

**Маслій В.В.,***к.е.н., доцент,**доцент кафедри аудиту, ревізії та аналізу,**Тернопільський національний економічний університет***Березька К.М.,***к.т.н., доцент,**доцент кафедри економіко-математичних методів,**Тернопільський національний економічний університет*

## ВИБІР ТА ОЦІНКА ARIMA-МОДЕЛІ ДЛЯ ПРОГНОЗУВАННЯ ОБСЯГІВ ПРЯМИХ ІНОЗЕМНИХ ІНВЕСТИЦІЙ

**Анотація.** У статті зроблено спробу оцінити можливість застосування ARIMA-моделей для прогнозування прямих іноземних інвестицій в Україну. Для вибору найбільш оптимальної моделі запропоновано модифікований алгоритм вибору ARIMA-моделі. Проаналізовано співвідношення ВІС-критерію та середньої відносної похибки MAPE із довжиною навчальної вибірки для моделей ARIMA різних порядків.

**Ключові слова:** прями іноземні інвестиції, прогнозування, ARIMA-модель, ВІС-критерій, АІС-критерій, середня відносна помилка MAPE, довжина навчальної вибірки.

**Постановка проблеми.** Значний інтерес до проблеми залучення прямих іноземних інвестицій (далі – ПІІ) останніми роками пояснюється не тільки потребою в додаткових обсягах капіталу для розв’язання проблем трансформації вітчизняної економіки, але й тим, що вони є одними з ключів до інтеграції України в ринкове та демократичне європейське співтовариство. Зростаюча роль прямих іноземних інвестицій в усіх сферах суспільно-економічного життя спричинює необхідність детального дослідження закономірностей процесу іноземного інвестування, зокрема закономірностей розвитку. Слід зазначити, що тенденція глобалізації, легкість переміщення капіталів через кордони держав, інформатизація економіки, зростаюча вартість знань тощо справляють кількісний та якісний вплив на формування взаємостосунків між господарюючими суб’єктами на ринку. Тенденції розвитку процесу іноземного інвестування можна вивчати як за допомогою методів аналізу динамічних рядів, так і на основі прогнозних моделей. В науковій літературі, присвяченій прогнозуванню, запропоновано значну кількість прогнозних моделей, проте останнім часом популярними стали ARIMA-моделі (Autoregressive Integrated Moving Average), які пояснюють поведінку часового ряду, виходячи лише з його значень в попередні моменти часу, а також добре описують як стаціонарні, так і нестаціонарні часові ряди.

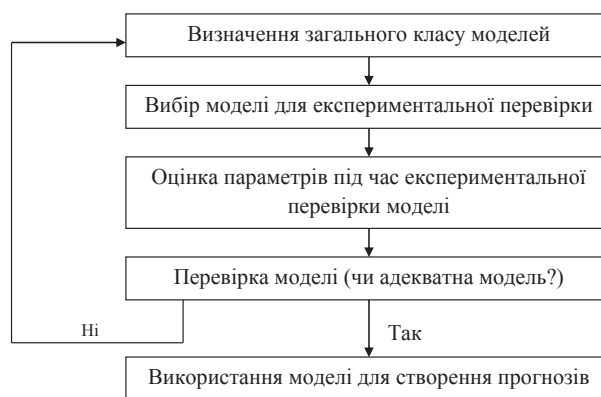
**Аналіз останніх досліджень і публікацій.** Такого роду прогнозні моделі знаходять своє застосування як в техніці, так і в економіці. Зокрема, можна виділити дослідження Є. Реуцького [1], О. Любич [2], Н. Рилової [3], В. Волохатої [4], О. Дзедзелюка [5], В. Мартинюка [6]. Як зазначає І. Лук’яненко, широкому застосуванню ARIMA-моделей сприяє те, що [7]:

1) не завжди попередня інформація про можливі взаємозв’язки між динамічними рядами може бути добре обґрунтована; у цьому випадку чисто статистична модель, що зв’язує

поточні та попередні значення досліджуваного показника, може використовуватися для короткострокових прогнозів;

2) інколи з добре відомих структурних моделей економічної теорії можна отримати моделі типу авторегресійних або моделей ковзного середнього, особливо у процесі оцінювання наведеної форми симультативних систем рівнянь, тобто під час вираження ендогенних змінних структурної моделі через попередньо визначені та екзогенні змінні.

**Виділення не вирішених раніше частин загальної проблеми.** Для вибору найбільш оптимальної ARIMA-моделі переважно застосовують алгоритм за методом Бокса-Дженкінса (рис. 1). Це передбачає побудову ARIMA-моделі на основі наявного динамічного ряду. В деяких випадках навіть акцентують увагу на тому, що чим більше довжина динамічного ряду, тим краще буде модель.



**Рис. 1.** Блок-схема стратегії вибору моделі за методом Бокса-Дженкінса [8, с. 454]

В наших попередніх дослідженнях [9; 10] ми зазначали, що ARIMA-моделі обсягів прямих іноземних інвестицій (ПІІ) високих порядків не дають можливості отримати якісний прогноз.

**Мета статті.** В дослідженні були поставлені завдання визначити:

1) чи доцільно для отримання надійних прогнозів будувати ARIMA-модель на основі усього ряду динаміки обсягів прямих іноземних інвестицій;

2) як змінюються критерії оптимальності вибору ARIMA-моделі зі зменшенням навчальної вибірки;

3) чи впливає довжина навчальної вибірки на якість прогнозних розрахунків.

Як інформаційна база застосовувалися дані про шоквартальні обсяги прямих іноземних інвестицій за період з I квар-

талу 1998 р. по IV квартал 2015 р. (72), які були опубліковані Національним банком України [11]. Їх графічне зображення подане на рис. 2.

**Виклад основного матеріалу дослідження.** Відповідно до завдання ми модифікували вищезазначений алгоритм вибору ARIMA-моделі таким чином (рис. 3):

На основі вихідних даних тестуються ARIMA-моделі нульового, першого та другого порядків. Для вибору найбільш оптимальної моделі застосовуються інформаційний критерій Акаїке (AIC-критерій) та критерій Шварца (BIC-критерій) [12]. Після цього довжина навчальної вибірки скорочується на 1 значення, а процес повторюється. Процедура триває доти, доки довжина навчальної вибірки не буде меншою 33 значень. Зазначимо, що

початкова довжина навчальної вибірки становить 72 значення, а прогнозна вибірка, на основі якої розраховуються показники якості прогнозу, – 4 значення.

На першому етапі на основі навчальної вибірки із 72 спостережень були побудовані вищезазначені моделі та обчислені AIC- та BIC-критерії (табл. 1).

Дані, наведені в табл. 1 свідчать про те, що найкращими (найменшими) є значення AIC- та BIC-критеріїв для моделі ARIMA (0, 1, 1). Проте слід зазначити, що всі розраховані вище критерії знаходяться у вузьких межах: AIC-критерій – від 16,628 до 16,830; BIC-критерій – від 16,692 до 16,895. Це свідчить про те, що будь-яка із запропонованих моделей ARIMA може бути використана як прогнозна.

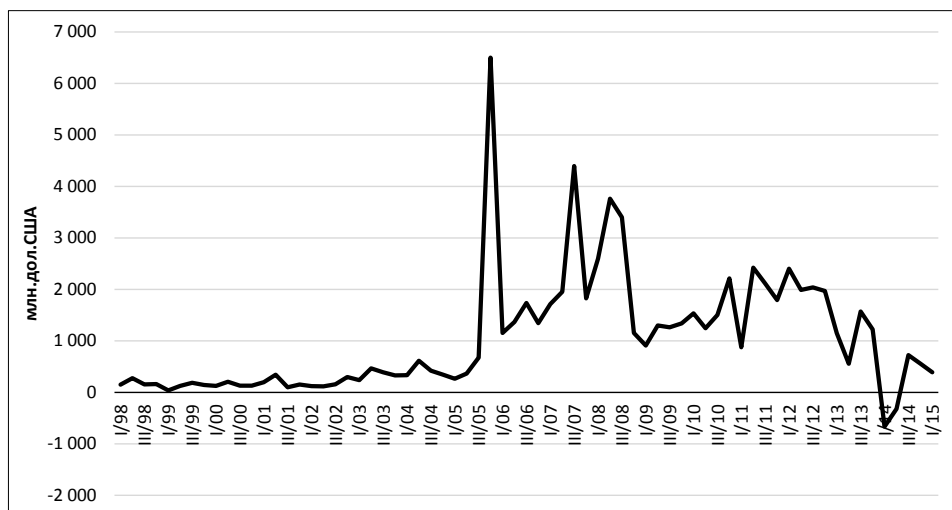


Рис. 2. Вихідний ряд динаміки обсягів прямих інвестицій в Україну

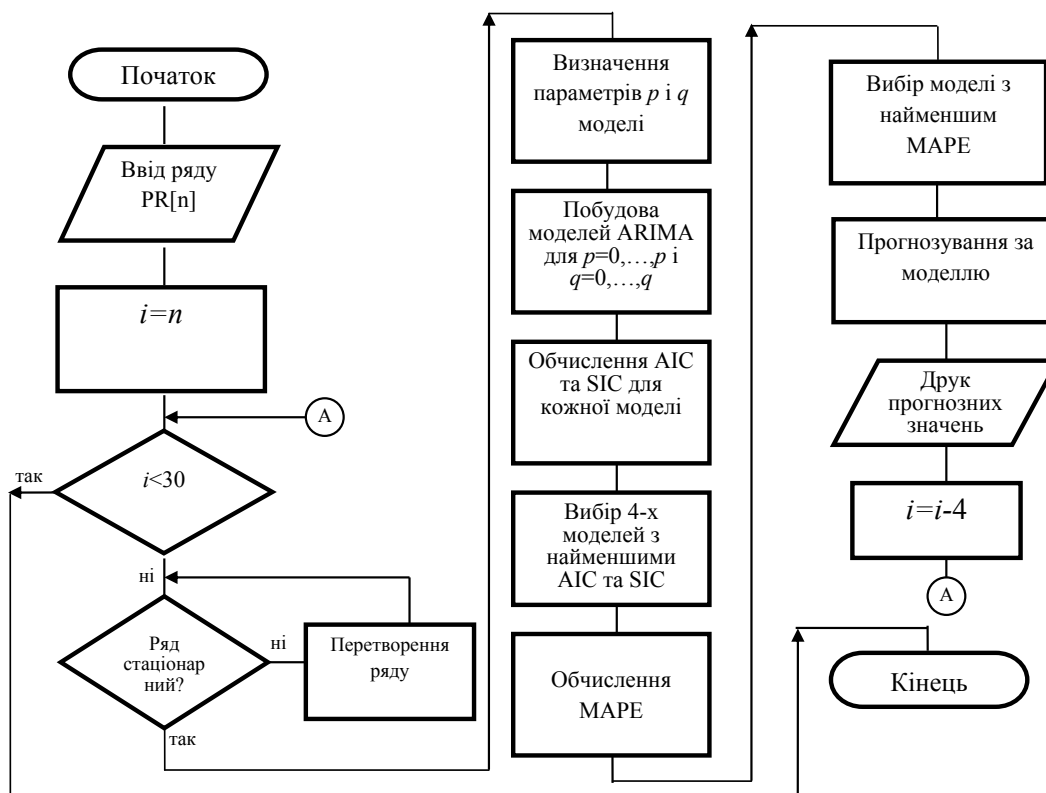


Рис. 3. Модифікований алгоритм вибору ARIMA-моделі

Таблиця 1  
Критерії надійності для моделей ARIMA<sub>72</sub>\*

| Модель          | AIC-критерій | BIC-критерій |
|-----------------|--------------|--------------|
| ARIMA (2, 1, 1) | 16,710       | 16,840       |
| ARIMA (2, 1, 2) | 16,692       | 16,853       |
| ARIMA (2, 1, 0) | 16,739       | 16,837       |
| ARIMA (0, 1, 2) | 16,654       | 16,749       |
| ARIMA (1, 1, 1) | 16,669       | 16,766       |
| ARIMA (0, 1, 1) | 16,628       | 16,692       |
| ARIMA (1, 1, 0) | 16,830       | 16,895       |
| ARIMA (1, 1, 2) | 16,649       | 16,777       |

\* Довжина навчальної вибірки становить 72 спостереження (з 1 кв. 1998 по IV кв. 2015 рр.)

Для оцінювання було обрано перші чотири моделі ARIMA з найменшими значеннями AIC- та BIC-критеріїв: ARIMA (0, 1, 1), ARIMA (0, 1, 2), ARIMA (1, 1, 1) та ARIMA (1, 1, 2). Порівняння прогнозової вибірки з отриманими даними дало можливість отримати такі оцінки якості прогнозних моделей (табл. 2).

Таблиця 2  
Помилки моделей ARIMA<sub>72</sub>

| Модель          | Прогнозна вибірка | Прогнозні дані              | Помилки моделі      | Значення                         |
|-----------------|-------------------|-----------------------------|---------------------|----------------------------------|
| ARIMA (0, 1, 1) | -319              | 608,7972                    | RMSE<br>MAE<br>MAPE | 960,1732<br>456,3515<br>53,40241 |
|                 | 762               | 558,3072                    |                     |                                  |
|                 | 556               | 664,1919                    |                     |                                  |
|                 | 346               | 735,7922<br><b>767,3286</b> |                     |                                  |
| ARIMA (0, 1, 2) | -319              | 664,4240                    | RMSE<br>MAE<br>MAPE | 958,8126<br>458,9593<br>53,17097 |
|                 | 762               | 597,1229                    |                     |                                  |
|                 | 556               | 721,0689                    |                     |                                  |
|                 | 346               | 765,7988<br><b>781,4287</b> |                     |                                  |
| ARIMA (1, 1, 1) | -319              | 665,1005                    | RMSE<br>MAE<br>MAPE | 965,7309<br>463,9523<br>53,40263 |
|                 | 762               | 597,2009                    |                     |                                  |
|                 | 556               | 716,8768                    |                     |                                  |
|                 | 346               | 766,6074<br><b>782,7870</b> |                     |                                  |
| ARIMA (1, 1, 2) | -319              | 662,2981                    | RMSE<br>MAE<br>MAPE | 942,3361<br>525,6630<br>99,65950 |
|                 | 762               | 535,5693                    |                     |                                  |
|                 | 556               | 750,9948                    |                     |                                  |
|                 | 346               | 686,3367<br><b>833,2727</b> |                     |                                  |

Незважаючи на те, що для розрахунків були обрані кращі за критеріями AIC- та BIC-моделі, помилки прогнозу, зокрема MAPE > 50, свідчать про незадовільний прогноз.

Розрахунки моделей проводились для інших навчальних вибірок, довжина яких кожної наступної ітерації вкорочувалась на 1 період (це дало змогу протестувати 50 моделей з довжиною навчальної вибірки від 72 до 33). Таке зменшення не вплинуло на клас моделей часових рядів: всі часові ряди, які застосовувались в дослідженні, є стаціонарними. Із двох запропонованих вище інформаційних критеріїв ми застосували BIC-критерій, адже цей критерій є аналогом AIC з більш строгою функцією штрафів, особливістю є його застосування у разі невеликого розміру спостережуваної вибірки [13]. Отримані значення BIC-критерію для моделей з різною довжиною навчальної вибірки наведено в табл. 3 (з метою уникнення переобтяження результатів дослідження цифровим матеріалом ми наводимо вибіркові результати розрахунків).

Графічне подання вищенаведених даних (рис. 4) дає можливість зробити висновки про однакову поведінку розрахованого для різних моделей BIC-критерію: спочатку скорочення навчальної вибірки приводить до незначного зростання, а поступове зменшення критерію спостерігається для навчальних вибірок обсягом 47 і менше. Найкращий результат отримано для моделі ARIMA (1, 1, 2) з обсягом навчальної вибірки 34, що дає можливість вибрати її як найбільш оптимальну для прогнозування обсягів іноземних інвестицій в Україну.

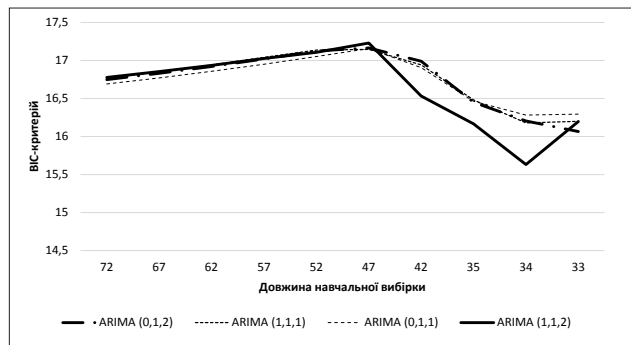


Рис. 4. BIC-критерій для моделей ARIMA з різною довжиною навчальної вибірки

Для того щоб визначити, чи впливає довжина навчальної вибірки на величину помилки прогнозу, були обчислені помилки MAPE для моделей ARIMA нульового, першого та другого порядків. Вибіркові дані, наведені в табл. 4, свідчать про те, що для моделей ARIMA, які побудовані на коротких динамічних рядах (34 та 33 елементи), найкращими є результати для моделі

BIC-критерії для моделей ARIMA

Таблиця 3

| Модель          | BIC-критерії для моделей ARIMA з довжиною навчальної вибірки |        |        |        |        |        |        |        |        |        |
|-----------------|--|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|                 | 72   | 67     | 62     | 57     | 52     | 47     | 42     | 35     | 34     | 33     |
| ARIMA (0, 1, 2) | 16,749   | 16,832 | 16,922 | 17,020 | 17,126 | 17,165 | 16,990 | 16,450 | 16,205 | 16,066 |
| ARIMA (1, 1, 1) | 16,766   | 16,850 | 16,942 | 17,041 | 17,137 | 17,146 | 16,943 | 16,482 | 16,179 | 16,200 |
| ARIMA (0, 1, 1) | 16,692   | 16,771 | 16,858 | 16,951 | 17,053 | 17,157 | 16,907 | 16,464 | 16,282 | 16,295 |
| ARIMA (1, 1, 2) | 16,777   | 16,855 | 16,936 | 17,026 | 17,108 | 17,230 | 16,533 | 16,167 | 15,632 | 16,197 |

ARIMA (1, 1, 2), адже зменшення довжини навчальної вибірки привело до зменшення помилки MAPE практично на 72,3% (для ряду довжиною 71 спостереження MAPE = 103,049, для 34 спостережень MAPE = 28,516), а також для моделі ARIMA (2, 1, 2), оскільки помилка зменшилась більш ніж на 73% (для ряду довжиною 71 спостереження MAPE = 100,347, для 33 спостережень MAPE = 26,923) (табл. 4).

Для того щоб виявити, як впливає довжина навчальної вибірки на величину помилки MAPE, для ARIMA-моделей різного порядку застосуємо лінійний коефіцієнт кореляції ( $r$ ). Висновки щодо взаємозв'язку між довжиною навчальної вибірки та величиною MAPE на основі результатів розрахунків, наведених в табл. 5, такі: для моделі ARIMA (0, 1, 1) існує помірний обернений зв'язок між досліджуваними статистичними параметрами, для моделей ARIMA (0, 1, 2), (1, 1, 1) та (2, 1, 1) – слабкий обернений зв'язок, для моделей ARIMA (1, 1, 2) та (2, 1, 2) – прямий зв'язок.

Проте слід зазначити, що навіть найменші значення помилки MAPE (26,923 та 28,516) не є підставою для того, щоб обрати моделі ARIMA (2, 1, 2) та ARIMA (1, 1, 2) для прогнозування, адже це тільки свідчить про те, що прогноз задовільний.

**Висновки.** Побудова прогнозних моделей для сучасного українського ринку ускладнюється ще й специфікою економіки, яку переживає Україна. Це є об'єктивною причиною, яка не дає змогу досягти задовільної точності будь-яких прогнозів.

Результати проведеного дослідження дають змогу зробити такі висновки: застосування ARIMA-моделі дає задовільний результат на вкороченій навчальній вибірці, отже, для отримання якісних прогнозів є доцільним застосування не всього динамічного ряду обсягів ПІІ, а тільки його другої половини, яка приблизно починається з 2007 р. (рис. 5).

В таких умовах для прогнозування обсягів ПІІ із використанням всіх наявних даних, на нашу думку, доцільно застосовувати сплайн-функції, або ж більш спрощені методи експоненційного прогнозування.

**Література:**

1. Реуцький Є. Метод прогнозування характеристик точності вимірювальних каналів енергетичних систем / Є. Реуцький // Енергетика і автоматика. – 2016. – № 3. – С. 171–181.
2. Любич О. Застосування моделі ARIMA для прогнозування доходів зведеного бюджету України / О. Любич, В. Домрачев // Екон.-

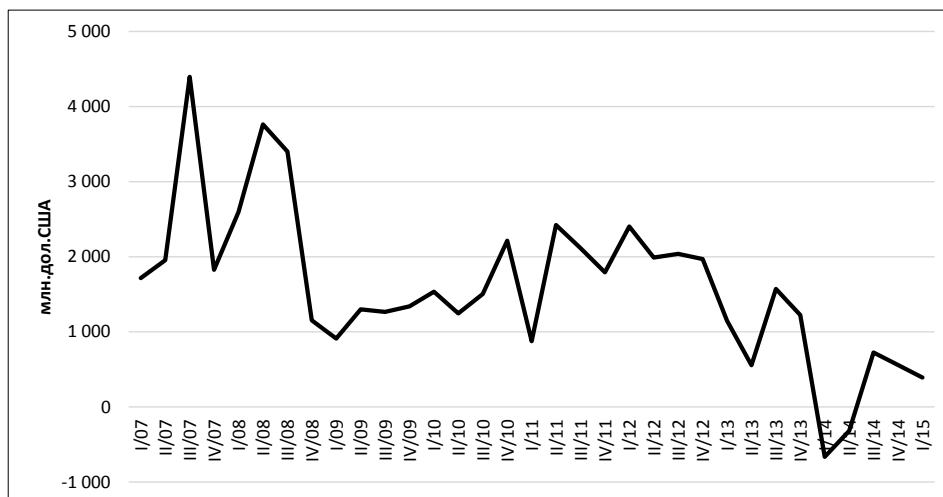


Рис. 5. Динаміка прямих інвестицій в Україну за 2007–2015 рр.

Таблиця 4

Помилки MAPE для моделей ARIMA

| Модель          | Помилки MAPE для моделей ARIMA з довжиною навчальної вибірки |         |        |        |        |        |        |        |        |        |
|-----------------|--|---------|--------|--------|--------|--------|--------|--------|--------|--------|
|                 | 72   | 71      | 65     | 60     | 55     | 50     | 45     | 40     | 34     | 33     |
| ARIMA (2, 1, 1) | 51,632   | 51,487  | 50,861 | 47,950 | 50,007 | 77,076 | 64,703 | 51,886 | 49,637 | 50,982 |
| ARIMA (2, 1, 2) | 94,738   | 100,347 | 92,321 | 73,222 | 56,822 | 61,427 | 52,939 | 32,075 | 47,032 | 26,923 |
| ARIMA (2, 1, 0) | 55,545   | 56,063  | 52,858 | 51,577 | 51,966 | 55,512 | 58,173 | 51,797 | 49,642 | 50,900 |
| ARIMA (0, 1, 2) | 53,171   | 53,475  | 50,338 | 53,024 | 49,675 | 54,690 | 78,263 | 47,494 | 48,522 | 63,208 |
| ARIMA (1, 1, 1) | 53,403   | 51,926  | 50,837 | 53,637 | 73,143 | 77,241 | 69,208 | 48,965 | 49,651 | 50,377 |
| ARIMA (0, 1, 1) | 53,402   | 53,677  | 50,435 | 53,096 | 49,615 | 53,438 | 80,689 | 50,294 | 62,422 | 55,182 |
| ARIMA (1, 1, 0) | 58,222   | 58,443  | 56,224 | 57,487 | 56,000 | 60,082 | 61,896 | 50,880 | 48,141 | 46,621 |
| ARIMA (1, 1, 2) | 99,660   | 103,049 | 89,547 | 93,403 | 78,247 | 62,472 | 63,971 | 48,833 | 28,516 | 46,674 |

Таблиця 5

Значення лінійного коефіцієнта кореляції при  $\alpha = 0,05$  та  $n = 40$

|       |         |         |         |         |         |         |         |
|-------|---------|---------|---------|---------|---------|---------|---------|
| ARIMA | 0, 1, 1 | 0, 1, 2 | 1, 1, 1 | 1, 1, 2 | 2, 1, 2 | 2, 1, 1 | 2, 1, 0 |
| $r$   | -0,444  | -0,266  | -0,017  | +0,965  | +0,942  | -0,226  | -0,127  |

- мат. моделювання соц.-екоп. систем : зб. наук. пр. – Вип. 16. – К. : МННЦТС НАН та МОН України, 2011. – С. 33–37.
3. Рилова Н. Синтез ARIMA-моделей для прогнозування коефіцієнтів виходу кондиційних напівпровідникових матеріалів / Н. Рилова, І. Оксаніч // Системи обробки інформації. – 2015. – Вип. 5 (130). – С. 102–107.
  4. Волохата В. Управління залученими ресурсами банку на основі цінового прогнозування з використанням мультиплікативних моделей ARIMA / В. Волохата // Інвестиції: практика та досвід. – 2014. – № 24. – С. 72–78.
  5. Побудова ARIMA-моделей часових рядів для прогнозування метеоданих на мові програмування R / [О. Дзєндзелюк, Л. Костів, В. Рабик] // Електроніка та інформаційні технології. – 2013. – Вип. 3. – С. 211–219.
  6. Мартинюк В. Прогнозування надходження податкових платежів до державного бюджету за допомогою використання ARIMA-моделі / В. Мартинюк // Наукові праці НДФІ. – 2011. – № 2 (55). – С. 46–55.
  7. Лук'яненко І. Прогнозування податкових надходжень за допомогою моделей корегування помилок / І. Лук'яненко, Ю. Городніченко // Фінанси України. – 2001. – № 7. – С. 89–99.
  8. Бизнес-прогнозирование / [Д.Э. Ханк, Д.У. Уичерн, А.Дж. Райтс] ; пер. с англ. – 7-е изд. – М. : Издательский дом «Вильямс», 2003. – 656 с.
  9. Березька К. Деякі аспекти прогнозування обсягів портфельних інвестицій в Україну / К. Березька, В. Маслій // Актуальні проблеми економіки 2014 : матеріали VIII міжнародної конференції (м. Київ, 12 грудня 2014 р.) – К., 2015. – С. 7–12.
  10. Березька К. Модифіковані ARIMA-моделі обсягів прямих іноземних інвестицій / К. Березька, В. Маслій // Актуальні проблеми економіки 2015–2016 : матеріали IX міжнародної конференції (м. Київ, 19 лютого 2016 р.) – К., 2016. – С. 15–20.
  11. Динаміка платіжного балансу відповідно до КПБ5 [Електронний ресурс]. – Режим доступу : [http://www.bank.gov.ua/control/uk/publish/article?art\\_id](http://www.bank.gov.ua/control/uk/publish/article?art_id).
  12. Аналіз використання інформаційних критеріїв у моделях оцінки надійності програмного забезпечення / [В.С. Яковина, Д.В. Федасюк, О.О. Нитребич] // Вісник НТУ «ХПІ». – 2014. – № 26 (1069). – С. 108–115.
  13. Schwarz G. Estimating the Dimension of a Model / G. Schwarz // Annals of Statistics. – 1978. – Vol. 6, № 2. – P. 461–464.
  14. Закс Л. Статистическое оценивание / Л. Закс ; пер. с нем. В. Варыгина ; под ред. Ю. Адлера, В. Горского. – М. : Статистика, 1976. – 600 с.

**Маслій В.В., Березька К.Н. Выбор и оценка ARIMA-модели для прогнозирования объемов прямых иностранных инвестиций**

**Аннотация.** В статье предпринята попытка оценить возможности применения ARIMA-моделей для прогнозирования прямых инвестиций в Украину. Для выбора наиболее оптимальной модели предложен модифицированный алгоритм выбора ARIMA-модели. Проанализировано соотношение BIC-критерия и средней относительной ошибки MAPE с длиной учебной выборки для моделей ARIMA разных порядков.

**Ключевые слова:** прямые иностранные инвестиции, прогнозирование, ARIMA-модель, BIC-критерий, AIC-критерий, средняя относительная ошибка MAPE, длина учебной выборки.

**Maslii V.V., Berezka K.M. Selection and evaluation of ARIMA-models for forecasting the amount of foreign direct investments**

**Summary.** The article is attempting to assess the use of ARIMA models for the forecasting of direct investment in Ukraine. To select the most optimal models, a modified algorithm of choosing the ARIMA model is proposed. Analysed the ratio of BIC-criterion and the mean absolute percentage error MAPE with the length of the training sample models of ARIMA different orders.

**Keywords:** foreign direct investments, forecasting, BIC-criterion, AIC-criterion, mean absolute percentage error MAPE, length of the training sample.