

## **ЕКОЛОГІЯ ТА ОХОРОНА НАВКОЛИШНЬОГО СЕРЕДОВИЩА**

**УДК 556.531 + 519.24.001**

**Артеменко В.А., Петрович В.В.,** канд. техн. наук

### **МЕТОД ВІДНОВЛЕННЯ ПРОПУСКІВ В ГІДРОХІМІЧНИХ ЧАСОВИХ РЯДАХ**

**Анотація.** Розроблено метод відновлення суцільних пропусків в гідрохімічних часових рядах.

Об'єкт дослідження: ряди середньодобових значень концентрацій нітратних та аммонійних йонів, *мг/л*.

Мета роботи: розробка методу відновлення надвеликих пропусків у часових рядах гідрохімічних даних по ряду витрат води.

Метод дозволяє відновлювати пропуски в рядах на термін один рік і більше.

**Ключові слова:** гідрохімічні та гідрологічні часові ряди, витрати води в річці, відновлення безперервних пропусків в гідрохімічних рядах за довгостроковий період.

**УДК 556.531+519.24.001**

**Артёменко В.А., Петрович В.В.,** канд. техн. наук

### **МЕТОД ВОССТАНОВЛЕНИЯ ПРОПУСКОВ В ГИДРОХИМИЧЕСКИХ ВРЕМЕННЫХ РЯДАХ**

**Аннотация.** Разработан метод восстановления сплошных пропусков в различных гидрохимических временных рядах.

Объект исследования: ряды среднесуточных значений концентраций нитратных и аммонийных йонов, *мг/л*.

Цель работы: разработка метода восстановления сверхбольших пропусков во временных рядах гидрохимических данных по ряду расходов воды.

Метод позволяет восстанавливать пропуски в соответствующих рядах на срок один год и более.

**Ключевые слова:** гидрохимические и гидрологические временные ряды, расход воды в реке, восстановление непрерывных пропусков в гидрохимических рядах за долгосрочный период.

**UDC 556.531+519.24.001**

**Artemenko V., Petrovych V.,** Cand. Eng. Sci. (Ph.D.)

## **METHOD FOR RECONSTRUCT UNCEASING GAPS IN HYDROCHEMICAL TIME SERIES**

**Abstract.** Designed method allows adequately reconstruct unceasing gaps in different hydrochemical time series for period one year and more.

Object of the study there are time series of water discharge and concentration of  $[NO_3^-]$  and  $[NH_4^+]$  ions.

Purpose of the study there is development of the method of the reconstruction very-large-scale gap given concentration time series by water discharge time series.

**Keywords:** hydrochemical time series, water discharge, unceasing gaps reconstruction.

*Требуются очень глубокие знания,  
чтобы заметить простейшие,  
но подлинные отношения вещей между собой.*

*Г.К. Лихтенберг*

### **Введение**

Процессы, происходящие в гидроэкологической системе, можно охарактеризовать при помощи различных временных рядов (гидрологических, климатических, гидробиологических, гидрохимических и других).

Одна из актуальных задач при этом является задача восстановления таких рядов в случае наличия неполных исходных данных.

Однако вопросы восстановления многолетних временных рядов различных гидроэкологических показателей, которые определяют основные характеристики водных объектов, остаются до настоящего времени мало изученными.

Численные эксперименты, проведенные авторами с использованием самых современных методов, основанных на различных подходах к прогнозированию,

показали, что большинство рядов с удовлетворительным качеством восстанавливаются всего на несколько точек.

При этом восстановить сверхбольшие пропуски в рядах обычно не удается.

Так, например, восстановить ряд среднесуточных значений концентраций химических веществ в воде по меньшей мере за один год (т. е. на 365 точек) или даже более, имея только данные по этим концентрациям за годы предыдущие, - задача в настоящее время практически нереальная.

Вместе с тем в [1] был приведен метод восстановления сверхбольших пропусков в рядах одного водного объекта по полным рядам другого водного объекта, где в качестве исходных данных фигурировали только гидрологические ряды.

Поскольку в любой экологической системе целый ряд параметров взаимосвязан [2], вполне возможно также восстановление одних временных рядов по качественно другим временным рядам.

Как известно, существует связь между величиной стока и концентрацией химических веществ в воде. Имея полный временной ряд среднесуточных расходов  $Q$  за  $N$  лет, а также ряд среднесуточных значений концентраций веществ в воде  $C$  за  $(N-j)$  лет, можно восстанавливать недостающие значения ряда концентраций.

### **Методика исследования**

В качестве примера рассмотрим метод восстановления сверхбольших пропусков во временных рядах нитратных и аммонийных йонов с помощью данных по расходам воды для одного из крупных водных объектов Украины.

Для выполнения расчетов воспользуемся базой данных по среднесуточным значениям расходов воды  $Q$ ,  $м^3/с$ , а также среднесуточным значениям концентраций нитратных  $[NO_3^-]$  и аммонийных  $[NH_4^+]$  йонов,  $мг/л$ , за период с 1991 по 2010 год включительно (без пропусков). Для упрощения вычислений значения данных за 29 февраля высокосных годов из рассматриваемых рядов исключались.

При проведении численных экспериментов из исходных данных искусственно удалялись среднесуточные значения по концентрациям веществ в реке за последний год наблюдений (с 01.01 по 31.12.2010 года).

В результате оперировали среднесуточными данными по расходам воды за 20-летний период и среднесуточными данными по концентрациям веществ в реке, но только за 19-летний период.

Располагая соответствующими данными для значений  $Q$ ,  $[NO_3^-]$  и  $[NH_4^+]$ , попытаемся восстановить данные по концентрациям  $C$  за весь 2010 год (365 значений).

Для этого первоначально сформируем траекторные матрицы для  $Q$ ,  $[NO_3^-]$  и  $[NH_4^+]$ , где каждая строка матрицы есть данные для одного года.

При этом, если строка  $I$  матрицы  $Q$  – это данные по расходам для конкретного  $N$ -го года, то и соответствующая строка матрицы  $C$  обязана быть данными для этого же года (см. рис. 1).

То есть обязательно должно выполняться соответствие по годам в строках матриц  $Q$  и  $C$ .

Далее ищем аналог для последней строки траекторной матрицы  $Q$  (т.е. для 2010 года) среди оставшихся 19-ти строк этой матрицы. Последовательно просматривая строки матрицы  $Q$ , находим наилучший аналог в этой матрице. Для поиска аналога можно использовать евклидову норму (метрику).

Аналог для расходов воды за 2010 год – это та строка траекторной матрицы  $Q$ , где норма разности строк (последней строки и текущей) принимает минимальное значение.

Заметим, что в работе использован термин «норма» как это понимают в математике, а точнее, в нотации языка программирования **MATLAB** (поскольку под «нормой», например, в гидрометеорологии понимают другое).

Определив индекс наилучшего аналога ( пусть, в данном случае,  $INDX=3$ , т.е. это третья сверху строка матрицы  $Q$ ), находим также год, соответствующий этому аналогу.

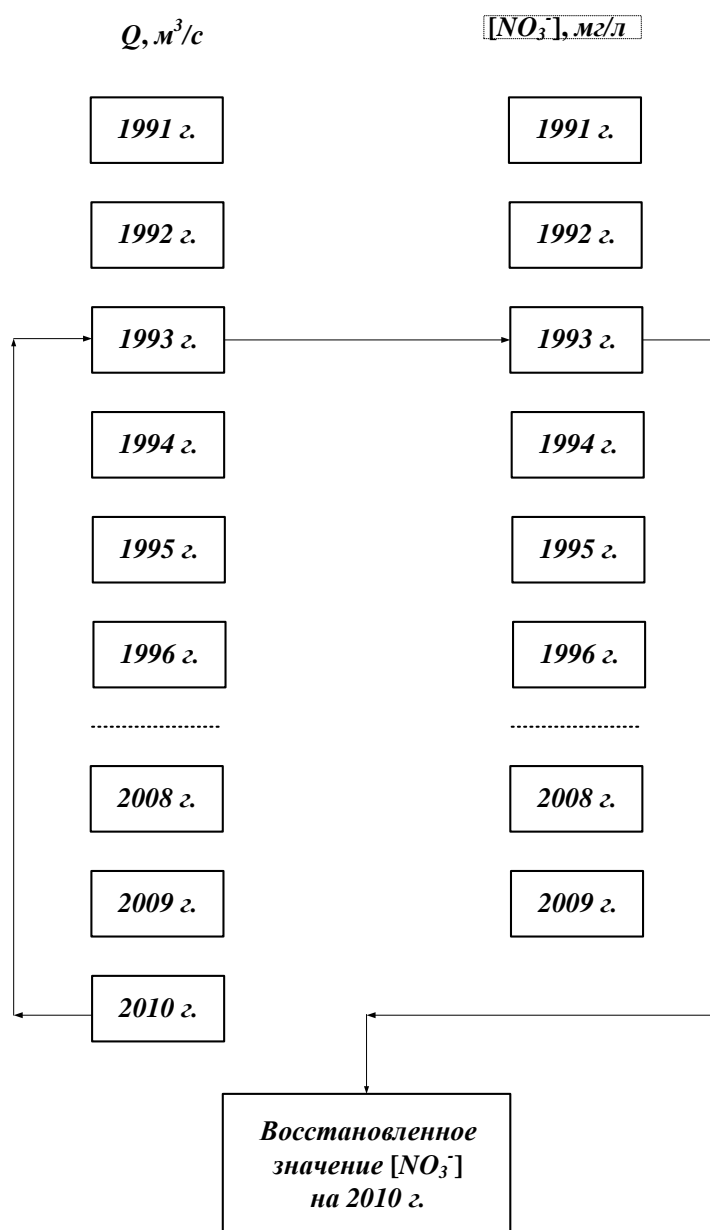
В нашем случае это 1993 год.

Далее будем полагать, что аналоги для  $Q$  порождают аналоги и для концентраций веществ  $C$ .

***Если найденному аналогу отвечает строка  $A$  матрицы  $Q$ , то соответствующая строка  $A$  матрицы  $C$  и будет искомой концентрацией.***

Другими словами, строка в матрице концентраций с индексом наилучшего аналога, найденного для матрицы  $Q$ , будет ответом данной задачи.

При этом можно искать не один, а несколько наилучших аналогов. В этом случае ответ получаем усреднением найденных аналогов в матрице  $C$  (обычным или взвешенным усреднением нескольких найденных значений в строках матрицы соответствующих концентраций).



**Рисунок 1** – Схема поиска вектора концентрации  $[NO_3^-]$  по наилучшему аналогу в матрице расходов

Заметим, что для адекватного восстановления концентраций по ряду расходов не требуется наличия значительных объемов данных. Как показали эксперименты, даже если база данных содержит значения по расходам всего за 5...7 лет, можно уже получать вполне приемлемые результаты.

Кроме того, данная модель восстановления гидрохимических показателей по временным рядам расходов позволяет в дальнейшем более правильно предсказывать наступление аномальных режимов в гидроэкологической системе.

Следует обратить внимание на следующее.

На первый взгляд может показаться, что разработанный метод восстановления пропусков в различных гидроэкологических рядах есть некоторым образом модифицированный метод локальной аппроксимации (метод *LA*), исследования по адаптации которого к природным рядам проводились авторами ранее [3].

Однако между методом *LA* и данным методом существует большая разница, причем на качественном уровне.

1. Метод *LA* оперирует только с данными одного вида. Если необходимо восстановить часть временного ряда, выполняем соответствующее восстановление ряда на 1; 2; 3; ... точки.

При этом другие ряды, характеризующие данный водный объект, принципиально не задействуются.

2. В методе *LA* четко прослеживается порог предсказуемости. В нашем методе понятие порога предсказуемости вообще не рассматривается.

Если, например, восстановление ряда  $[NO_3^-]$  методом *LA* возможно на 5...7 точек, то для данного метода вполне реально восстановление ряда на 365 и даже более точек!

3. Метод *LA* полагает, что подлежащий восстановлению ряд порождается динамической системой, где в полной мере возможно применение процедуры Такенса и других соответствующих методов анализа и прогноза. Однако, как показал наш опыт работы, временной ряд концентраций не является в полной мере динамической системой. Так, используя данную программу восстановления ряда  $Y$  по ряду  $X$  одной и той же динамической системы модельного хаотического ряда, мы не получаем результатов даже удовлетворительного качества.

В то же время приведенный метод долгосрочного восстановления различных природных временных рядов по качественно другим природным рядам в своём большинстве дает достаточно хорошие результаты.

Так, проведенные эксперименты по восстановлению ряда концентраций одних химических веществ по рядам концентраций других химических веществ в реке также дали результаты вполне удовлетворительного качества.

### **Особенности программной реализации метода**

Программа выполнена на языке программирования *MATLAB* и таким образом, что может быть также реализована в её бесплатных аналогах *OCTAVE*

и *FREEMAT*. При этом возможна работа не только в *OC WINDOWS*, но также в *OC LINUX*.

Поскольку авторами ранее разработано большое количество процедур для прогнозирования различных природных временных рядов на основе разновидностей метода локальной аппроксимации (метода *LA*), часть из них была использована и в данном случае.

Так, практически без модификаций была использована процедура нахождения аналогов.

Программа не обязательно требует последовательного расположения данных в строках матриц год за годом и в направлении увеличения времени сверху вниз (см. рис. 1), допуская также произвольное расположение данных по годам. Однако при этом требует для всех матриц одного и того же закона в расположении данных по годам.

Вследствии особенностей языка программирования *MATLAB* какие-либо специальные процедуры для загрузки исходных данных, сохранения результатов расчетов, вывода графиков и индексации массивов (матриц) не выполнялись.

Программа работает с различными природными рядами и показала устойчивость в работе при максимальной простоте в эксплуатации.

### **Оценка достоверности полученных результатов**

Вопрос о том, что зависимости вида  $[NO_3^-]=\Phi(Q)$  и  $[NH_4^+]=\Psi(Q)$  действительно реальны, выясняется исключительно при восстановлении одного ряда по другому ряду (в данном случае ряда концентраций по ряду расходов).

Будем считать, что такие зависимости существуют, если ряд реальный, данные по которому у нас также есть (см. выше), и ряд восстановленный имеют линейный коэффициент корреляции, равный 0,7 или более.

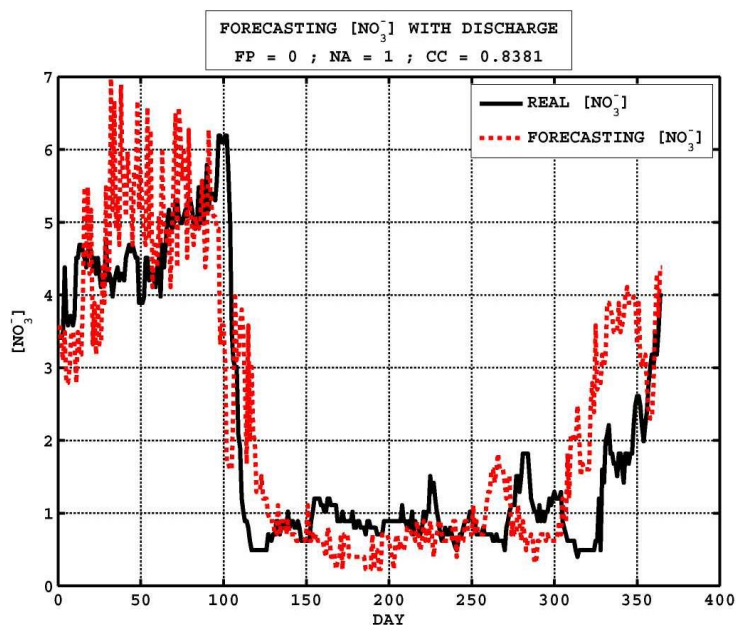
В результате проведенных расчетов оказалось, что коэффициент корреляции между реальными и восстановленными значениями для рядов  $[NO_3^-]$  оставался всегда существенно большим, чем значение 0,7.

Однако данный метод восстановления работает не для всех природных рядов. Так, восстановление временного ряда  $[NH_4^+]$  по ряду расходов  $Q$  выполнить не удалось (см. ниже).

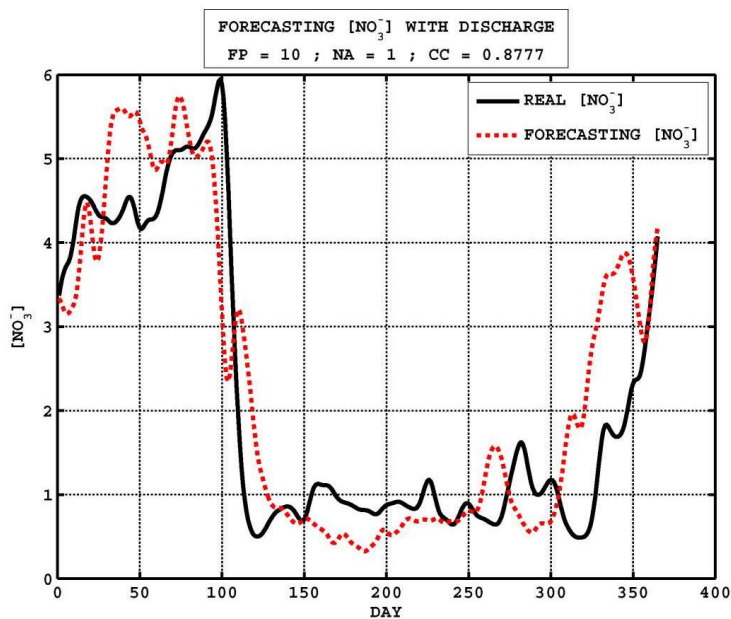
### **Результаты численных экспериментов и их обсуждение**

Было выполнено восстановление временного ряда концентраций ионов  $[NO_3^-]$  по временному ряду расходов  $Q$  на 2010 год (рис. 2).

На рис. 2 а приведенные как истинные, так и восстановленные данные для этого ряда.



а)



б)

а) фильтрация исходного ряда отсутствует (Filter Parameter = 0); б) при 10-кратной фильтрации исходного ряда

**Рисунок 2** – Результаты восстановления ряда  $[NO_3^-]$  по временному ряду расходов на 2010 год для случая одного аналога (NA=1)

Линейный коэффициент корреляции  $CC$  между реальными и восстановленными значениями получился достаточно высоким ( $CC=0,8381$ ), что заметно выше установленного порога 0,7.



Учитывая тот факт, что осцилляции необработанных рядов делают их практическое использование затруднительными, в дальнейшем проводилась также фильтрация этих рядов.

Фильтрация осуществлялась с помощью классического линейного трехточечного фильтра, работающего согласно соотношения

$$SS(J) = \frac{S(J-1) + S(J) + S(J+1)}{3},$$

где SS – профильтрованный ряд;

S – исходный временной ряд.

Поскольку цифровая фильтрация проводилась в разумных пределах, качественно общие тенденции ряда она не меняла.

Так, после проведенной фильтрации (FP=10) график стал более пригодным для практического использования (рис. 2б). Был получен еще более высокий коэффициент корреляции  $CC=0,8777$ .

Что касается результатов восстановления ряда  $[NH_4^+]$  по ряду значений  $Q$ , то восстановить этот ряд данным методом практически не получилось (линейный коэффициент корреляции  $CC=-0,1609$ ).

Фильтрация исходных данных (FP=10) только несколько улучшила визуальное восприятие графика ( $CC=-0,1774$ ).

Следует особо отметить, что восстановить значения ряда аммонийного азота даже на несколько точек другими известными методами также не удавалось.

### **Выводы**

1. Разработан метод восстановления временного ряда концентраций различных химических веществ по ряду расходов воды в реке. Выполнено в качестве примера восстановление ряда концентраций одного из важнейших биогенных ионов, - йона  $[NO_3^-]$  по ряду расходов воды.

Предложенным методом можно восстанавливать сплошные пробелы в исходных данных, превышающие по времени несколько лет (как последовательно, так и выборочно по годам).

Линейный коэффициент корреляции между реальными и восстановленными значениями ряда  $[NO_3^-]$  всегда оставался большим значения 0,7.

2. В большинстве случаев рассмотренным методом можно адекватно восстанавливать многие гидрохимические ряды. При этом следует отметить, что отдельные временные ряды концентраций химических веществ данным методом восстановить не удавалось (например, ряды концентраций  $[NH_4^+]$ ).

3. Проведенные предварительные численные эксперименты по восстановлению ряда концентраций одних химических веществ по временным рядам концентраций других химических веществ в реке также дали результаты вполне удовлетворительного качества.

### Литература

1. Артёменко В.А. Метод восстановления сверхбольших пропусков в гидрологических временных рядах/ В.А. Артёменко, В.В. Петрович// Автомобільні дороги і дорожнє будівництво, вип. 93. – К.: Вид-во Націон. трансп. ун-ту. – 2015. –С. 150-156.
2. Одум Ю.Экология. –Т.1. –М.: Мир, 1986. -328 с.; –Т.2. –М.: Мир, 1986. -376 с.
3. Артеменко В.А. Прогнозування нерегулярних часових рядів методом локальної аппроксимації/ В.А. Артеменко, В.В. Петрович// Автомобільні дороги і дорожнє будівництво, вип. 86. – К.: Вид-во Націон. трансп. ун-ту. – 2012. –С. 176-195.

#### Рецензенти:

Кузло М.Т., д-р техн. наук, Національний університет водного господарства та природокористування.

Хрутьба В.О., д-р техн. наук, Національний транспортний університет.

#### Reviewers:

Kuzlo M.T., Dr. Tech. Sci., National University of Water and Environmental Engineering.

Hrutba V.O., Dr. Tech. Sci., National Transport University.

Стаття надійшла до редакції: **20.05.2017 р.**