

УДК 504.3

С. М. ЮРАСОВ канд. техн. наук, доц., **О. А. АЛЕКСЕЄНКО**

Одеський державний екологічний університет, м. Одеса

ул. Львовская, 15 м. Одесса, 65016

lovely_lena@ukr.net

СТАТИСТИЧНІ ІНСТРУМЕНТИ АНАЛІЗУ ЧАСОВИХ ЗАКОНОМІРНОСТЕЙ ЗМІНИ ЯКОСТІ ВОДНИХ ОБ'ЄКТІВ

Виконано детальний статистичний аналіз показників якості водного середовища річки Великий Куяльник Котовського району Одеської області. Виявлені закономірності розподілу показників якості водного середовища та отримані тренди, що відбивають загальні тенденції зміни показників за період з 2004 по 2014 років.

Ключові слова: показники забруднення водного середовища, інструменти статистичного аналізу

Urasov S. N., Alekseenko E. A. STATISTICAL TOOLS FOR ANALYSIS OF TEMPORAL CHANGES IN WATER QUALITY

The research is about detailed statistical analysis of water quality for the river 'Big Kuyalnik' which is Odessa region. The water quality indexes were analyzed and their distributions were assessed. The time trends of water indexes were obtained for 2004-2014.

Key words: water quality indexes, statistical analysis tools

Юрасов С. Н., Алексеенко Е. А. СТАТИСТИЧЕСКИЕ ИНСТРУМЕНТЫ АНАЛИЗА ВРЕ- МЕННЫХ ЗАКОНОМЕРНОСТЕЙ ИЗМЕНЕНИЯ КАЧЕСТВА ВОДНЫХ ОБЪЕКТОВ

Выполнен детальный статистический анализ качества водной среды реки Большой Куяльник Одесской области. Определены закономерности распределения показателей качества водной среды и получены тренды, отражающие общин тенденции изменения показателей за период с 2004 по 2014 годы.

Ключевые слова: показатели загрязнения водной среды, инструменты статистического анализа

© Юрасов С. М., Алексеенко О. А., 2015

Вступ

Постановка проблеми та її зв'язок із важливими практичними завданнями.

Одним із шляхів отримання інформації щодо якості водного середовища є детальний статичний аналіз відповідних показників. Однак на першому ж етапі виникає необхідність вибору найбільш зручних інструментів, застосування яких відповідає поставленій меті. Таким чином вирішення задачі аналізу якості водного середовища повинна враховувати: визначення наявних часових рядків показників якості водного середовища та визначення інструментарію статистичної обробки. В даний час існує велика кількість алгоритмів, що дозволяють виконати всебічний аналіз інформації, однак не існує єдиної схеми комбінування цих методів, оскільки кожна окрема проблема потребує індивідуального підходу. В рамках даного дослідження розглянуті питання оцінки якості водного середовища, і пошук оптимальної комбінації статистичних методів, що нададуть максимально повний опис якості водного середовища для окремого водного об'єкту.

Аналіз останніх досліджень і публікацій. Статистичному аналізу інформації,

що характеризує якість водного середовища в теперішній час присвячено багато досліджень, але значна частина не ставить своєю метою вирішення проблеми формування універсальної сукупності алгоритмів, що можуть бути застосовані для отримання повної картини, що відображає якість статистичних даних. Так в роботах, аналогічних [1] і [2] увага переважно надається апроксимації законів розподілу показників якості водного середовища чи пошуку окремих закономірностей просторового розподілу показників [3]. Це свідчить про актуальність пошуку оптимальної комбінації методик математичної статистики [4, 5] для всебічного аналізу експериментальних даних. Деякі попитки вирішити аналогічні завдання предприняті рядом західних авторів [6, 7], і відповідно потребують прикладного застосування для вітчизняних умов.

Мета роботи - визначення комбінації послідовного застосування набору статистичних інструментів на прикладі якості води річки Великий Куяльник Котовського району Одеської області.

Матеріали і методи дослідження

Вирішення проблеми оцінки якості довкілля може бути представлено у вигляді наступної загальної схеми [Statistical Guidance for Determining Background Ground Water Quality and Degradation] (рис. 1). Відповідно мети, в даному дослідженні основна увага надається центральному блоку схеми – визначення інструментів статистичного аналізу.

Вихідними даними є інформація щодо якості води малої річки у Котовському районі Одеської області за наступними показниками: рН, розчинений кисень, NH_4^+ , NO_2^- , NO_3^- , БСК₅. Дані отримані за період з 2004 по 2014 роки. Часові ряди представлені середньомісячними значеннями.

В якості інструментів статистичного аналізу використані наступні методи:

Boxplot (діаграма розмаху). Такий вид діаграми в зручній формі показує медіану, мінімальне і максимальне значення вибірки, а також викиди. Кілька таких «ящиків» можна намалювати пліч-о-пліч, щоб візуально порівнювати один розподіл з іншим; їх можна розташовувати як

горизонтально, так і вертикально. Відстані між різними частинами ящика дозволяють визначити ступінь розкиду (дисперсії) і асиметрії даних і виявити викиди.

Scatter plots (діаграма розсіювання) Рівняння регресії. На діаграмі розсіювання кожному спостереженню (або елементарній одиниці набору даних) відповідає точка, координати якої (в декартовій системі координат) дорівнюють значенням двох певних параметрів відповідно осі параметрів цього спостереження. Якщо передбачається, що один з параметрів залежить від іншого, то зазвичай значення незалежного параметра відкладається по горизонтальній осі, а значення залежного – по вертикальній. Діаграми розсіювання використовуються для демонстрації наявності або відсутності кореляції між двома змінними. При наявності лінійного зв'язку між величинами доцільно після отримання діаграм розсіювання розраховувати рівняння регресії та будувати пряму регресії.

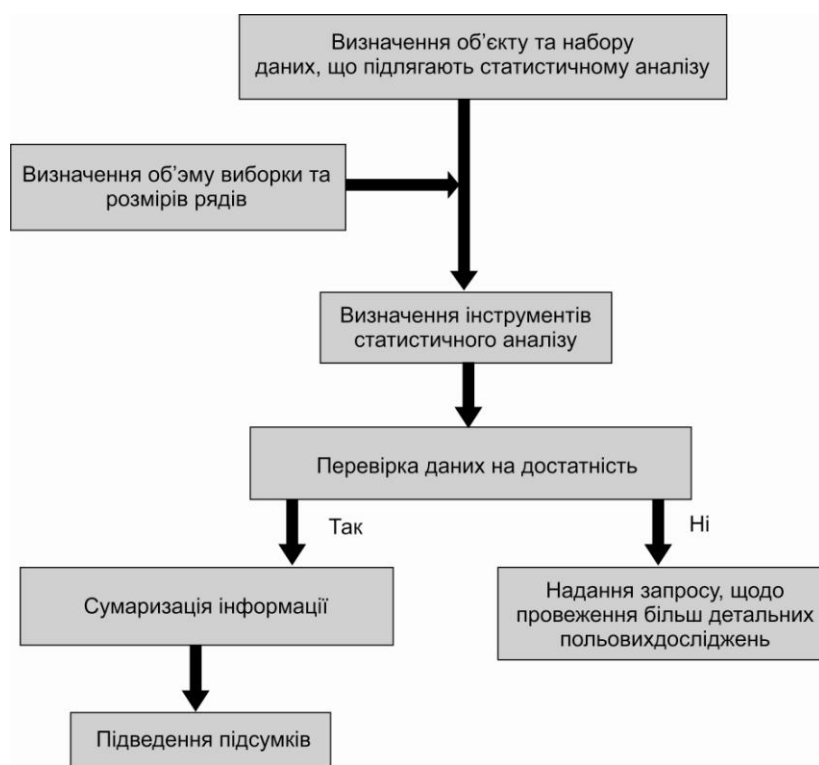


Рис. 1 – Загальна схема статистичної обробки даних стану якості довкілля

Q-Q plots. (Графік Q-Q) — імовірнісний графік, який являє собою графічний метод для порівняння двох розподілів ймовірностей, ставлячи їх квантилі один проти одного. По-перше, набір інтервалів для квантилів — заданий. Точки (x, y) на графіку відповідає один з квантилів другого розподілу (y -координата), побудована разом з аналогічним квантилем першого розподілу (x -координата). Таким чином, лінія є параметричною кривою з параметром, який є кількістю інтервалів для квантилів. Якщо два порівнюваних розподілів схожі, точки в графіці Q-Q будуть приблизно лежати на прямій $y = x$. Якщо розподіли лінійно пов'язані, точки в графіці Q-Q будуть приблизно лежати на одній прямій, але не обов'язково на прямій $y = x$. Графік Q-Q також може бути використана в якості графічного засобу оцінки параметрів в масштабі сімейства розподілів.

Trend detection (визначення трендів). загальна тенденція при різно-направленому русі, визначена загальною спрямованістю змін показників часового ряду. Графіки можуть описуватись різними рівняннями – лінійними, логарифмічними, степеневими і т. ін. Фактичний тип графіка встановлюють

за графічним зображенням даних часового ряду, шляхом усереднення показників часового ряду, на основі статистичної перевірки гіпотези про сталість параметрів графіка. В рамках даного дослідження застосований метод ковзаючого середнього для отримання сгладжування для коротких часових коливань і побудова тренда із застосуванням полінома третього ступеню для довгочасових коливань.

Ковзаюче середнє застосовується для розрахунку значень в періоді, що прогнозується, на базі середнього значення змінної для встановленого числа попередніх періодів. Ковзаюче середнє, на відміну від простого середнього для всієї вибірки містить інформацію щодо тенденції зміни даних. Розрахунок проводиться за наступною формулою:

$$F_{t+1} = \frac{1}{N} \sum_{j=1}^N A_{t-j+1}$$

де:

- N — кількість попередніх періодів, що входять до ковзаючого середнього;
- A_j — фактичне значення у момент часу j ;
- F_j — прогнозне значення у момент часу j .

Результати дослідження

Перший етап аналізу даних припускає виявлення інтервалів розмаху величин, що розглядаються. На рис. 2 наведені діаграми розмаху для шести інгредієнтів, що відповідають за якість водного середовища. Як видно із діаграм найбільший розмах значень спостерігався для періоду з 2010 по 2014 роки. Також слід зазначити відносно

високий рівень розсіювання для O_2 , NH_4^+ та BCK_5 (рис. 2. б, в, є).

В якості наступного кроку аналізу висунуто гіпотезу щодо наявності лінійного зв'язку між наданими інгредієнтами. Для цього побудовано матрицю попарних кореляцій (табл. 1).

Як видно із таблиці, переважна кількість коефіцієнтів попарної кореляції

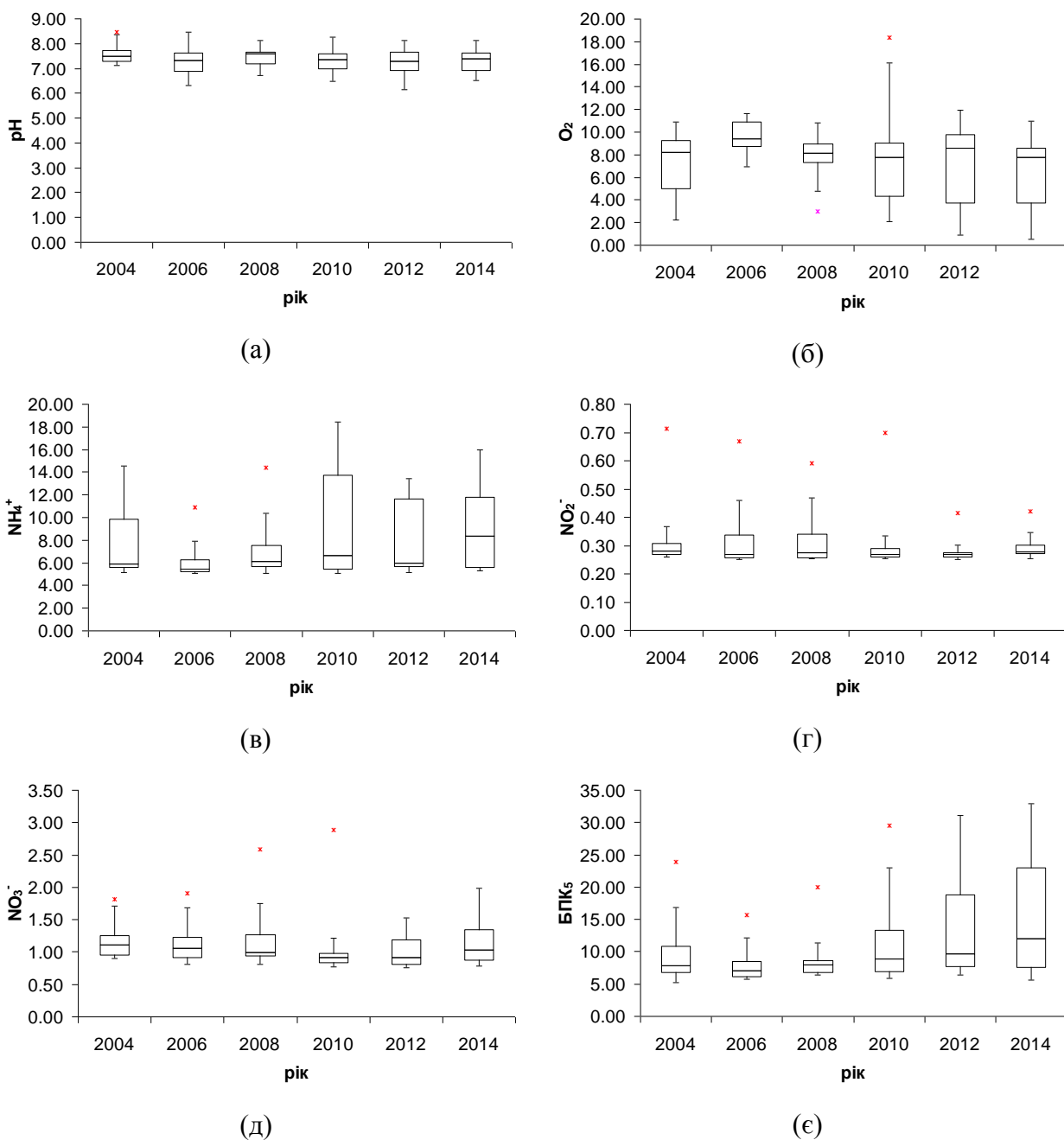


Рис. 2 – Діаграми розмаху для показників якості водного середовища, мг/дм³

досить низка, але перевірка на значимість визначила, наявність лінійного зв'язку між трьома показниками: БПК₅, O₂, NH₄⁺. Це свідчило про доцільність побудови рівняння регресії, що характеризує залежність се-

реднього значення БПК₅ від значень концентрації O₂, NH₄⁺. На рисунках 3 і 4 наведені хмари точок, що відповідають диграмам розсіювання та відповідні лінії регресії.

Таблиця 1

Матриця кореляцій показників якості водного середовища

Показник	pH	O ₂	NH ₄ ⁺	NO ₂ ⁻	NO ₃ ⁻
pH	1.00				
O ₂	-0.07	1.00			
NH ₄ ⁺	0.22	-0.54	1.00		
NO ₂ ⁻	-0.13	-0.13	0.00	1.00	
NO ₃ ⁻	0.07	0.13	0.12	0.15	1.00
БСК ₅	0.11	-0.66	0.70	0.07	0.01

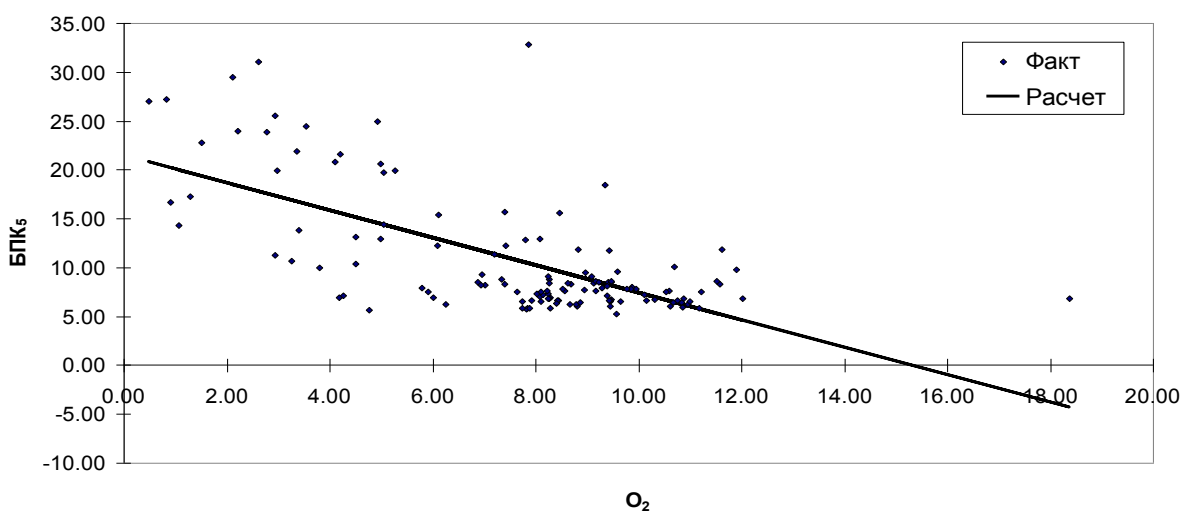


Рис. 3 – Лінія регресії, що відповідає залежності БПК₅ від O₂, мг/дм³

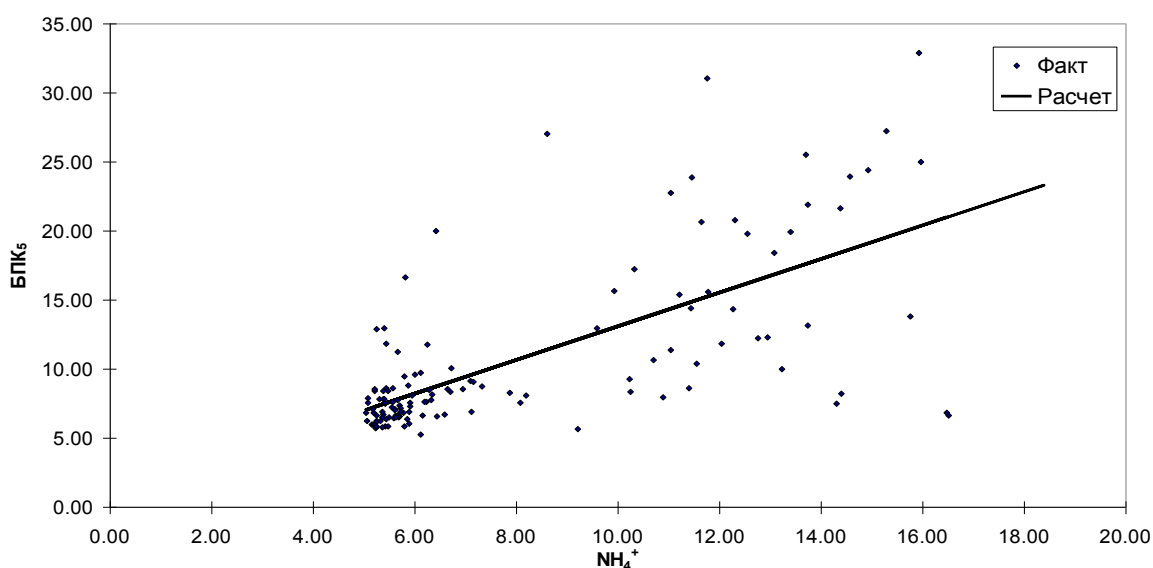


Рис. 4 – Лінія регресії, що відповідає залежності БПК₅ від NH₄⁺, мг/дм³

Як видно із наведених графіків поле точок досить упорядковане уздовж ліній регресії, що підтверджує гіпотезу о наявності лінійного зв'язку між значеннями даних показників.

Наступний етап – апроксимація зако-

нів розподілу показників, що розглядаються, нормальному розподілу. Для цього застосовані Q-Q графіки. Як видно із графіків найбільш близькими к нормальному розподілу є ряди значень рН, NO_3^- і БСК₅ (рис. 5, а, д, є).

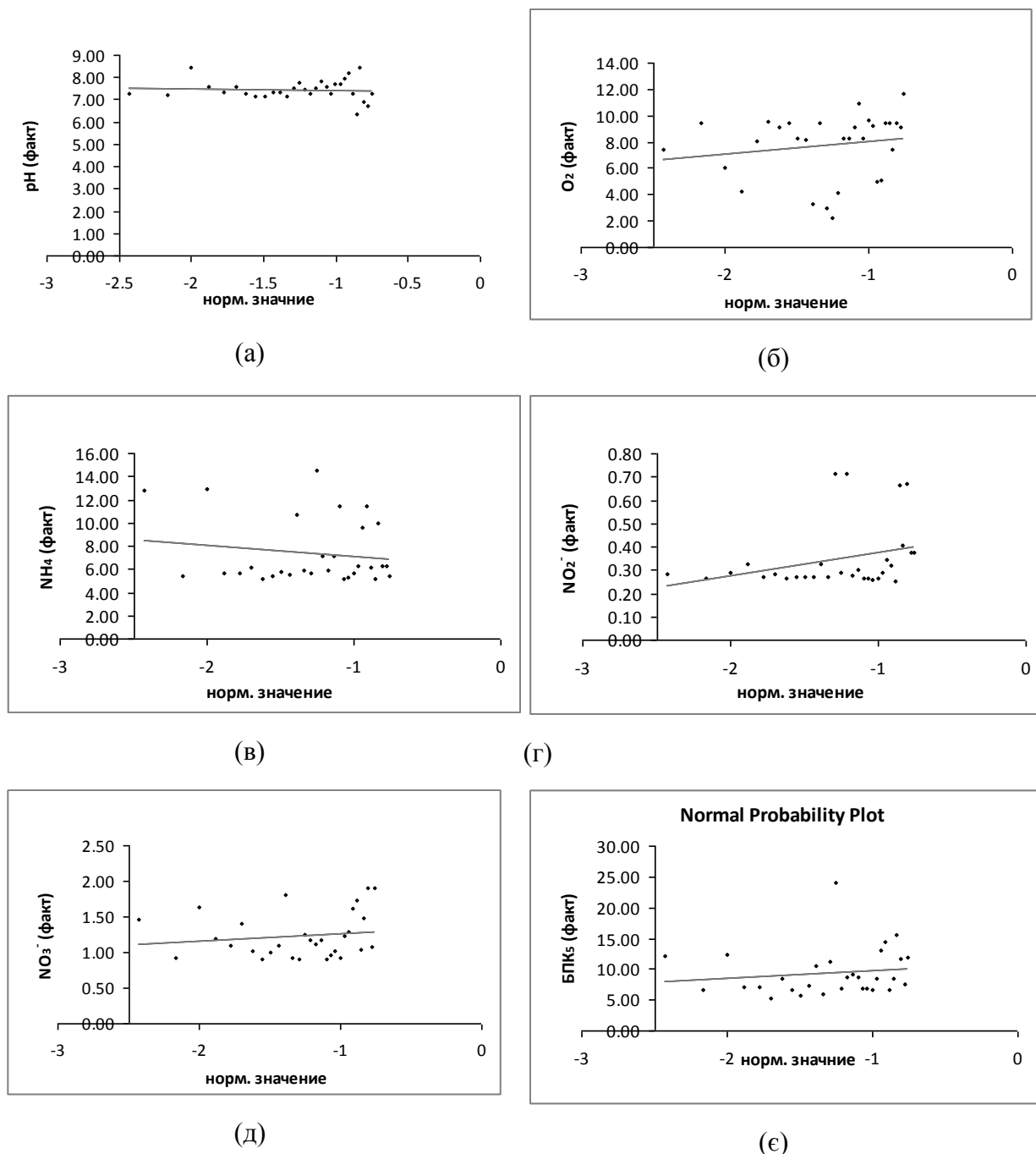


Рис. 5 – Графіки Q-Q для показників якості водного середовища, мг/дм³

Останнім кроком дослідження є пошук загальних тенденцій змінення значень показників якості водного середовища. Для

цього використані метод ковзаючого середнього та поліном третього ступеню. Даний вибір оснований на тому, що метод ковзаю-

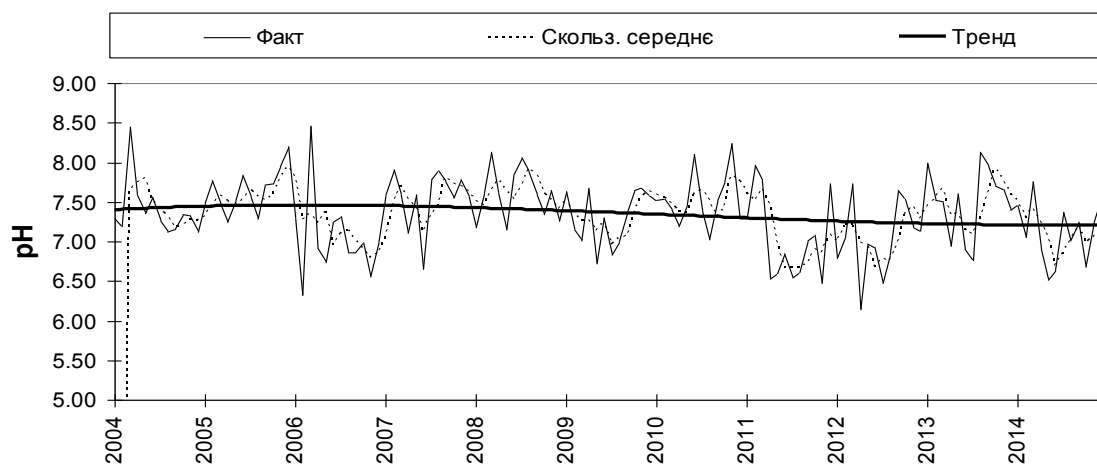


Рис. 6 – Часовий хід та трендові криві рН

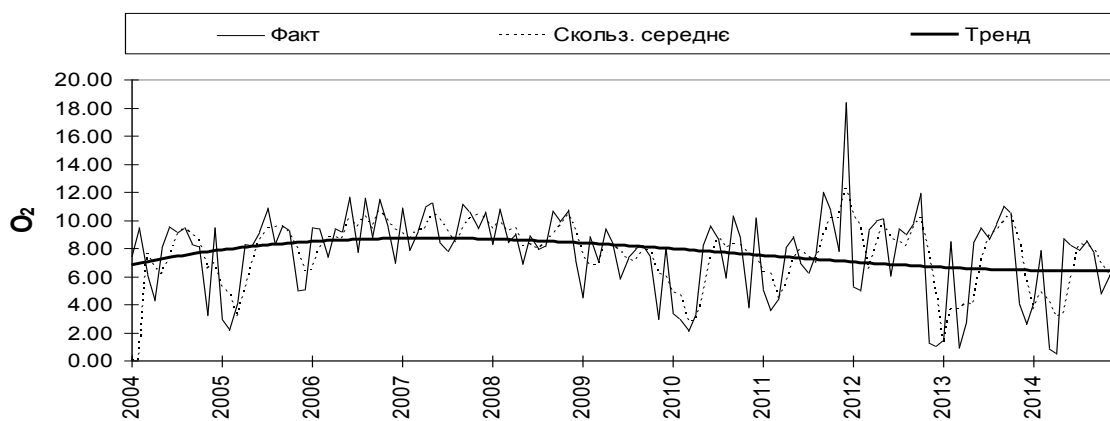


Рис. 7 – Часовий хід та трендові криві O_2 , мг/дм³

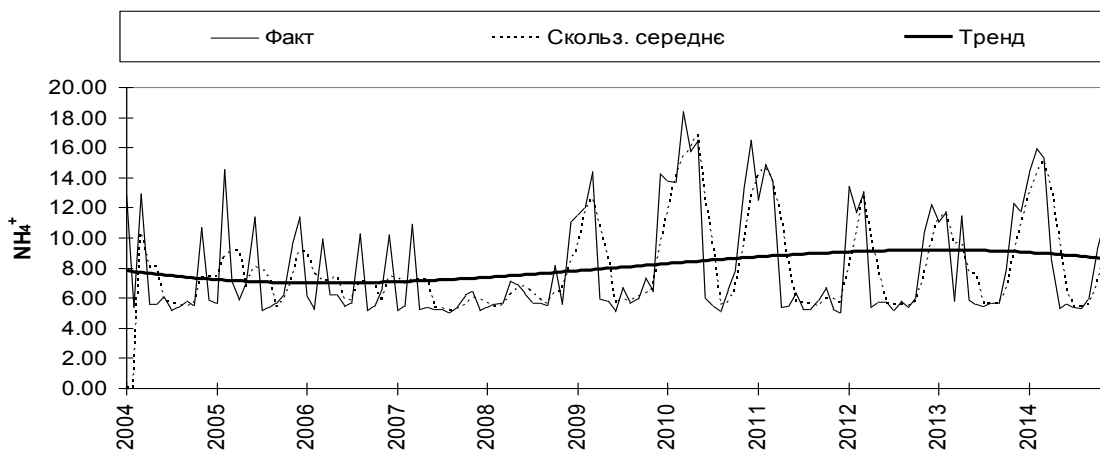


Рис. 8 – Часовий хід та трендові криві NH_4^+ , мг/дм³

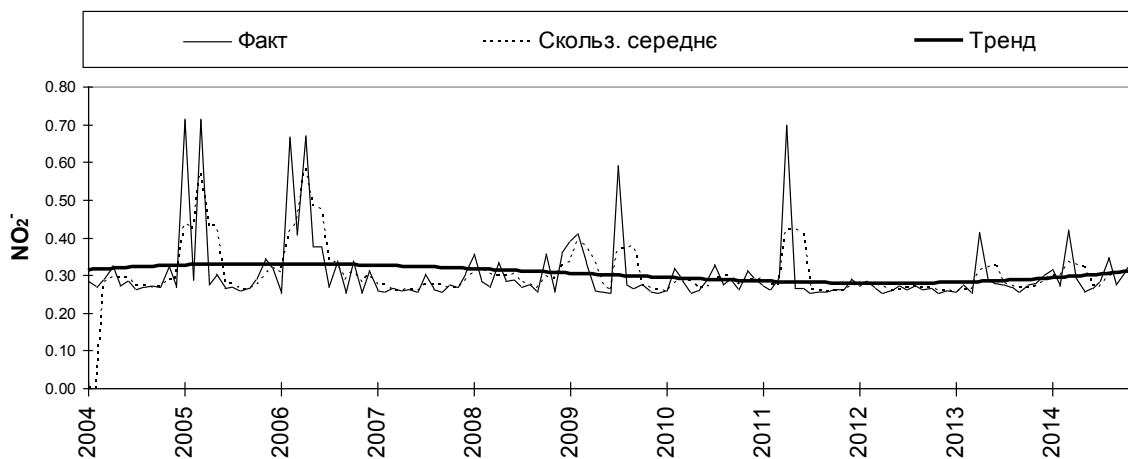


Рис. 9 – Часовий хід та трендові криві NO_2^- , мг/дм³



Рис. 10 – Часовий хід та трендові криві NO_3^- , мг/дм³



Рис. 11 – Часовий хід та трендові криві БСК_5 , мг/дм³

чого середнього враховує короткострокову інерційність випадкової величини, а застосування поліному третього ступню надає можливість виявити загальну тенденцію. На рисунках 6-11 приведені трендові криві, що відповідають показникам якості водного середовища. Для рН O_2 за розглянутий період спостерігається тенденція поступового

зменшення починаючи із 2007 року. Для NH_4^+ найбільші значення спостерігались з 2011 по 2013 роки. Значення показника NO_2^- за останній період 2011-2014 характеризувалися порівняно різким зменшенням значень. Для NO_3^- та BCK_5 спостерігалась тенденція зростання.

Висновки

Статистичні інструменти важливі при аналізі якості води, тому що якість води визначається чисельними показниками. В результаті дослідження запропонована комбінація статистичних методів, що дозволяють надати відносно повну картину стосовно якості водного середовища. Розрахунки проведені на прикладі малої річки Великий Куяльник (Котовський район, Одеської області). На основі проведених розрахунків визначено:

- найбільш високий рівень розсіювання

значень показників спостерігався для NH_4^+ та BCK_5 у період з 2010 по 2014 рр;

- наявність лінійного зв'язку відповідає парам показників: BPK_5 , O_2 та BPK_5 від NH_4^+ ;
- нормальному розподіленню відповідають статистичні сукупності рН, NO_3^- і BCK_5 ;
- тенденція збільшення за останні роки відповідає показникам NO_3^- і BCK_5 .

Література

1. Рідей Н. М. Екологічний контроль за якістю поверхневих водних джерел озер навчально-дослідного господарства «Великоснітинське» ім. Музиченка / Н. М. Рідей, І. В. Захаркевич. // Вісник Запорізького національного університету. – №2 – Запоріжжя: 2008. – С. 172-176.

2. Ковальчук Л. А. Вероятностно - статистическое оценивание качества поверхностных вод по категориям / Л. А. Ковальчук, Н. Н. Осадчая, В. И. Осадчий // Наук. праці УкрНДГМІ -№257 – К.: 2008. – С.162-175

3. Матеріали семінару «Основи природоохоронного законодавства України та Європейського співтовариства: водні ресурси». – К.: Державний інститут підвищення кваліфікації та перепідготовки кадрів Мінекобезпеки України, 1997. – 71 с.

4. Козлов М. В., Прохоров А. В. Введение в

математическую статистику. – М.:Изд-во МГУ, 1987. – 264 с.

5. Боярский А. Я. Теоретические исследования по статистике / А. Я. Боярский: сборник научн. Трудов. – М.: Статистика, 1974 – 304с.

6. Liya Fu1, You-GanWang. Statistical Tools for Analyzing Water Quality Data. / Water Quality Monitoring and Assessment. InTech Europe, 2012 – P. 143-168.

7. V. Z. Antonopoulos, D. M. Papamichail, K. A. Mitsiou. Statistical and trend analysis of water quality and quantity data for the Strymon River in Greece. Hydrology and Earth System Sciences Discussions, Copernicus Publications, 2013, 5 (4), – pp.679-692.

Надійшла до редколегії 29.09.2015

