

**ІНДУКТИВНЕ МОДЕЛЮВАННЯ РИЗИКІВ ЗБИТКІВ ВІД
РУЙНІВНИХ ПОВЕНЕЙ В БАСЕЙНІ Р. ТИСА ЗА ЕМПІРИЧНИМИ ДАНИМИ
З ВИКОРИСТАННЯМ МОДЕЛЕЙ РЕГРЕСІЙНОГО ТИПУ**

***Анотація.** За даними рядів динаміки значень збитків від руйнівних повеней в басейні р. Тиса та максимальних витрат води р. Тиса, що визначалися на гідрометричному посту Вилок, розглянуто задачу індуктивного моделювання ризиків збитків як добутоків збитків та ймовірностей перевищення максимальних витрат води з використанням моделей регресійного типу.*

***Ключові слова:** ймовірність перевищення, індуктивне моделювання, максимальні витрати води, ризики збитків, ряди динаміки, ситуаційні регресійні моделі.*

***Аннотация.** На основе динамических рядов убытков от разрушительных паводков в бассейне р. Тиса и максимальных расходов воды р. Тиса, которые определялись на гидрометрическом посту Вилок, рассмотрено задачу индуктивного моделирования рисков убытков как произведений убытков и вероятностей превышения максимальных расходов воды с использованием моделей регрессионного типа.*

***Ключевые слова:** вероятность превышения, индуктивное моделирование, максимальные потери воды, риски убытков, ряды динамики, ситуационные регрессионные модели.*

***Abstract.** According to the time series of values of losses of devastating floods in the basin of Tisza and maximum water discharges of the river Tisza, which were identified on the hydrometric post of Vylok, the problem of inductive modeling the risks of losses as products of damage and probability of exceeding the maximum water discharge using the regression type models was examined.*

***Keywords:** probability of exceeding, inductive modeling, maximum water loss exposures, a series of dynamics, situational regression model.*

Вступ

Традиційні підходи до побудови математичних моделей за емпіричними даними ґрунтуються на принципі оптимізації. Однак навіть в найбільш простих задачах моделювання за емпіричними даними, з використанням моделей регресійного типу, принцип оптимізації, який передбачає вихід на ряд граничних обмежень, не завжди виконується [1, 2]. Задача оптимізації може ускладнюватися зі збільшенням кількості емпіричних даних, розмірності моделі за рахунок врахування додаткових факторів і параметрів, нелінійних ефектів тощо. Все це може призводити до посилення невизначеності прогнозування [3, 4].

Введення в модель додаткових незалежних змінних в умовах немонотонних рядів динаміки зазвичай не здатне зняти проблему гетероскедастичності моделі (неоднорідності дисперсії). При цьому додатковою проблемою може стати кореляція між пояснюючими (екзогенними) змінними багатфакторної моделі (мультиколінеарність). В умовах мультиколінеарності коефіцієнти регресії стають у край нестійкими до малих змін у даних, що порушує стійкість розв'язків при відшукуванні невідомих параметрів моделі.

Побудова таких моделей регресійного типу як авторегресійні, дистрибутивно-лагові тощо, в яких враховується авторегресія та фактор часу також не завжди приносить успіх.

1. Новий підхід до моделювання за емпіричними даними з використанням регресійних моделей

Пропонується підхід до моделювання різних економічних, екологічних та соціальних процесів за емпіричними даними, що ґрунтується на побудові ситуаційних і індуктивних моделей регресійного типу. Ситуаційні моделі, зазвичай у вигляді простих, однофакторних регресій, адаптуються до вибіркового ряду динаміки залежної і незалежної змінних моделі, які характеризуються монотонністю або квазістаціонарністю (ми це називатимемо однорідним прогнозним фоном) на відповідних часових інтервалах. Побудовані таким чином ситуаційні моделі вважаються адекватними на цих часових інтервалах. Далі результати ситуаційного моделювання (модельні дані) формують підставу (базу) для наступної побудови індуктивних моделей, де враховуються особливості еволюції ендогенної змінної в часі. Загальна математична модель при такій постановці задачі являє собою сімейство індуктивних моделей як моделей «рівнів», представлених у вигляді регресій модельних значень, трендів або комбінацій трендів та регресій для залишків трендів модельних значень ендогенної змінної, що встановлюються за ситуаційними моделями минулих періодів, в тому числі, при необхідності, – від деяких похідних параметрів незалежних змінних. При цьому прогнозування може здійснюватися як на основі ситуаційних моделей, за допомогою яких відслідковується поведінка ендогенної змінної в межах інтервалів часу з однорідним прогнозним фоном, за даними яких ситуаційні моделі можуть вважатися адекватними, так і на основі індуктивних моделей, за допомогою яких відслідковується еволюція ситуаційних моделей як фазових портретів минулих станів динамічної системи і можуть встановлюватися ситуаційні моделі для визначення залежної змінної для майбутніх періодів [5, 6].

При такій постановці задачі в якості рівнянь зв'язку при побудові ситуаційних і індуктивних моделей можуть використовуватися відносно прості залежності, що легко адаптуються до нових даних та змін прогнозного фону в рядах динаміки. При необхідності можуть враховуватися лаги між змінними моделей. Допускається модифікація змінних моделі з метою побудови адекватних ситуаційних і індуктивних моделей для цілей прогнозування.

Під ситуаційною моделлю далі будемо розуміти модель, адаптовану до певної ситуації (прогнозного фону) і яка є адекватною в цій ситуації, що розгортається протягом обмеженого періоду часу.

Прогнозний фон характеризується як сукупність зовнішніх і/або внутрішніх умов, істотних для вибору структури моделі з метою прогнозування.

Ситуаційні моделі, по суті, відображають окремі фазові стани динамічної системи, що досліджується, на різних інтервалах часу. При цьому перехід від однієї регресійної моделі до іншої, які визначають сусідні фазові стани системи, може відбуватися немонотонно (стрибкоподібно) [5].

Ситуаційні моделі будуються на основі рядів актуальних (фактичних) емпіричних даних – фізичних даних, отриманих в результаті безпосередніх спостережень, вимірювань тощо. Ситуаційні моделі можуть використовуватися для оперативного прогнозування.

Індуктивна модель – це модель, отримана з узагальнення (ансамблю) кількох моделей (ситуаційних або індуктивних). Будується на основі рядів модельних даних, що являють собою результати статистичної обробки даних і/або ситуаційного моделювання.

Індуктивна модель, отримана з узагальнення кількох (сімейства) ситуаційних моделей, може охоплювати кілька кластерів актуальних даних і, таким чином, відображає еволюцію модельних значень результуючої змінної стану динамічної системи в часі. Індуктивні можуть використовуватися для довгострокового прогнозування.

Метою роботи є демонстрація принципової можливості адекватного довгострокового прогнозування за допомогою побудованих відповідним чином індуктивних моделей за рядами емпіричних даних.

2. Моделювання ризиків збитків від руйнівних повеней в басейні р. Тиса

Ряд динаміки зафіксованих збитків від руйнівних повеней в басейні р. Тиса в межах Закарпатської області України наведено на рис. 1. Ряд збитків в цілому є нестационарним, немонотонним і характеризується загальною тенденцією до наростання збитків в часі на фоні стаціонарного, без вираженого тренду, ряду максимальних витрат води р. Тиса, що визначалися на гідрометричному посту Вилोक (рис. 2).

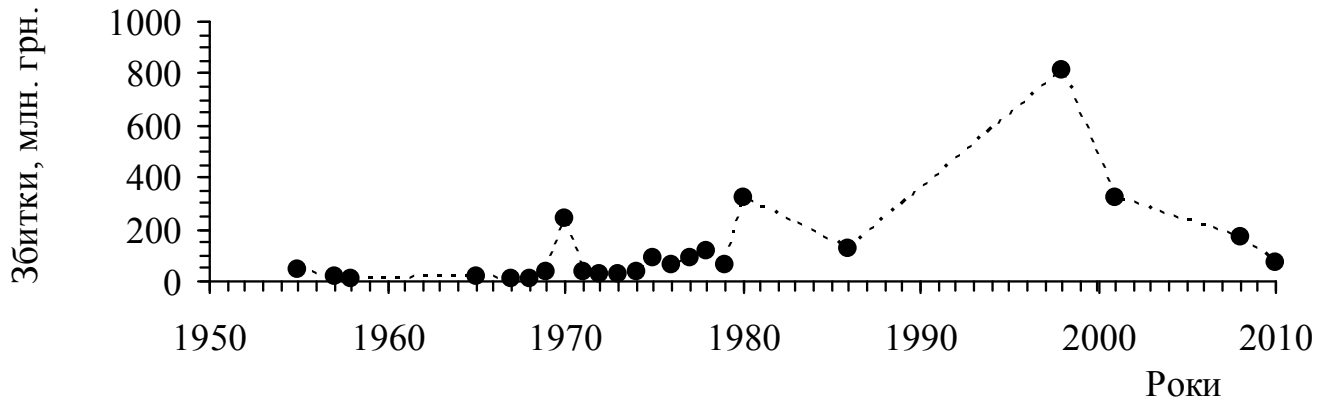


Рисунок 1 – Ряд динаміки збитків від руйнівних повеней в басейні р. Тиса

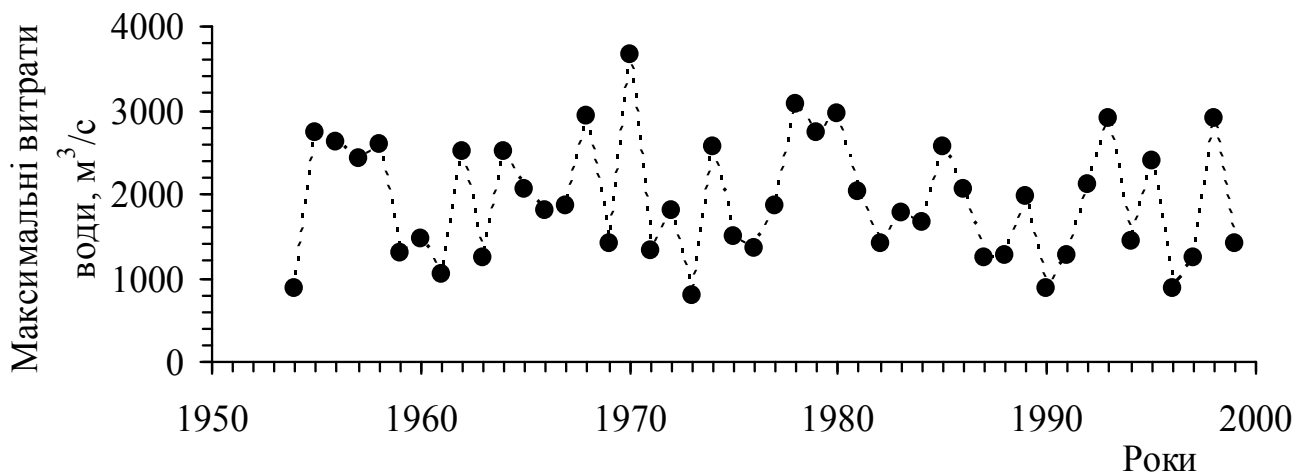


Рисунок 2 – Спостережені максимальні витрати води паводків на р. Тиса (гідрометричний пост Вилोक, Закарпатська область)

Однак візуальний аналіз вибіркового ряду динаміки збитків (рис. 3) на проміжку 1957–1979 р. р. (рис. 3), дозволяє виділити три характерні кластери відносної стаціонарності або монотонності ряду збитків:

- 1) 1957–1969 р. р.;
- 2) 1970–1974 р. р.;
- 3) 1975–1979 р. р.

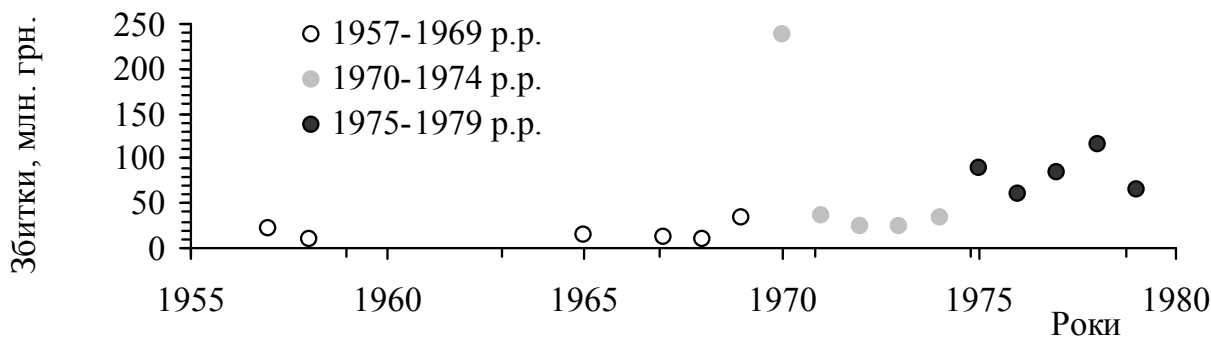


Рисунок 3 – Виділення характерних кластерів відносної стаціонарності або монотонності ряду динаміки збитків

З метою можливості використання відповідних вибірових рядів динаміки (кластерів) для побудови ситуаційних регресійних моделей було проведено кореляційний аналіз даних. Аналізувалися кореляційні зв'язки між збитками та максимальними витратами води на відповідних проміжках часу: 1) з 1957 р. по 1969 р.; 2) з 1970 р. по 1974 р.; 3) з 1975 р. по 1979 р. В результаті на всіх трьох виділених проміжках часу було встановлено добру кореляційну залежність між значеннями максимальної витрати води, як екзогенної змінної, та значеннями ризику збитків $R(D_i)$, як ендогенної змінної, які визначалися як добуток значень збитків D_i від повеней та емпіричних ймовірностей перевищення \hat{P}_i відповідних їм значень максимальних витрат води, що при цьому спостерігалися:

$$R(D_i) = D_i \cdot \hat{P}_i. \tag{1}$$

Тут ймовірності \hat{P}_i визначалися за формулою Н.Н. Чегодаєва [7]:

$$\hat{P}(m) = \frac{m - 0,3}{n + 0,4}, \tag{2}$$

де m – порядковий номер члена упорядкованого у порядку спадання варіаційного ряду; n – загальна кількість членів ряду максимальних витрат води р. Тиса, що спостерігалися на гідрологічному посту Вилوک з 1954 р. по 1979 р. Результати розрахунків зведено в табл. 1.

Таблиця 1 – Ретроспективні ризики збитків від руйнівних повеней в басейні р. Тиса за даними 1957–1979 р. р.

| Роки | Збитки, млн. грн. | Максимальна витрата, м ³ /с | \hat{P}_i , 1/рік | $R(D_i)$, млн. грн. / рік |
|------|-------------------|--|---------------------|----------------------------|
| 1 | 2 | 3 | 4 | 5 |
| 1957 | 21,3 | 2410 | 0,405303 | 8,632955 |
| 1958 | 9,8 | 2600 | 0,253788 | 2,487121 |
| 1965 | 14,2 | 2070 | 0,443182 | 6,293182 |
| 1967 | 10,9 | 1860 | 0,481061 | 5,243561 |
| 1968 | 10,4 | 2930 | 0,102273 | 1,063636 |
| 1969 | 34,4 | 1420 | 0,708333 | 24,36667 |

| 1 | 2 | 3 | 4 | 5 |
|------|-------|------|----------|----------|
| 1970 | 237 | 3650 | 0,026515 | 6,284091 |
| 1971 | 35,5 | 1310 | 0,784091 | 27,83523 |
| 1972 | 24 | 1790 | 0,594697 | 14,27273 |
| 1973 | 23,5 | 783 | 0,973485 | 22,87689 |
| 1974 | 33,6 | 2560 | 0,291667 | 9,8 |
| 1975 | 87,9 | 1500 | 0,632576 | 55,60341 |
| 1976 | 60,1 | 1350 | 0,746212 | 44,84735 |
| 1977 | 85,2 | 1860 | 0,481061 | 40,98636 |
| 1978 | 115,8 | 3060 | 0,064394 | 7,456818 |
| 1979 | 64,4 | 2720 | 0,17803 | 11,46515 |

Коефіцієнти кореляції r між ризиками збитків та максимальними витратами води для виділених часових інтервалів, відповідно, склали: 1) з 1957 р. по 1969 р., $r = -0,8161$; 2) з 1970 р. по 1974 р., $r = -0,8852$; 3) з 1975 р. по 1979 р., $r = -0,9632$.

За результатами кореляційного аналізу було побудовано три ситуаційні регресійні моделі залежності ризиків збитків від максимальних витрат води, які наведено на рис. 4.

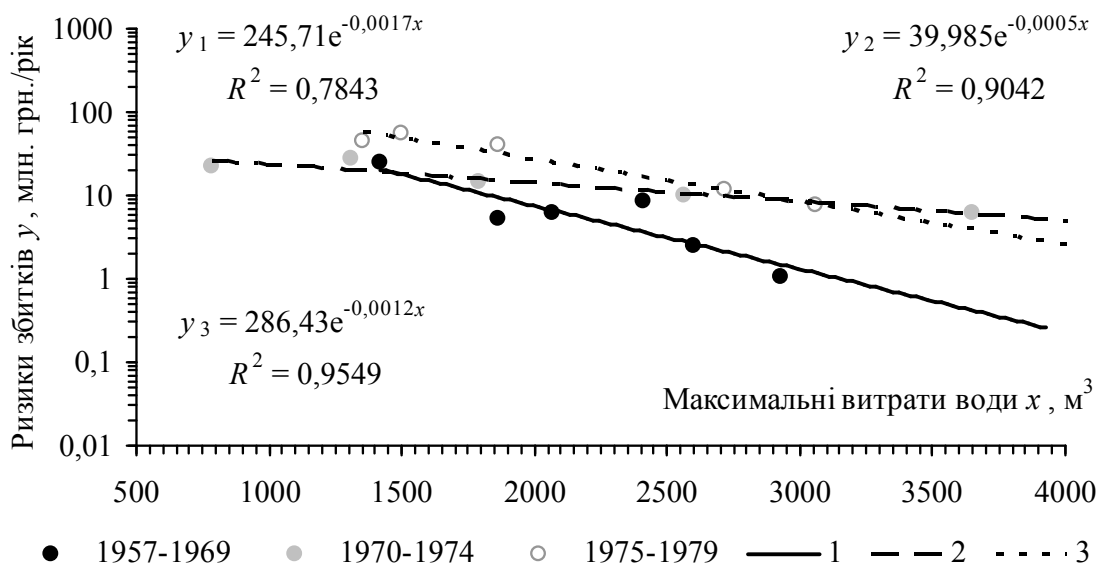


Рисунок 4 – Ситуаційні регресійні моделі залежності ризиків збитків в басейні р. Тиса від максимальних витрат води р. Тиса на гідрологічному посту Вилोक на інтервалі 1957–1979 р. р.

Ситуаційні моделі залежності ризиків збитків в басейні р. Тиса від максимальних витрат води на гідрологічному посту Вилोक, які зображено на рис. 4, володіють високими коефіцієнтами детермінації (R^2) і можуть вважатися адекватними на вибраних інтервалах: 1) з 1957 р. по 1969 р.; 2) з 1970 р. по 1974 р.; 3) з 1975 р. по 1979 р. Відповідно, на основі модельних значень ризиків збитків, отриманих при фіксованих витратах води за допомогою цих ситуаційних моделей, можуть бути побудовані індуктивні моделі (моделі «рівнів»), які, відображаючи еволюцію модельних значень ризику збитків, дозволять відтворити перспективні ситуаційні моделі ризику на заданих періодах упередження прогнозу.

На рис. 5 наведено сімейства індуктивних моделей, виконаних у вигляді трендів на кінець деякого інтервалу часу (а) та на його початок (б) (модельні значення згідно з

рівняннями регресій, які зображено на рис. 4, встановлювалися, відповідно, на кінець і на початок виділених часових інтервалів).

З метою тестування адекватності індуктивних моделей (рис. 5), що будувалися на основі ситуаційних моделей (рис. 4), виконувалася перевірка відповідності між фактичними ризиками збитків від руйнівних повеней, що сталися в 1980, 1986 та 1998 рр. (значення збитків 325 млн грн, 127,9 млн грн, 810 млн грн., відповідно) та їх прогнозними значеннями. Прогнозування на основі індуктивних моделей здійснювалося при спостережених максимальних витратах води р. Тиса на гідрологічному посту Вилоч на 1980, 1986 та 1998 рр., які склали, відповідно, 2970 м³/с, 2050 м³/с, 2900 м³/с. Емпіричні ймовірності перевищення (0,05819; 0,403017; 0,122845), 1/рік, відповідних витрат води перераховувалися для ряду, який охоплював період спостережень з 1954 р. по 1998 р.

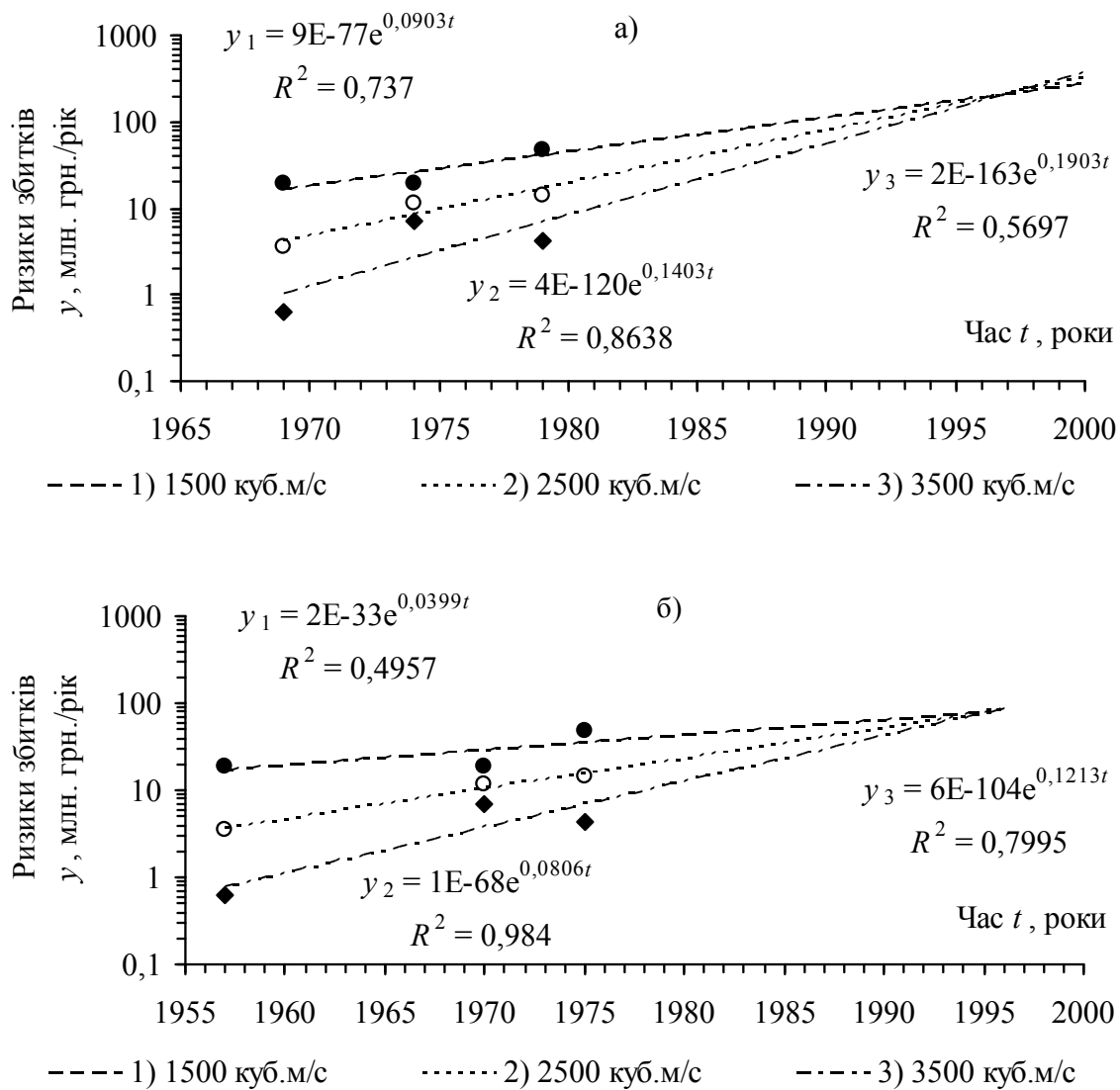


Рисунок 5 – Індуктивні моделі залежності ризиків збитків в басейні р. Тиса від максимальних витрат води р. Тиса на гідрологічному посту Вилоч (за модельними даними з рис. 4)

Результати співставлення фактичних ризиків збитків та їх прогнозних значень наведено на рис. 6.

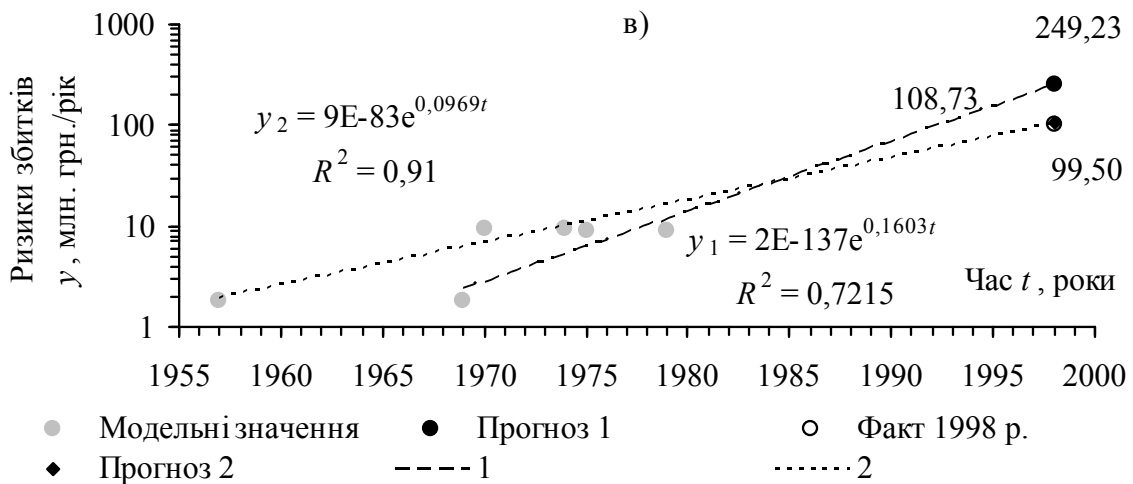
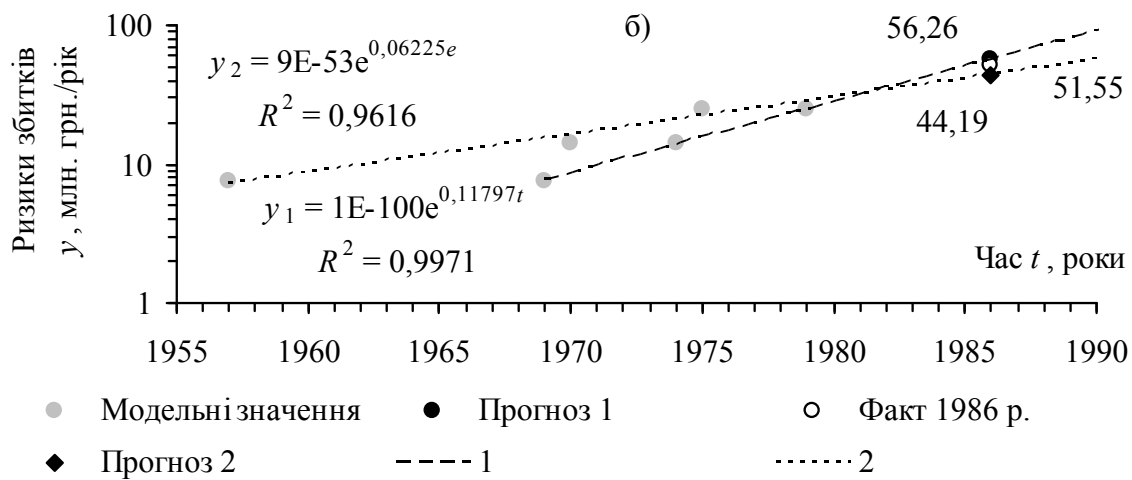
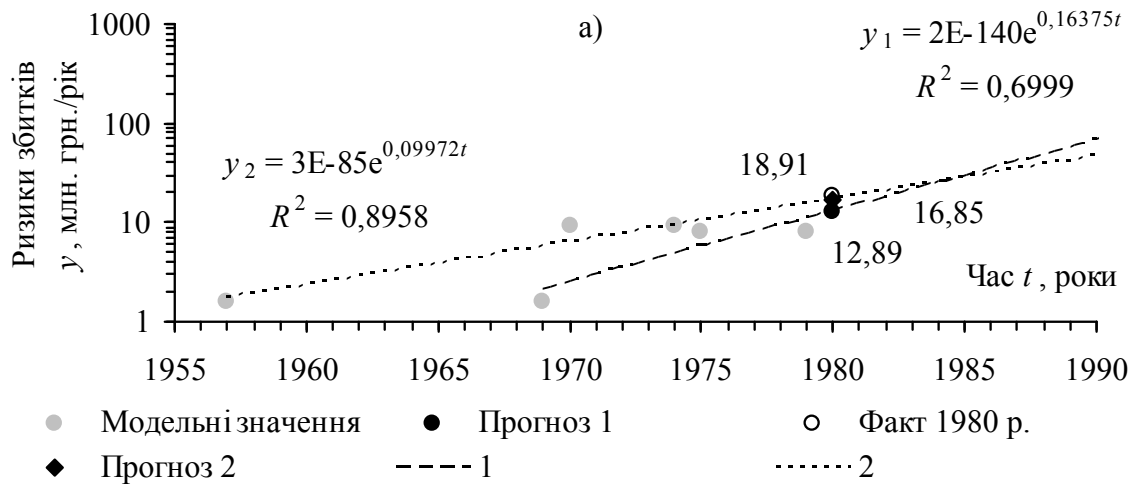


Рисунок 6 – Результати співставлення фактичних ризиків збитків та результатів їх прогнозування для повеней 1980 р. (а), 1986 р. (б), 1990 р. (в)

Виконувалися два прогнози:

- прогноз 1 – прогнозна оцінка ризику збитків для відповідного року як для останнього року на відповідному ситуаційному інтервалі;
- прогноз 2 – прогнозна оцінка ризику збитків для відповідного року як для першого року на відповідному ситуаційному інтервалі.

Фактичний ризик збитків в 1980 р. склав 18,91 млн грн/рік. Найближче до фактичного значення ризику дав прогноз на 1980 р. як на початок нового ситуаційного інтервалу при похибці прогнозування близько 11%.

Фактичний ризик збитків в 1986 р. склав 51,55 млн грн/рік. Найближче до фактичного значення ризику дав прогноз на 1986 р. як на кінець нового ситуаційного інтервалу при похибці прогнозування трохи більше за 9%.

Фактичний ризик збитків в 1998 р. склав 99,50 млн грн/рік. Найближче до фактичного значення ризику дав прогноз на 1998 р. як на початок нового ситуаційного інтервалу при похибці прогнозування близько 9,3%.

Слід відмітити, що на точність прогнозування, особливо на віддалену перспективу, коли ми не можемо достеменно встановити рамки окремих ситуаційних інтервалів на заданому горизонті прогнозування, може суттєво вплинути вибір індуктивної моделі як прогнозної моделі на початок чи кінець ситуаційного інтервалу, що вкладається у вибраний горизонт прогнозування. Зокрема, похибка прогнозування на 1998 р. як на кінець деякого перспективного ситуаційного інтервалу в нашому прикладі перевищила 150%.

Висновки

Встановлено, що вибіркоким рядам динаміки збитків від повеней, що відбуваються в басейні р. Тиса, можуть відповідати адекватні ситуаційні регресійні моделі залежності ризиків збитків від максимальних витрат води р. Тиса на гідрометричному посту Вилок. Показано, що на основі модельних значень ризиків збитків, отриманих за допомогою ситуаційних регресійних моделей, можна побудувати індуктивні моделі, які дозволять відтворити перспективні ситуаційні моделі ризиків збитків від повеней в басейні р. Тиса на заданих періодах упередження прогнозу.

СПИСОК ЛІТЕРАТУРИ

1. Тихонов А.Н. Регуляризирующие алгоритмы и априорная информация / А.Н. Тихонов, А.В. Гончарский и др. – М: Наука, 1983. – 198 с.
2. Моисеев Н.Н. Математические задачи системного анализа / Н.Н. Моисеев. – М.: Наука, 1981. – 487 с.
3. Стефанишина-Гаврилюк Ю.Д. Використання нечіткої міри для подолання невизначеності довгострокових прогнозів на основі екстраполяцій / Ю.Д. Стефанишина-Гаврилюк, Д.В. Стефанишин // Системні дослідження та інформаційні технології. – 2013. – № 4. – С. 99 – 110.
4. Стефанишин Д.В. Про невизначеність прогнозування за даними спостережень / Д.В. Стефанишин // Всеукраїнська науково-практична конференція (ПІКТ – 2013). Проблеми інформатики та комп'ютерної техніки. Тези доповідей. ЧНУ. 27–31 травня, 2013 р. – Чернівці: Видавничий дім «Родовід», 2013. С. 61 – 62.
5. Stefanyshyn D.V. A Method of Forecasting of Indexes of Dynamic System that evolves slowly, based on Time Series Analysis / D.V. Stefanyshyn // Proc. of 4th Int. Conf. on Inductive Modelling. Kyiv, Ukraine, September 16-20, 2013. – P.P. 221 – 224.
6. Стефанишина-Гаврилюк Ю.Д. Про один підхід до прогнозного моделювання поведінки складної системи за натурними даними / Ю.Д. Стефанишина-Гаврилюк, Д.В. Стефанишин // Проблеми інформатики та комп'ютерної техніки (ПІКТ – 2014). Праці III-ї міжнародної науково-практичної конференції. Тези доповідей. ЧНУ. 27–30 травня, 2014 р. – Чернівці: Видавничий дім «Родовід», 2014. С. 72 – 73.
7. Железняков Г.В. Гидрология, гидрометрия и регулирование стока / Под ред. Г.В. Железнякова. – М.: Колос. – 1984. – 205 с.

Стаття надійшла до редакції 17.11.2014