена отримують відповідно до плану факторного експерименту: по два значення на кожен параметр плюс центральна точка.

Трипараметрична система. Система має три параметри $\alpha_1, \alpha_2, \alpha_3$, для яких план факторного експерименту, будується аналогічним чином. Всього точок $2^3 + 1$.

На основі статистичних критеріїв не можна встановити значну перевагу розглянутої схеми над звичайною моделлю Уінтерса. Проте для рядів, що характеризуються високою автокореляцією, запропонована модель дає більш точні прогнози, ніж модель Уінтерса. З урахуванням того, що в економічних рядах часто присутня автокореляція [2], ця модель в багатьох випадках є корисною. Крім того, реакція на стандартні вхідні потоки – одиничний імпульс, ступінчаста зміна рівня, лінійно-наростаюча функція – в даній схемі краща (з точки зору прогнозів), ніж у моделі Уінтерса, і її варто застосовувати, коли дано недостатньо точне представлення структури досліджуваного ряду.

Інший прийом адаптаційної оптимізації – симплекс-планування – використовується в адаптивних моделях прогнозування Д. Монтгомері [15]. Цей метод дає можливість заздалегідь визначити чіткі правила прийняття рішень про те, куди та коли рухатись з даної точки параметричного простору. На відміну від методу еволюційного планування, що потребує обчислення $2^k + 1$ значень в кожній точці (k – кількість керованих методів), для методу симплекс-планування потрібно в кожній точці робити лише один прогноз. В цьому методі використовуються лінійні симплекс-плани. Симплекс – найпростіша фігура. На площині – це трикутник, у тривимірному просторі – тетраедр і так далі.

Розглянемо правильний симплекс S_0 з вершинами d_1, \dots, d_{k+1} і центром c_0 . На кожній грані симплекса S_0 можна побудувати новий симплекс S_j з центром c_j і k вершинами $d_1, \dots, d_{j-1}, d_{j+1}, \dots, d_{k+1}$, що належать множині S_0 і одною новою вершиною d_j^* , яка є дзеркальним відображенням точки d_j відносно грані, спільної для обох симплексів.

Щоб знайти ту чи іншу координату точки d_j^* , потрібно взяти подвоєне середнє відповідних координат точок $d_1, \dots, d_{j-1}, d_{j+1}, \dots, d_{k+1}$, і відняти від

нього відповідну координату точки d_j . У векторному позначенні це записується так:

$$d_j^* = \frac{2}{k} (d_1 + \dots + d_{j-1} + d_{j+1} + \dots + d_{k+1}) - d_j.$$

Тепер припустимо, що кожна точка d_i симплексу характеризує значення параметрів згладжування і кожній точці після накопичення результатів (в кожен момент часу зробиться лише один прогноз) поставлена у відповідність абсолютна похибка прогнозу $|e_i|$. Звідси слідує, що до початку ітеративної процедури необхідно $k + 1$ кроків для отримання перших прогнозів і визначення їхніх похибок у вершинах початкового симплекса. Припустимо, що $|e_p|$ – похибка в точці d_p – найбільше значення серед $k + 1$ значень, отриманих для симплекса. Тоді зрозуміло, що потрібно рухатись в напрямку точки d_p^* . Рух з центра симплекса за грань, що знаходиться проти точки, буде збігатись з напрямком крутого зниження, розрахованого за результатами спостережень у вершинах правильного симплекса.

Стратегія симплекс-планування, відповідно до якої здійснюється еволюція параметрів згладжування, сформульована у трьох простих правилах:

- 1) Відібрати найбільше значення $|e_p|$ зі значень $|e_1|, \dots, |e_{k+1}|$ заміряних у точках, що утворюють симплекс S_0 . Перейти до нового симплекса S_p , замінивши в симплексі S_0 точку d_p , що відповідає $|e_p|$, точкою d_p^* . Обчислити прогноз на наступний момент часу, використовуючи значення параметрів, що визначаються точкою d_p^* .
- 2) Якщо застосування першого правила призводить до того, що одна i -та точка (вершина) збереглась у $k + 1$ послідовних симплексах, то слід припинити застосування цього правила і розрахувати прогноз на наступний

момент часу, використовуючи значення параметрів, що визначені точкою d_i . Потім застосовувати перше правило.

- 3) Якщо значення $|e_p|$ було найбільшим у попередньому симплексі, а значення $|e_p^*|$ виявилось найбільшим у наступному симплексі, то слід рухатись з останнього симплекса, не повертаючись до попереднього, і при цьому відкинути (замінити її на дзеркальну) точку, якій відповідає значення похибки, найближче до найбільшого.

Таким чином, в цьому методі чітко вказано, коли і куди рухатись, що дозволяє повністю автоматизувати процес керування параметрами.

Д. Монтгомері експериментував з двопараметричною моделлю, що враховує лінійний тренд, і трипараметричною моделлю, аналогічною до моделі Уінтерса.

Початковий симплекс з довільною орієнтацією і довжиною ребра зазвичай задають матрицею D , рядки якої визначають вершини симплекса. Матриця D може бути побудована з останніх k стовпчиків матриці $\sqrt{k+1} \cdot H$, де H – будь-яка ортогональна матриця розмірності $(k+1) \cdot (k+1)$, елементи першого стовпчика якої рівні.

У двопараметричній моделі, довільно покладаючи $\alpha_1 = 0,100$ і $\alpha_2 = 0,075$, Монтгомері використовував матрицю:

$$D = \begin{matrix} & \alpha_1 & \alpha_2 \\ \begin{bmatrix} 0,100 & 0,075 \\ 0,168 & 0,093 \\ 0,118 & 0,143 \end{bmatrix} \end{matrix}$$

Довжина ребра в цьому конкретному випадку $l = \left[(d_i - d_j)' (d_i - d_j) \right]^{\frac{1}{2}} = 0,07$. Верхня і нижня межі зміни параметрів тут також були встановлені 0,95 та 0,05 відповідно.

У трипараметричній моделі, покладаючи $\alpha_1 = 0,100$, $\alpha_2 = 0,075$, $\alpha_3 = 0,100$, як початковий він використовував симплекс, визначений матрицею:

$$D = \begin{array}{c} \alpha_1 \quad \alpha_2 \quad \alpha_3 \\ \left[\begin{array}{ccc} 0,100 & 0,075 & 0,100 \\ 0,166 & 0,092 & 0,117 \\ 0,117 & 0,141 & 0,117 \\ 0,117 & 0,092 & 0,166 \end{array} \right] \end{array}$$

Тут теж $l = 0,07$, $0,095 \geq \alpha_1, \alpha_2, \alpha_3 \geq 0,05$.

Метод, що був розглянутий, є альтернативою методу Роберта і Ріда. Його перевага в тому, що він набагато більш економічний з точки зору обчислень.

3. АДАПТИВНІ КОМБІНОВАНІ МОДЕЛІ

3.1 Адаптивна селективна модель

Цілком логічно припускати, що реальний процес час від часу переживає певні докорінні зміни. Змінюються рівень та динамічні властивості ряду. Причому на одних ділянках зберігається приблизно постійний рівень, а на інших ряд зростає з певною швидкістю або характеризується появою прискорення. Тому будь-які спроби підібрати одну прогнозну модель для всього ряду призводить до деякої усередненої моделі з надмірними дисперсіями параметрів, що оцінюються, та похибками прогнозу. Коли зміна структури відбувається різким стрибком, досліднику доволі просто відкинути застарілі дані та будувати моделі тільки по свіжій інформації. Однак, зазвичай, зміни основних властивостей ряду відбуваються не одразу, а неперервно. Причому ряд змінює тенденцію, рухається зигзагами, тому буває складно, а при одночасному дослідженні великої кількості рядів навіть неможливо, швидко аналізувати таку динаміку та змінювати одну модель на іншу.

Ю. П. Лукашин [9] пропонує дві процедури адаптації структури моделі, які можуть виконуватись програмно. Одна з них базується на принципі неперервної селекції.

Припустимо, що розглядається процес, який по черзі генерується моделлю з постійним рівнем, моделлю з лінійним зростанням та квадратичною моделлю. Враховуючи це, побудована адаптивна комбінована модель (АКМ), яка включає в свій базовий набір три моделі: експоненційного середнього, лінійного зростання та квадратичну.

Обчислення майбутніх значень виконується по кожній з них окремо, але в якості прогнозу вибирається величина, що отримана за моделлю, яка найкраще описує реальний процес на даному проміжку часу. Найкраща модель обирається

відповідно до заданого критерія селекції. Найкращою є та модель, яка дає мінімальну абсолютну похибку прогнозу поточного члена ряду при заданому періоді лагу τ .

Коли автор розглядав реакції на ступінчасту зміну, то до ступінчастого стрибка і через $\tau - 1$ одиниць часу після його появи всі три моделі з базового набору предикторів дають однакові τ -прогнози. Це призводить до того, що складно обрати прогнозуючу модель. У такому випадку для більшої визначеності можна взяти модель лінійного зростання в якості компромісу між моделлю ковзного середнього та квадратичною моделлю. Коли стає зрозуміло, що квадратична модель дає результати, ближчі до реальних точок, АКМ переключасться на неї. В подальшому через помітне перевищення цією моделлю рівня ступінчастого стрибка АКМ переходить на модель лінійного зростання та модель експоненційного середнього, але цей перехід відбувається не одразу, а з лагом $\tau = 3$. Така затримка погіршує загальний результат. Ясно, що найбільш ефективно в даному випадку критерій селекції має працювати при малих τ .

В розглянутому прикладі АКМ випробовувалась на вхідному потоці даних, що містять детермінований ступінчастий стрибок. Якщо ж провести такий самий експеримент в більш життєвих умовах, при наявності шуму, то настільки простий критерій селекції буде не завжди задовольнятися, оскільки через випадкову складову найкращий результат буде то в одній, то в іншій моделі. Таким чином, не можна в загальному випадку виходити тільки з останньої похибки прогнозування. Необхідно враховувати деяку їхню множину. Існує доволі багато варіантів побудови критерія селекції. Розглянемо два приклади таких критеріїв.

Критерій К. Перехід на дану модель здійснювати тоді, коли K її останніх прогнозів є найкращими порівняно з прогнозами за іншими моделями з базового набору АКМ.

Критерій В. Перехід на дану модель здійснювати тоді, коли її експоненційно згладжений квадрат похибки прогнозу B мінімальний порівняно до аналогічного показника для інших моделей з базового набору.

Критерій B формується так:

$$B_t = (1 - \alpha_B)B_{t-1} + \alpha_B e_t^2(t - \tau),$$

де $0 < \alpha_B \leq 1$ – параметр згладжування; $e_t(t - \tau)$ – похибка прогнозу, зробленого в момент $t - \tau$ на τ кроків вперед.

Легко бачити, що при $\alpha_B = 1$ критерій B еквівалентний критерію K при $K = 1$.

Якщо в моделі експоненційного типу параметр згладжування часто буває близьким до 0, то тут цього сказати не можна. Параметр B характеризує інерційність переходу, а, як було видно, затримка переходу може призвести до поганої роботи АКМ, порушити відповідність структури моделі до динаміки процесу. Те саме стосується K .

Власне, величину B_t можна назвати оцінкою поточного значення дисперсії похибки прогнозу на τ одиниць часу вперед. Але враховуючи, що α_B зазвичай досить велике, згладжування квадратів похибок буде недостатнім. Тому надавати величині B такий сенс в загальному випадку було б не виправданим. За малих значень параметра згладжування ця інтерпретація є допустимою і аналогічна процедура може використовуватись для грубої оцінки довірчих рівнів прогнозів при прогнозуванні нестационарних процесів.

Принцип неперервної селекції досить простий. Але у зв'язку з тим, що перехід з моделі на модель залежить від предикторів, що входять в базовий набір АКМ, та специфічних динамічних властивостей ряду, загальний теоретичний

аналіз ефективності АКМ є складною задачею, яку можна спробувати розв'язати експериментально.

3.2 Адаптивна гібридна модель

Адаптивна селективна модель розрахована на вибір одного предиктора з деякої їх множини. Але очевидно, що такий вибір однозначно та ефективно можна здійснити, тільки якщо моделі в базовому наборі суттєво відрізняються. Для тих випадків, коли в АКМ входять моделі, які дають відносно близькі результати, і селекція ускладнюється, то можна запропонувати гібридну АКМ [9], прогноз за якою є зваженою сумою прогнозів, отриманих за всіма предикторами. Ваги прогнозів ω_i пропонується взяти адаптивними, обернено пропорційними до $B_{i,t}$ (обчислюється так само, як і в селективній АКМ):

$$\omega_{i,t} = \frac{g_t}{B_{i,t}}.$$

Коефіцієнт пропорційності g_t визначається з рівності суми ваг одиниці. Наприклад для базового набору з трьох моделей ваги будуть визначатись так:

$$\omega_{1,t} = \frac{B_{2,t}B_{3,t}}{B_{2,t}B_{3,t} + B_{1,t}B_{3,t} + B_{1,t}B_{2,t}};$$

$$\omega_{2,t} = \frac{B_{1,t}B_{3,t}}{B_{2,t}B_{3,t} + B_{1,t}B_{3,t} + B_{1,t}B_{2,t}};$$

$$\omega_{3,t} = \frac{B_{1,t}B_{2,t}}{B_{2,t}B_{3,t} + B_{1,t}B_{3,t} + B_{1,t}B_{2,t}}.$$

Отриманий в даному випадку прогноз буде ближчим до результату, що отримується то за одною, то за іншою моделлю, що робить його деякою адаптивною рівнодійною.

На відміну від звичайної АКМ, гібридна модель здійснює перехід більш плавно, з великою кількістю проміжних положень. В результаті виходить неперервний спектр можливих конструкцій прогнозу.

Базовий набір АКМ формується дослідником або автоматично з будь-якого поєднання предикторів; доцільно брати 3-4 моделі.

Автоматичне формування базується на неперервному порівнянні середніх квадратів похибок прогнозів, отриманих за різними моделями до поточного моменту часу. Погані моделі можуть випадково наближатись до реального процесу і короткочасно давати хороші прогнози, що призводить до переходу АКМ саме на них. В результаті з'являються великі похибки, що погіршує ефективність АКМ. Такі моделі краще одразу вивести з базового набору. Тому якщо базовий набір формується автоматично, то з програмного набору моделей в нього включаються лише ті моделі, середні квадрати похибок яких не більше ніж в m разів перевищують мінімальну з них:

$$D_i(\tau) \leq mD_{min}(\tau).$$

Зазвичай $m \in [1,2; 1,5]$. Причому менше значення m відповідає малим τ , оскільки при збільшенні періоду затримки дисперсія середніх квадратів похибок різних моделей, як правило, збільшується.

При автоматичному формуванні базовий набір не є раз і назавжди визначеним. Якщо з плином часу змінюється співвідношення між середніми квадратами похибок, то змінюється відповідно і набір.

Таким чином, при прогнозуванні модель піддається двом випробуванням. Одне, засноване на аналізі середніх квадратів попередніх похибок, дає відповідь на питання про доцільність включення моделі до базового набору предикторів. Друге, шляхом порівняння поточних властивостей моделей за критерієм B ,

визначає ту модель з базового набору, від якої можна очікувати найкращий прогноз.

4. ПОБУДОВА ТА АНАЛІЗ МОДЕЛЕЙ

Для аналізу та реалізації моделей було взято дані про ВВП США [16] (рис. 4.1-4.7) та ВВП України [17] (рис. 4.8-4.14) з відкритих джерел Світового банку.

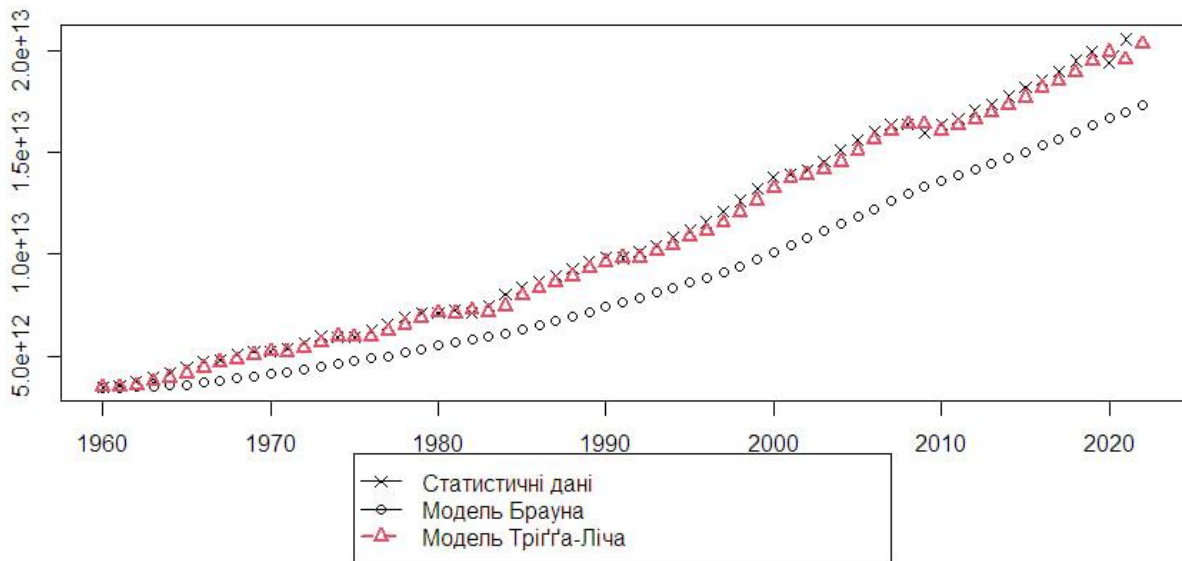


Рисунок 4.1 – Порівняння прогнозів моделі Брауна ($\alpha = 0.1$) та моделі Трігґа-Ліча ($\tau = 1$)

Можна побачити, що найпростіша модель Брауна значно відстає за швидкістю реакції, в той час як модель з адаптивним параметром адаптації Трігґа-Ліча досить успішно підлаштовується під зміну ряду статистичних даних ВВП США (рис. 4.1).

Модель Рао-Шапіро, адаптивна селективна та гібридна моделі також показують набагато кращі результати за модель зі сталим параметром адаптації (рис. 4.2-4.5). Порівнюючи похибки на рис. 4.6-4.7 можна сказати, що адаптивна селективна модель у даному випадку найкраще прогнозує обраний макроекономічний показник.

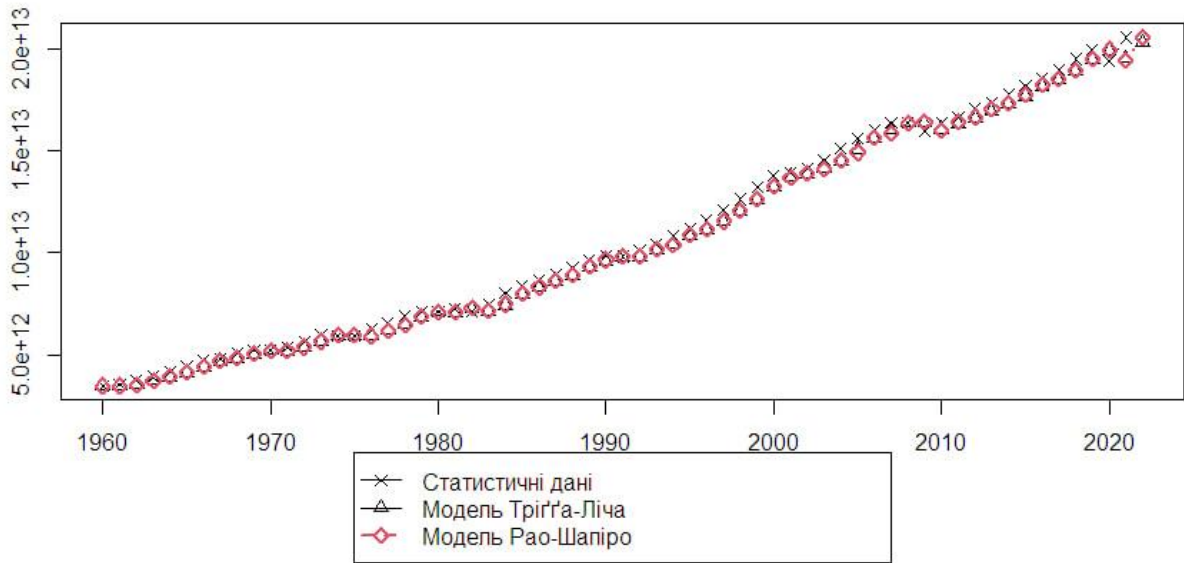


Рисунок 4.2 – Порівняння прогнозів моделі Трігґа-Ліча ($\tau = 1$) та моделі Рао-Шапіро ($N = 10, m = 5$)

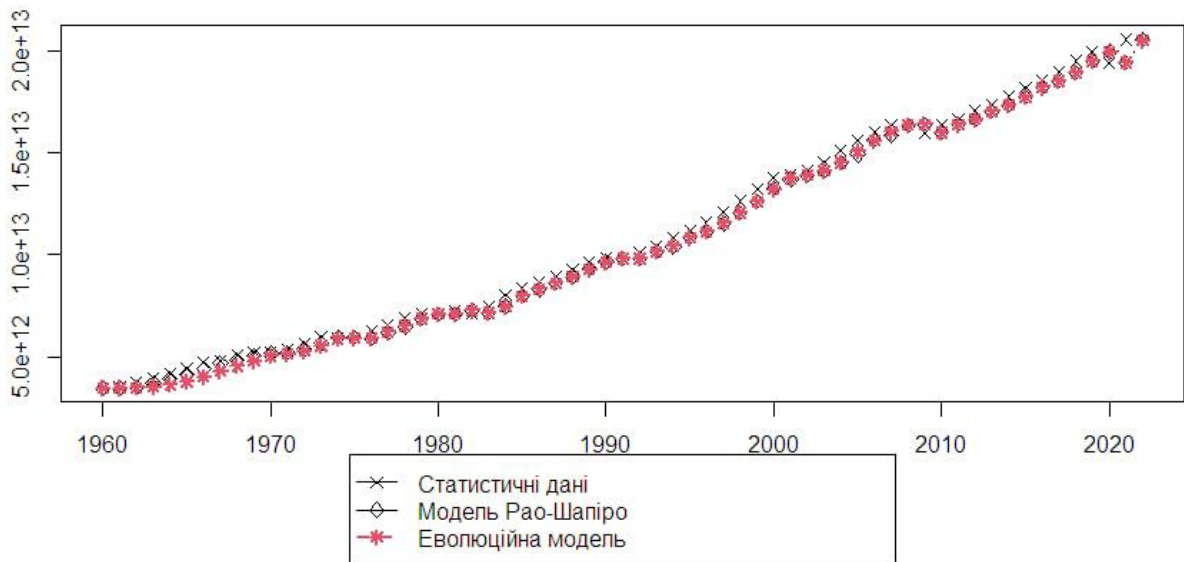


Рисунок 4.3 – Порівняння прогнозів моделі Рао-Шапіро ($N = 10, m = 5$) та Еволюційної моделі

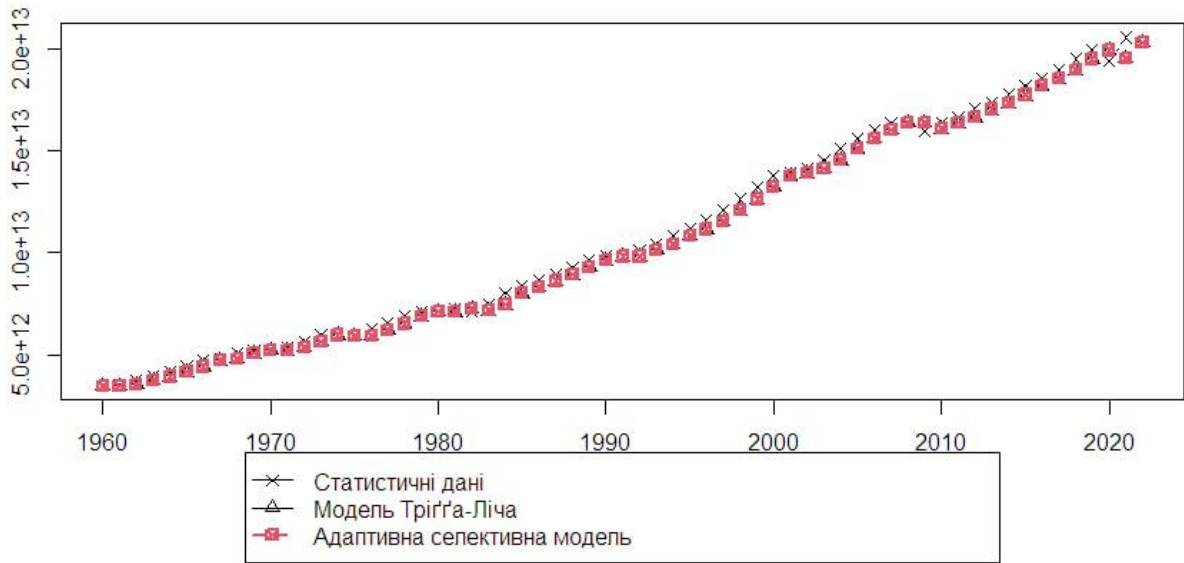


Рисунок 4.4 – Порівняння прогнозів моделі Трігга-Ліча ($\tau = 1$) та адаптивної селективної моделі (2 моделі в базовому наборі)

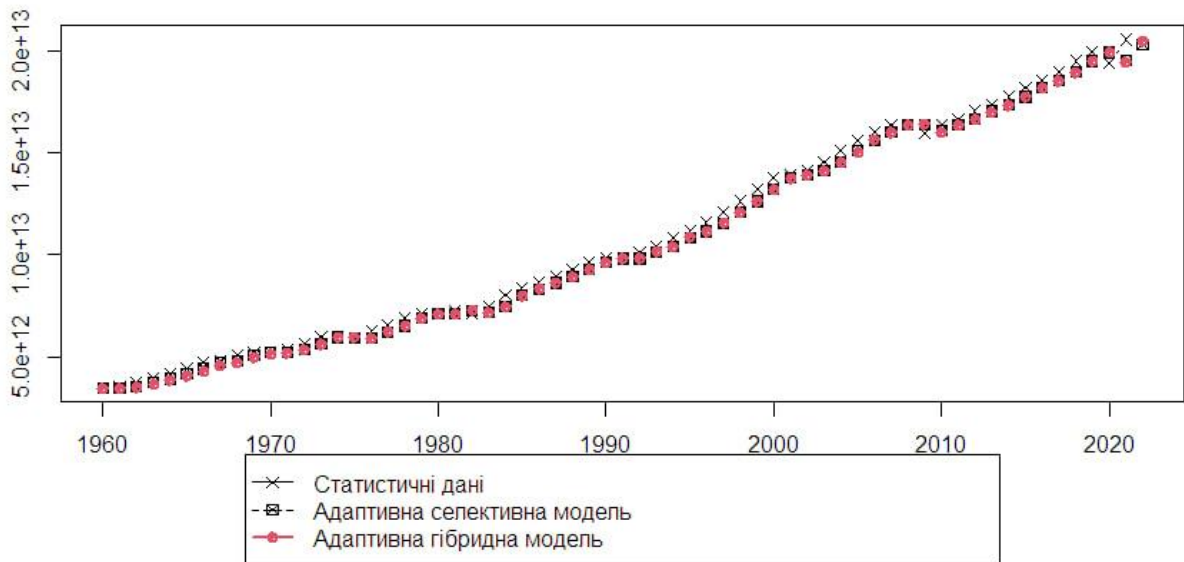


Рисунок 4.5 – Порівняння прогнозів адаптивної селективної моделі (2 моделі в базовому наборі) та адаптивної гібридної моделі (3 предиктори в наборі)

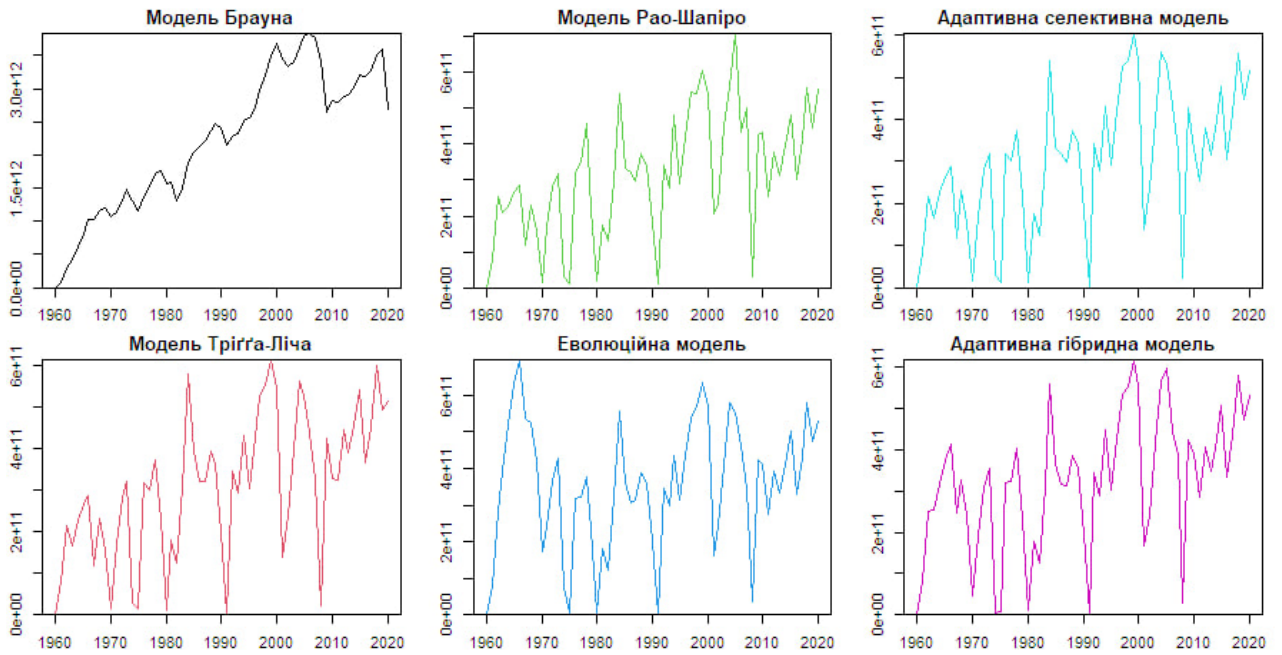


Рисунок 4.6 – Похибки прогнозів моделей з рис. 4.1-4.5

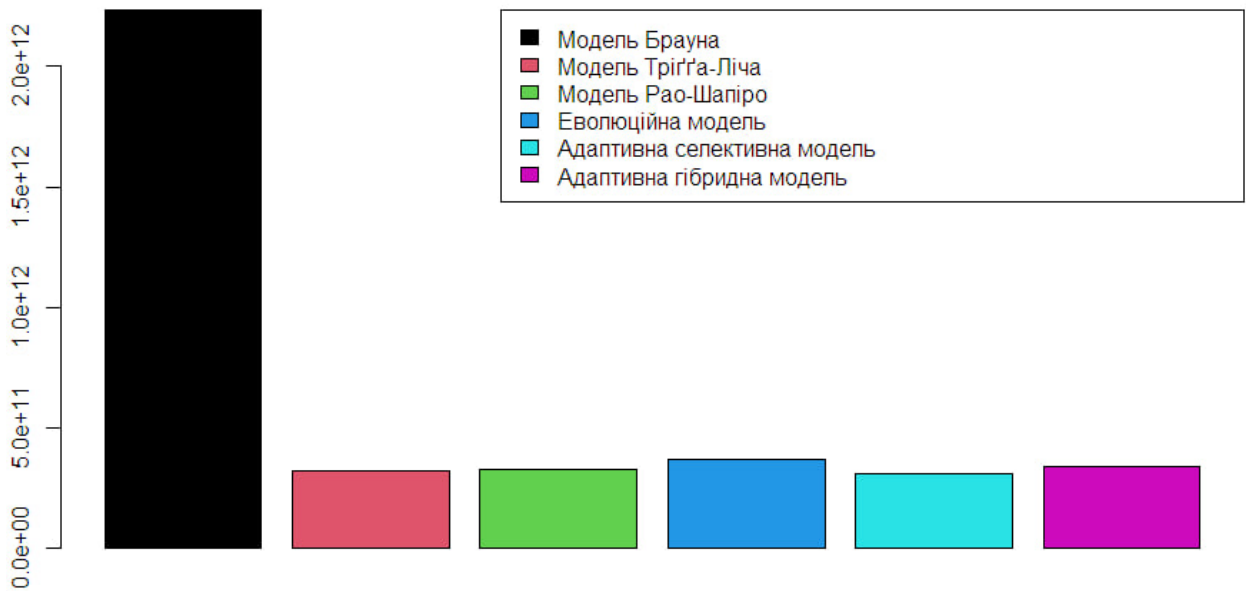


Рисунок 4.7 – Стовпчикова діаграма середніх похибок прогнозів моделей з рис. 4.1-4.5

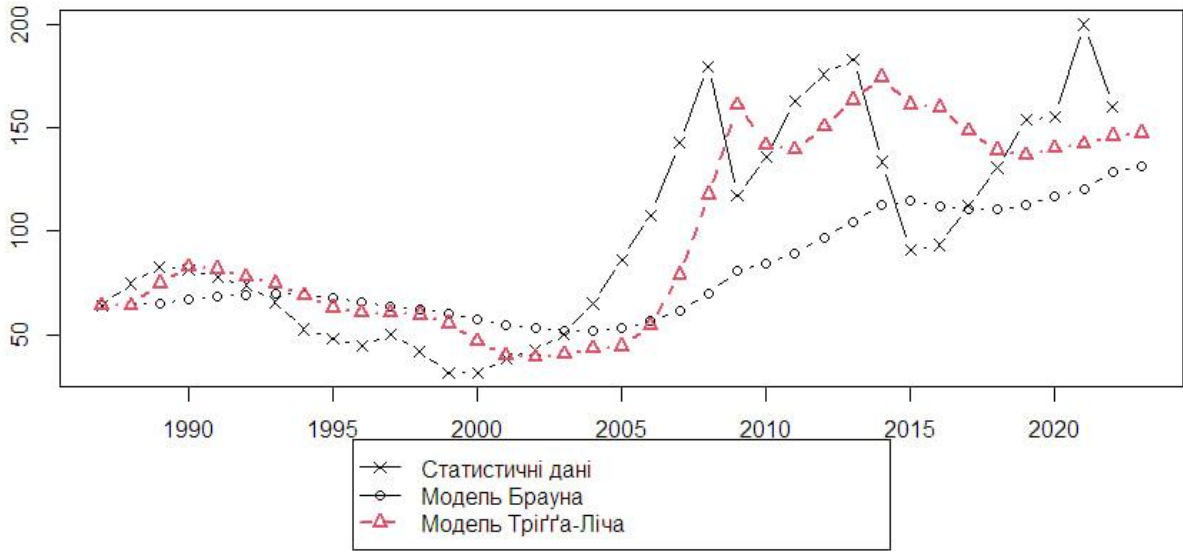


Рисунок 4.8 – Порівняння прогнозів моделі Брауна ($\alpha = 0.1$) та моделі Трігґа-Ліча ($\tau = 1$)

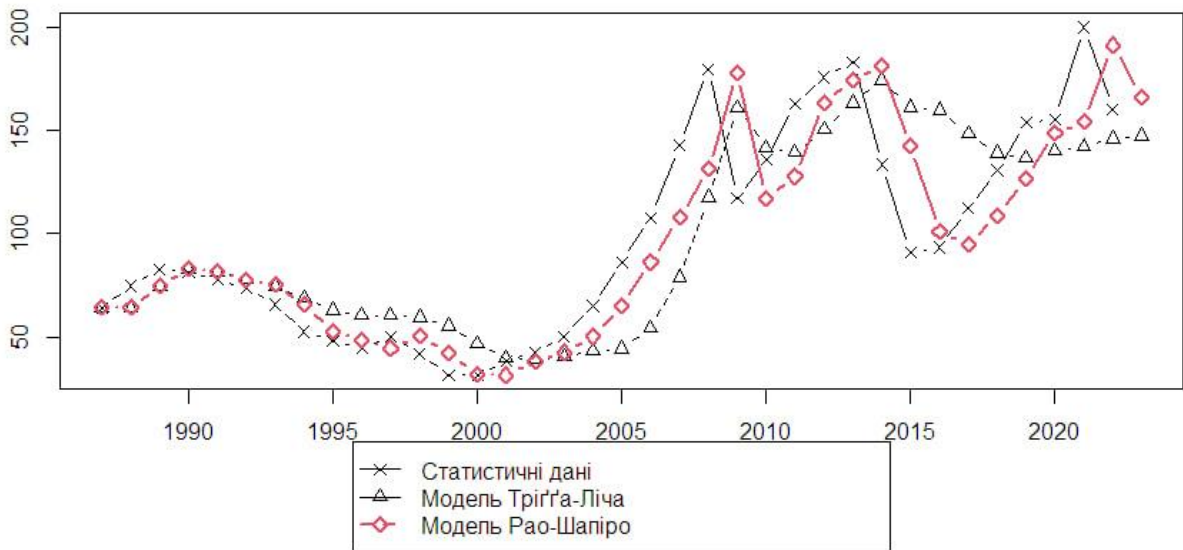


Рисунок 4.9 – Порівняння прогнозів моделі Трігґа-Ліча ($\tau = 1$) та моделі Рао-Шапіро ($N = 10, m = 5$)

Аналогічно до випадку з ВВП США, на рис. 4.8 видно, що модель Брауна і прогнозуючи такий показник для України значно повільніше підлаштовується під стрімкі зміни.

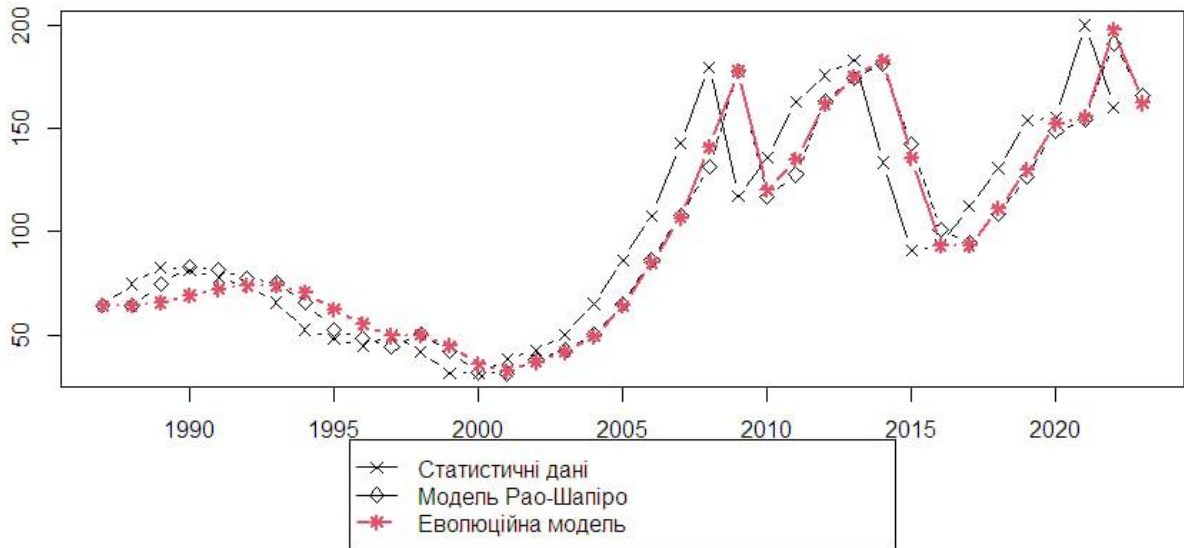


Рисунок 4.10 – Порівняння прогнозів моделі Рао-Шапіро ($N = 10, m = 5$) та Еволюційної моделі

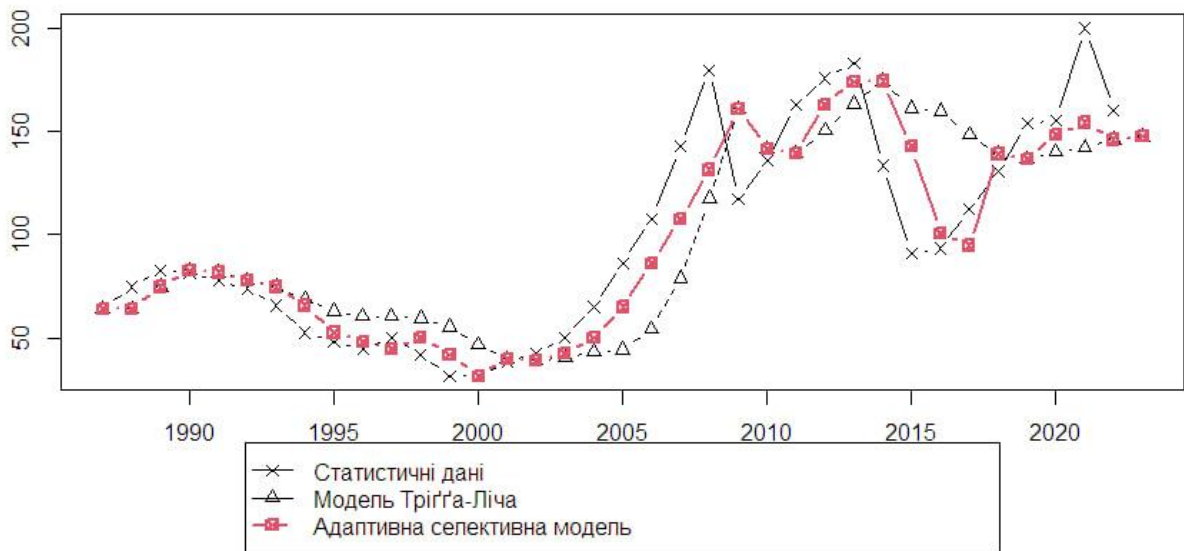


Рисунок 4.11 – Порівняння прогнозів моделі Трігга-Ліча ($\tau = 1$) та адаптивної селективної моделі (2 моделі в базовому наборі)

Статистичні дані ВВП України показують більшу мінливість рівня ряду. Тому тут на рис. 4.9 та 4.11 модель Трігга-Ліча помітно поступається прогнозуванню за моделлю Рао-Шапіро та адаптивною селективною моделлю. Еволюційна модель демонструє схожі результати до моделі Рао-Шапіро (рис 4.10).



Рисунок 4.12 – Порівняння прогнозів адаптивної селективної моделі (2 моделі в базовому наборі) та адаптивної гібридної моделі (3 предиктори в наборі)

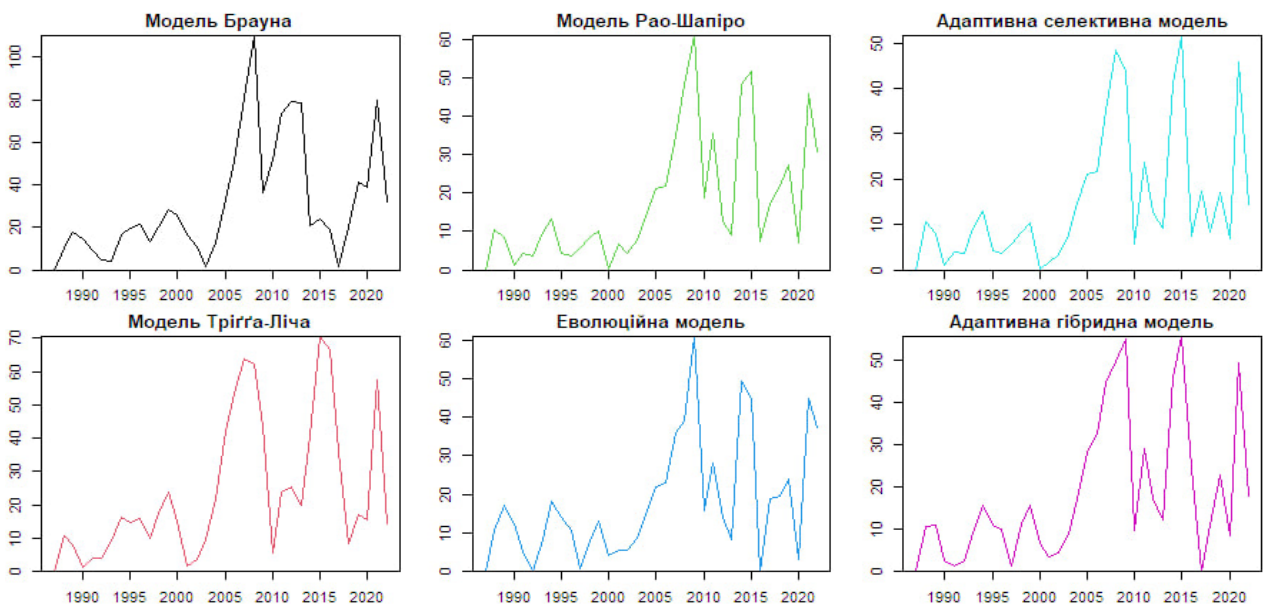


Рисунок 4.13 – Похибки прогнозів моделей з рис. 4.8-4.12

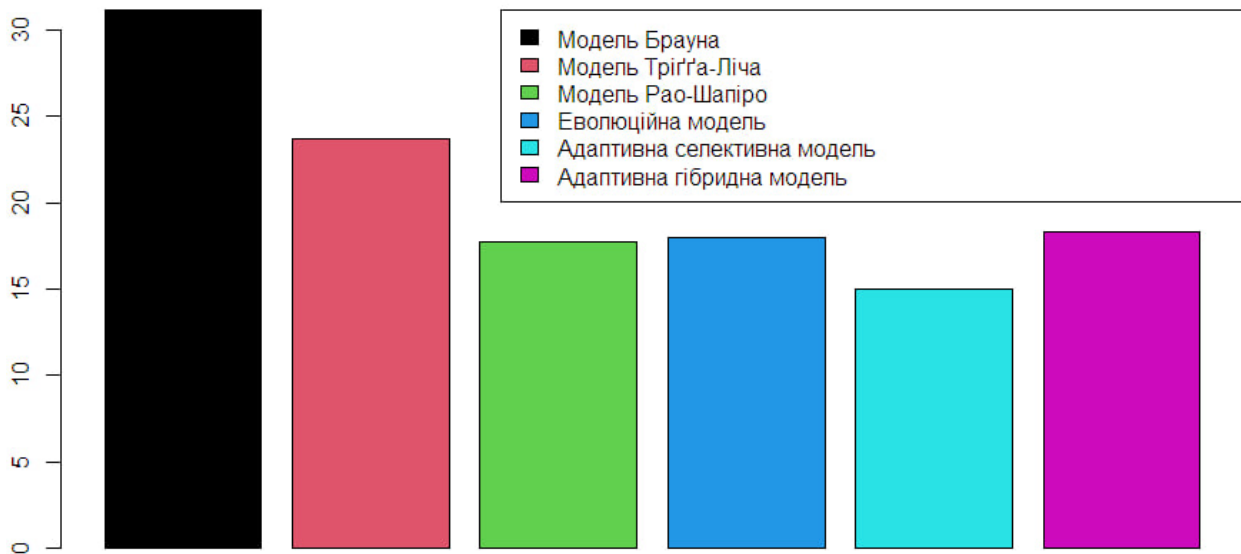


Рисунок 4.14 – Стівпчикова діаграма середніх похибок прогнозів моделей з рис 4.8-4.12

З рис. 4.14 видно, що адаптивна селективна модель має найменшу середню похибку, а отже, найкраще прогнозує дані в обох випадках. Для порівняння використовувалась найпростіша модель Брауна, з чого можна зробити висновок, що модифіковані моделі можуть давати досить непогані результати і без попереднього аналізу даних, в той час коли точність прогнозування за моделлю Брауна сильно залежить від того, наскільки вдало було підібрано математичну модель.

Результати можна покращити, взявши більш точні математичні моделі для моделі Тригга-Ліча та більшу кількість моделей у базовому наборі для адаптивної селективної та гібридної моделей.

ВИСНОВКИ

Адаптивні моделі – дуже корисний інструмент макроекономічного прогнозування. Оскільки такі моделі здатні враховувати нестационарність економічних процесів, адаптуватися до інерційності характеристик.

Ці моделі відображують поточні властивості ряду і можуть постійно враховувати еволюцію динамічних характеристик процесів. Мета таких методів полягає в тому, щоб побудувати економіко-математичні моделі, що можуть самостійно враховувати умови, що з часом змінюються. З цих причин такі моделі найчастіше використовуються для прогнозування на короткі періоди часу.

У роботі було розглянуто два типи модифікованих моделей: моделі з адаптивними параметрами адаптації та адаптивні комбіновані моделі.

Моделі з адаптивними параметрами адаптації (модель Трігга-Ліча, модель Рао-Шапіро, моделі Чоу та Монтгомері) базуються на змінюванні параметра адаптації залежно від похибки прогнозу. Такі моделі краще прогнозують процеси зі складною динамікою, бо, за рахунок адаптивного параметра, швидше пристосовуються до змін часового ряду.

Адаптивні комбіновані моделі передбачають обчислення декількох прогнозів на кожному кроці. Адаптивна селективна модель перемикається між різними моделями з базового набору, залежно від похибки прогнозу, що дозволяє краще описувати макроекономічні процеси, оскільки вони можуть мати різні тенденції на різних проміжках. За гібридною моделлю прогноз визначається як зважена сума прогнозів всіх моделей з базового набору.

Проаналізовані в роботі методи мають певні властивості:

- можуть застосовуватись для широкого спектру задач;

- адаптивне прогнозування базується на інтенсивному аналізі інформації, що міститься в окремих часових рядах, тобто не потребує великого об'єму інформації;
- модель, як правило, має чітку і просту математичну форму;
- адаптивна еволюція параметрів реагує на неоднорідність часових рядів.

Однак адаптивні моделі та методи доцільно використовувати лише для обробки рядів з помірними змінами в часі. Такі моделі не є дуже точними для довгострокових прогнозів.

Для більш точних результатів можна або провести попередній аналіз даних для визначення кращої математичної моделі процесу, або використовувати більше предикторів, таким чином збільшуючи шанси використання більш точної моделі без безпосереднього втручання дослідника в процесі прогнозування.

Ці модифіковані адаптивні моделі стали хорошою основою для подальшого розвитку авторегресійних моделей, моделей на основі машинного навчання, гібридних моделей тощо.

Перспективи розвитку модифікованих адаптивних моделей для прогнозування макроекономічних показників полягають у поліпшенні їхньої точності і здатності до адаптації до нових умов. Це може включати в себе використання більш складних архітектур нейронних мереж, врахування більш широкого спектру факторів, таких як соціальні та політичні чинники, і вдосконалення методів тренування і оцінки моделей.

Додатковою перспективою є комбінування моделей із класичними економетричними методами, що дозволяє поєднувати переваги обох підходів. Наприклад, можна використовувати адаптивні моделі для отримання загальних

тенденцій і шаблонів, а потім використовувати економетричні методи для оцінки впливу конкретних факторів та проведення прогнозів.

Розвиток модифікованих адаптивних моделей для прогнозування макроекономічних показників є активним напрямом досліджень і важливою галуззю для економічного прогнозування та прийняття рішень.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Кендалл М., Стюарт А. Багатовимірний статистичний аналіз та часові ряди, 1976.
2. Вітлінський В.В. Адаптивні моделі в економіці, КНЕУ, 2013
3. Brown R. G. Smoothing forecasting and prediction of discrete time series, New York, 1963. – p. 159-168.
4. Гіхман І., Скороход А., Ядренко М. Теорія ймовірностей та математична статистика, 1979.
5. Meyer R. F. An adaptive method for routine short-term forecasting, International Federation of Operational Research Societies, Oslo, 1963.
6. Trigg D. W. Monitoring a forecasting system, Operational Research Quarterly, 1964.
7. Trigg D. W., Leach A. G. Exponential smoothing with an adaptive response rate, Operational Research Quarterly, 1967.
8. Скороход А. В. Елементи теорії ймовірностей та теорії випадкових процесів, Наукова думка, 1975.
9. Лукашин Ю. П. Адаптивные методы краткосрочного прогнозирования временных рядов, 2003.
10. Shone M. L. Viewpoint, Operational Research Quarterly, 1967. – p. 318.
11. Rao A. G., Shapiro A. Adaptive smoothing using evolutionary spectra, Management Science, 1970.
12. Chow W. M. Adaptive control of the exponential smoothing constant, Journal of Industrial Engineering, 1965.
13. Roberts S. D, Reed R. Jr. The development of a self-adaptive forecasting technique, AIE Transactions, 1969.
14. Winters P. R. Forecasting sales by exponentially weighted moving averages, New York, 1960.

15. Montgomery D. C. Adaptive control of exponential smoothing parameters by evolutionary operation, AIIE Transactions, 1970.
16. The World Bank – United States Data: <https://data.worldbank.org/country/US>
17. The World Bank – Ukraine Data: <https://data.worldbank.org/country/ukraine>