

ПРОБЛЕМИ ВЕРИФІКАЦІЇ ХАРАКТЕРИСТИК СТИЧНОЇ ВОДИ

І. С. Єремєєв, докт. техн. наук (Державна академія житлово-комунального господарства, м. Київ), А. О. Дичко, канд. техн. наук (НТУУ «КПІ»)

Застосування порівняльного аналізу даних «найближчих сусідів» забезпечує уточнення характеристик стічної води, а також дає змогу виявити справжні причини змін показників та їхній взаємний вплив.

Ключові слова: стічні води, верифікація, невизначеність параметрів.

Применение сравнительного анализа данных «ближайших соседей» обеспечивает уточнение характеристик сточных вод и дает возможность выявить истинные причины изменения показателей и их взаимное влияние.

Ключевые слова: стоки, верификация, неопределенность параметров.

Application of comparing analysis of “the nearest neighborhoods” method provides specification of wastewater characteristics and allows to determine real causes of values change and their interference.

Key words: wastewater, verification, parameters uncertainty.

Вступ. Інформаційне забезпечення комплексів, призначених для повної біологічної очистки стічних вод та обробки затриманих забруднень, характеризується великою кількістю даних, що збираються, як за номенклатурою, так і з точки зору діапазону та кількості точок відбирання проб. Умови відбирання проб, супутні чинники, чутливість приладів та методик вимірювання, перешкоди тощо впливають на якість інформації і можуть призвести навіть до неадекватної оцінки ситуації, що склалася. Тому проблема верифікації отриманих даних має суттєве значення.

Постановка проблеми в загальному вигляді і її зв'язок з важливими науковими і практичними завданнями. Під час збору і оцінювання даних можуть траплятися такі характерні ситуації (особливо у разі аналізу води на виході комплексу очисних споруд):

- компонент, що аналізується, може бути присутнім у пробі, але його значення знаходиться в межах нечутливості детектора або недосконалості методики, яка використовується для завдань аналізу (НЧ);
- має місце ефект матриці (ЕМ);
- компонент має «фонове» значення (ФЗ), тобто значення, характерне для «пустої» проби;
- дані не можуть бути використані через недосконалість процедури аналізу і невідповідність критерію якості (НЯ).

В усіх подібних ситуаціях отримані дані повинні супроводжуватися відповідними маркерами, як зазначено вище.

Аналіз останніх досліджень і публікацій, в яких започатковано розв'язання даної проблеми і на які спирається автор. Поводження зі стічними водами характеризується використанням осереднених за часом та за обсягом даних, що пов'язане із статистичними похибками типу стандартного відхилення σ , коли у межах подвійного відхилення 2σ існує невизначеність

щодо істинного значення того чи іншого параметра, а також з наявністю так званих довірчих інтервалів. Іншими словами, мають місце певні чинники, які призводять до неоднозначності оцінок.

Можна підрахувати комбіновану стандартну невизначеність даних за допомогою такого спрощеного рівняння (якщо прийняти, що усі невизначеності підсумовуються), яке характеризує результат процесу вимірювання:

$$y = a_1x_1 + a_2x_2 + \dots + a_Nx_N,$$

де x_i – оцінки вхідних даних X_i , а a_i – певні константи (коефіцієнти ваги тих чи інших складових вимірювань).

При цьому стандартна інтегральна невизначеність може бути охарактеризована виразом

$$u_c^2(y) = a_1^2 u^2(x_1) + a_2^2 u^2(x_2) + \dots + a_N^2 u^2(x_N),$$

а відносна стандартна невизначеність кожного з параметрів x_i $u_r(x_i)$ характеризується як $u_r(x_i) = u(x_i)/|x_i|$, де $|x_i|$ – абсолютне значення x_i , яке не дорівнює нулю. При цьому відносна комбінована стандартна невизначеність може бути охарактеризована як $u_{c,r}(y) = u_c(y)/|y|$, де $|y|$ – абсолютне значення y , яке не дорівнює нулю [1].

Виділення не розв'язаних раніше питань і формулювання цілей дослідження. Якщо імовірнісний розподіл результату вимірювання y і його комбінована стандартна невизначеність $u_c(y)$ характеризуються нормальним (гаусовим) законом розподілу і $u_c(y)$ є надійною оцінкою стандартного відхилення y , то інтервал від $y - u_c(y)$ до $y + u_c(y)$ є очікуванням вмісту близько 68 % значень, які можуть бути надійно віднесені до величини Y , що є оцінкою y . Такий підхід дозволяє робити достовірні висновки щодо меж, в яких знаходяться ті чи інші параметри, але всередині цих меж параметр вважається незмінним. В більшості випадків цього цілком достатньо для оцінювання ситуації в статистиці або в межах окресленого інтервалу часу, але в разі необхідності визначення причинно-наслідкових зв'язків і прогнозування подальшого розвитку подій подібний метод не є плідним. Замість цього варто розглянути підхід, який базується на оцінюванні значень «найближчих сусідів», які знаходяться в прийнятих межах 2σ . Суть його полягає у тому, що зіставляються дані про характеристики об'єкта досліджень у точці, яка викликає зацікавленість, з даними у найближчих (у часі чи просторі) до неї сусідніх точках об'єкта. Якщо спостерігається узгоджений рух показників станів в усіх сусідніх точках у межах стандартного відхилення, можна вважати, що має місце певна тенденція, якій можна довіряти.

Метою роботи є застосування методу «найближчих сусідів» при верифікації характеристик стічної води.

Викладення основного матеріалу дослідження. Для оцінювання згаданої вище тенденції можна використати таку евристику:

ЯКЩО ($[MCTAB]$ I $[SIGN\Delta ST_{i+j}EQ]$), ТО $[СПТ]$,

де $MCTAB$ – ситуація, коли величина, що вимірюється, не виходить за межі стандартного відхилення; $SIGN\Delta ST_{i+j}EQ$ – результат спостереження, який свідчить, що знак зміни стану при поточному вимірюванні в усіх точках зліва та справа від i , тобто у точках від i до $i-j$ та від i до $i+j$, однаковий відносно стану у тих же точках під час попереднього вимірювання або у просторовому розподілі; $СПТ$ – спостерігається певна тенденція (рис. 1).

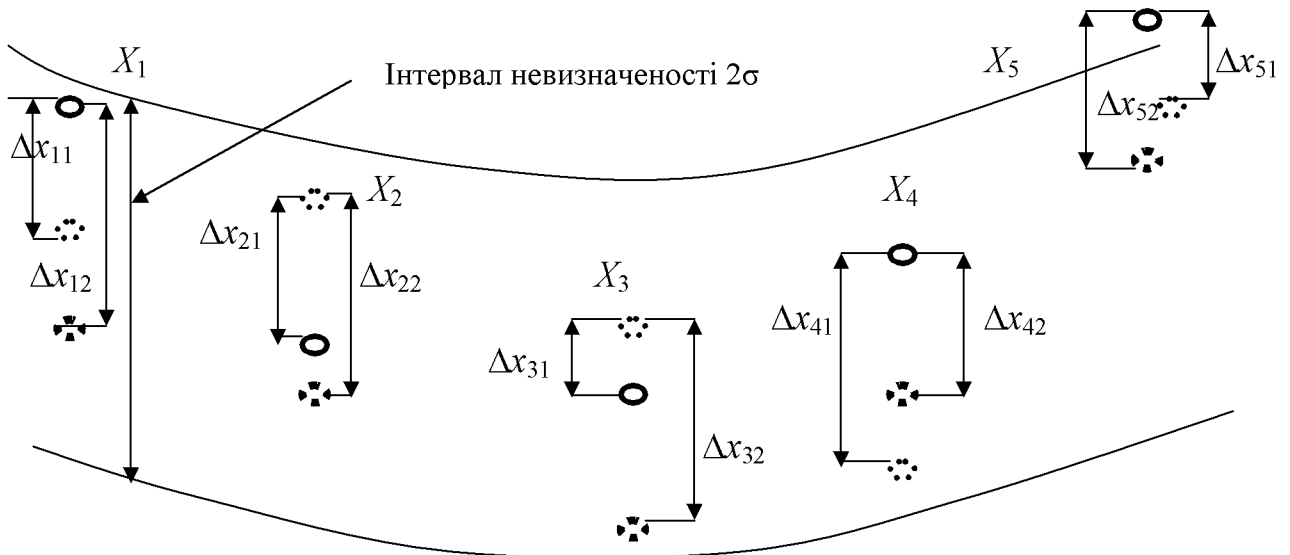


Рис. 1. Графічна інтерпретація методу порівняльного аналізу даних «найближчих сусідів»

Цікаво розглянути співвідношення кількості різних забруднень в стічних водах Бортницької станції аерації м. Києва, які були виявлені впродовж 2001–2006 рр., а також потік води крізь очисні споруди з точки зору їхньої невизначеності. Дані щорічних (осереднених) значень забруднень і витрати води та їхня статистична обробка наведені на рис. 2 і 3. З цих результатів очевидно, що спостерігається сталий позитивний тренд усіх видів забруднень, зокрема азоту амонійного та фосфатів. Це свідчить у першу чергу про зростання населення Києва, а також про зміну питомої ваги певних забруднювачів у скидах і повинно прийматися до уваги як при плануванні ресурсного забезпечення процесів біологічного очищення, так і при плануванні чергової реконструкції очисних споруд. У даному разі наведені довірчі інтервали 95 % можна розглядати як вихідні дані для отримання стандартної невизначеності, яка може бути розрахована (за умов реального нормального розподілу даних) шляхом поділу інтервалу ($a_+ - a_-$) на коефіцієнт 1,96. У разі, коли нормальний розподіл розглядається лише як певне наближення, можна використати інший підхід, коли $u_c(y)$ визначається за допомогою формули

$$u_c(y) = a/3,$$

де $a = (a_+ - a_-)/2$ і a_+ та a_- – відповідно верхня та нижня границя довірчого інтервалу.

Використання коридору 2σ або довірчих інтервалів не можна вважати ознакою недостовірності даних. Навпаки, знаходження параметру в межах коридору є тим фактом, який підтверджує достовірність отриманих даних, але в той же час свідчить про те, що обмеження методики або властивостей детекторів не дозволяють уточнити дані всередині означених меж.

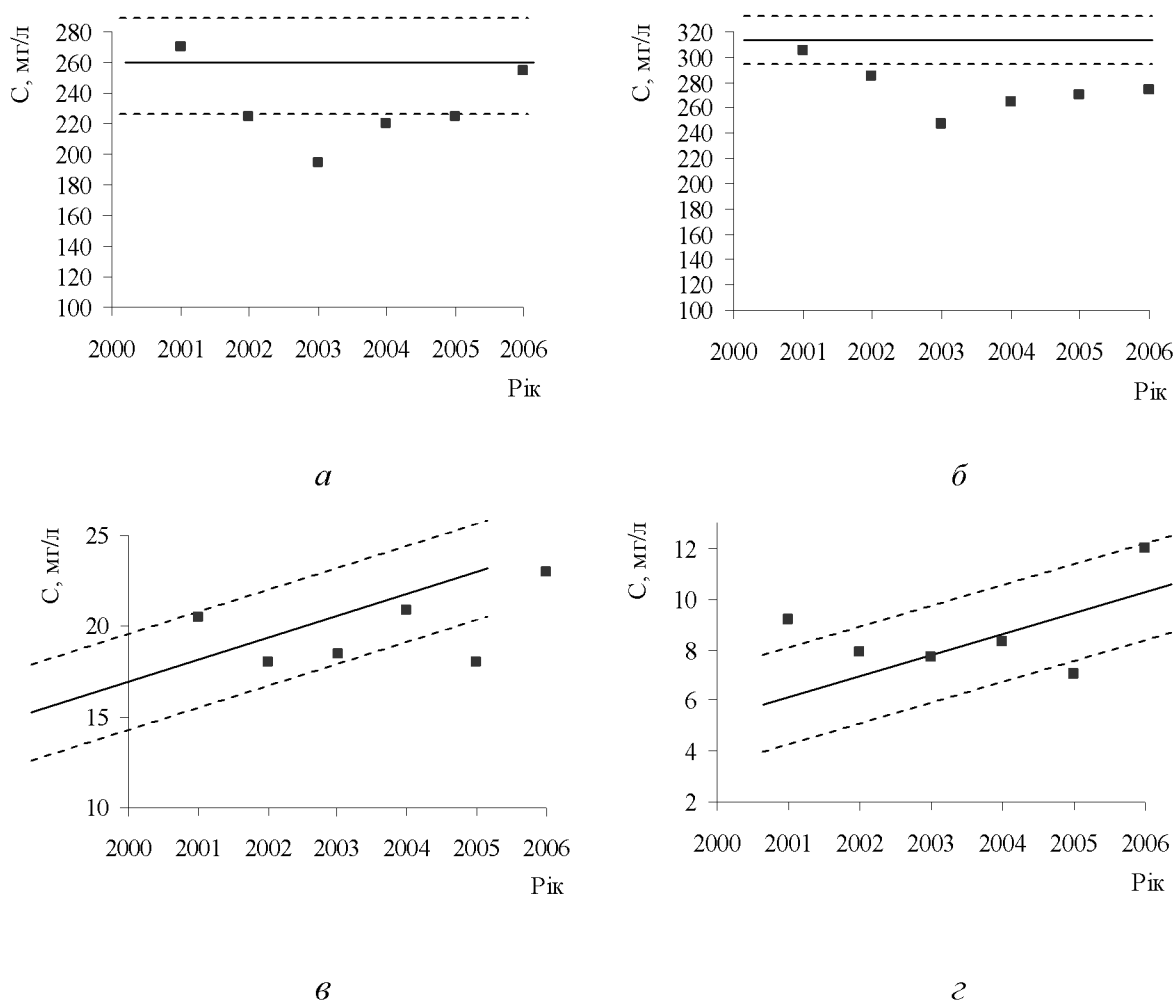


Рис. 2. Зміна концентрації забруднюючих речовин в стічних водах (C , мг/л) на Бортницьких очисних спорудах протягом 6 років: *а* – завислих речовин; *б* – БСК₂₀; *в* – азоту амонійного; *г* – фосфатів

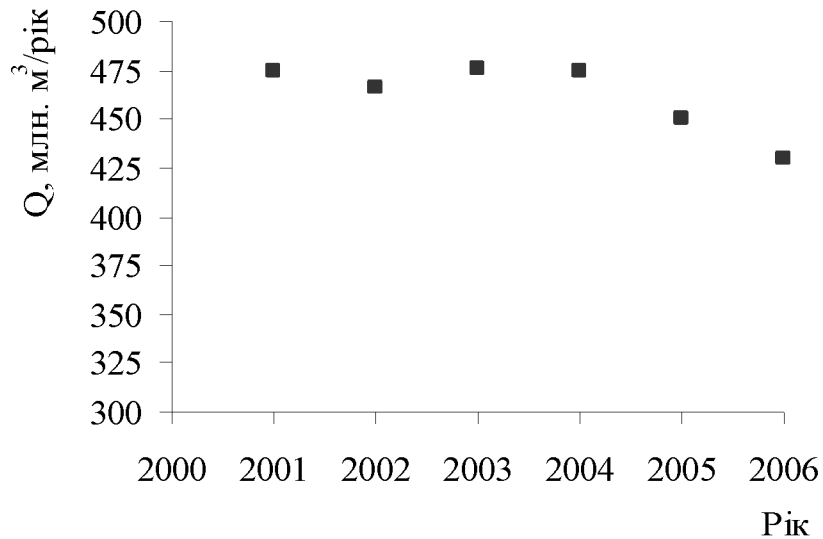


Рис. 3. Зміна потоку води (Q , млн. м³/рік) крізь Бортницькі очисні споруди впродовж 6 років

У той же час аналіз поведінки «найближчих сусідів» (дані про концентрацію завислих речовин та БСК₂₀ впродовж 2001–2006 рр. на рис. 2, *a* і *b*) дають підстави вважати, що з 2001 по 2003 рр. концентрація падала, а з 2003 по 2006 рр. – зростала. У той же час аналіз концентрації азоту амонійного та фосфатів (рис. 2, *в* і *г*) виявив тенденції падіння концентрації серед «найближчих сусідів» лише в 2001–2003 рр., в той час як протягом 2003–2006 рр. будь-яка тенденція не відслідковується. Відсутня кореляція між даними рис. 2 та даними рис. 3.

Висновки з даного дослідження і перспективи подальших розробок у даному напрямі. Застосування порівняльного аналізу даних «найближчих сусідів» не тільки дозволить уточнити характеристики стічної води, але й дасть змогу виявити справжні причини зміни показників та їхній взаємний вплив. З метою верифікації характеристик стічної води потрібно розглянути вихідні дані, на основі яких робилося щорічне осереднення, з метою виявлення поведінки «найближчих сусідів» і кореляції між різними показниками води.

1. *Guide to the Expression of Uncertainty in Measurement*: First edition. – ISO, Switzerland. – 1993. – 101 pp.