

УДК 556.048:625.745

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

**ПОВЫШЕНИЕ КАЧЕСТВА ПРОГНОЗИРОВАНИЯ
ГИДРОЛОГИЧЕСКИХ ВРЕМЕННЫХ РЯДОВ**

Анотація. Розглянуті особливості прогнозування коротких гідрологічних часових рядів методом локальної аппроксимації.

Об'єкт дослідження: добові ряди витрат води з різними інтервалами усереднення.

Мета роботи: підвищення якості прогнозування коротких гідрологічних часових рядів.

З метою кращої прогнозованості часових рядів методом **LA** запропонований простий спосіб їх попередньої **SVD**-фільтрації.

Аналізуються результати прогнозування попередньо фільтрованого гідрологічного ряду.

Ключові слова: гідрологічні часові ряди, метод локальної аппроксимації, **SVD**-фільтрація одномірних часових рядів, прогнозування фільтрованого гідрологічного ряду.

Аннотация. Рассмотрены особенности прогнозирования коротких гидрологических временных рядов методом локальной аппроксимации (методом **LA**).

Объект исследования : суточные ряды расходов воды с разными интервалами усреднения.

Цель работы: повышение качества прогнозирования коротких гидрологических временных рядов.

С целью лучшей прогнозируемости временных рядов методом локальной аппроксимации предложен простой способ их предварительной **SVD**-фильтрации.

Анализируются результаты прогнозирования предварительного фильтрованного гидрологического ряда.

Ключевые слова: гидрологические временные ряды, метод локальной аппроксимации, **SVD** – фильтрация одномерных временных рядов, прогнозирование фильтрованного гидрологического ряда.

Annotation. In article are considered particularities of the forecasting short hydrological time series by Local Approximation Method (**LA**).

Object of the study: it is streamflows with different intervals of the averaging.

Purpose of the investigation: increasing forecast quality of short hydrological time series. Preliminary **SVD**-filtering was used for improvement forecasting quality.

Was learning the action to preliminary filtering on quality of the forecasting and engineering method was designed.

Keywords: hydrological time series, the Local Approximation forecasting method, Singular Value Decomposition, **SVD**-filtration method for one-dimension time series, forecasting of filtering hydrological time series (streamflows).

Введение

Как известно, речной сток является одной из важнейших базовых характеристик при проектировании мостовых переходов и различных гидротехнических сооружений.

В настоящее время гидрологическое прогнозирование базируется на метеорологической информации.

При этом надежность такого прогноза во многом зависит от качества метеорологических данных и возможностей их доступа к архивным и оперативным материалам.

Как показывает практика применения программных комплексов на основе использования метеорологических характеристик, достоверные

результаты при прогнозировании получаются только на первые (максимум вторые) сутки прогноза.

Новые возможности открывают современная математическая теория нелинейной динамики, которая предлагает различные приёмы и методы прогнозирования. При этом не требуется знания аналитической зависимостей, описывающих поведение объекта во времени, а исследуется только его конкретный временной ряд.

С точки зрения нелинейной динамики в таком временном ряде уже содержится и закон эволюции ряда, и начальные условия, т.е. все достаточные для его прогнозирования.

Также это практически единственный способ, который можно использовать в ситуации, когда никакой другой информации о поведении объекта кроме значений временного ряда, нет.

Одним из наиболее эффективных и простых в реализации методов прогнозирования временных рядов является метод локальной аппроксимации (**LA**-метод).

В [1] были рассмотрены особенности прогнозирования временных рядов **LA**-методом нулевого порядка.

Одно из преимуществ прогнозирования этим методом по сравнению с другими методами состоит в том, что предварительная обработка исходных временных рядов в принципе не требуется.

Однако такое преимущество прежде всего относится к случаям, когда временные ряды имеют достаточную для реализации метода длину (порядка 10000 точек ряда и более).

На практике обычно приходится работать со значительно более короткими рядами.

В этой связи представляет особый интерес оценка возможностей **LA**-метода для прогнозирования коротких, предварительно не обработанных гидрологических рядов.

Для повышения качества прогноза таких рядов предлагается также дополнительно использовать предварительную «неглубокую» **SVD** – фильтрацию исходного временного ряда.

Разработки осуществляются в рамках специализированного программного комплекса анализа и прогнозирования гидрологических временных рядов на основе методов нелинейной динамики и детерминированного хаоса.

В исследовании использовалась векторно-матричная нотация, принятая в системе **MATLAB** и её аналогах.

Прогнозирование коротких гидрологических рядов методом локальной аппроксимации

Рассмотрим процесс прогнозирования расхода воды, $\text{м}^3/\text{с}$, на 1 месяц (30 суток) вперёд методом локальной аппроксимации.

В качестве исходных данных был принят временной гидрологический ряд среднесуточных расходов по данным одного из стоковых постов р. Десна.

Использовались параметры прогнозирования согласно [1]:

DIM - размерность реконструированного фазового пространства, равная числу столбцов траекторной матрицы:

NNV – число аналогов, находимых для текущего вектора состояния;

N – общее число точек, на которое прогнозируется данный ряд;

NFS – число точек (шагов), на которое осуществляется прогноз за одно обращение(один такт) к процедуре **LA** - прогнозирования;

LS – общая длина ряда, используемая для прогноза, точек.

При этом на вход процедуры прогнозирования поступала только часть исходного ряда, определяемая разностью **LS-N**.

Временная задержка **TAY**, необходимая при формировании траекторной матрицы, была принятой равной единице [1].

Как показали расчеты, для прогнозирования на 30 суток вперед необходимо сразу получить 30 прогнозных значений (точек), что оказалось в данном случае невыполнимым.

Поэтому далее использовался ряд расходов, усредненный по 5 суткам (одна точка временного ряда на графике соответствует среднему за 5 суток). То есть такие данные подразумевают дискретность в 5 суток.

Обозначим такой ряд как **DESNA-005**.

Применяя одношаговое прогнозирование, например, 300 раз, можно получать прогнозные значения на протяжении 1500 суток на основе значений реального ряда.

Результаты одношагового прогнозирования ряда **DESNA-005** при длине исходного ряда **LS=7920** значений (точек) и параметрах прогнозирования **NFS=1**, **DIM=1**, **NNV=300** и **N=300** приведены на рис.1.

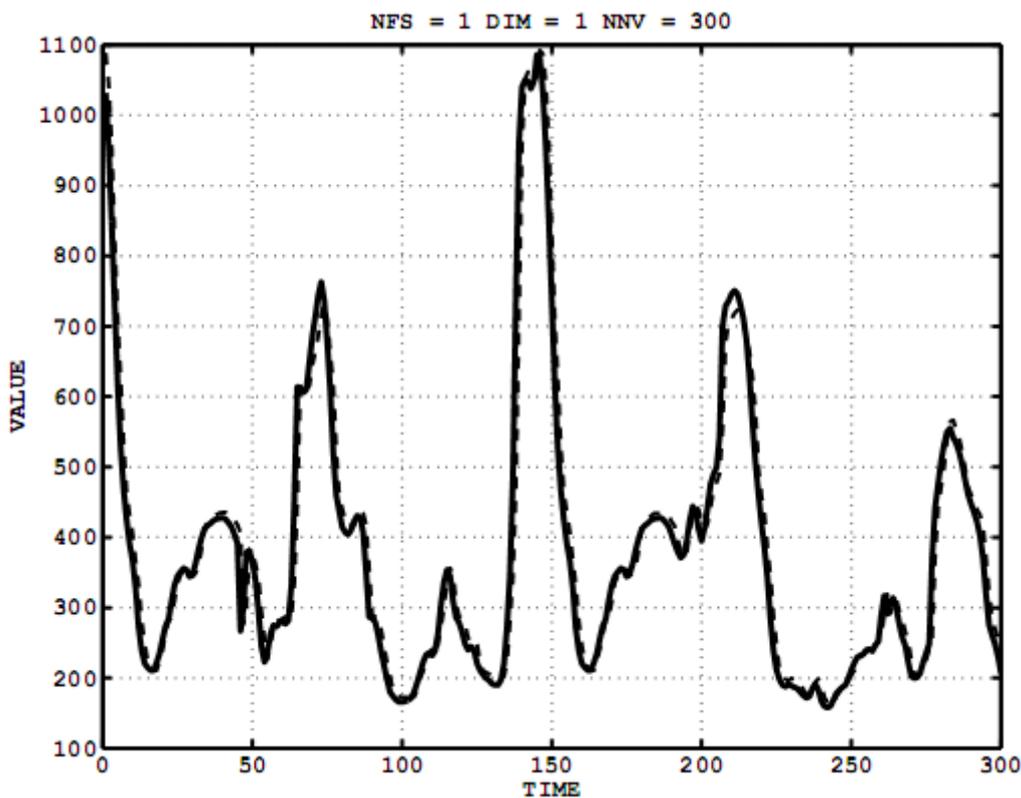


Рисунок 1

В данном случае, как видно из рис.1, наблюдается почти полное совпадение значений исходного ряда (сплошная линия) и рассчитанных прогнозных значений (штриховая линия) даже для таких рядов с фрактальным видом кривой.

Мерой качества прогноза предварительно выбран коэффициент корреляции **C** между соответствующими значениями участка реального (исходного) ряда и прогностического ряда (**N**).

Применение коэффициента корреляции является весьма удобным, поскольку в данном случае достигается прогноз достаточно высокого качества

($C=0,9782$), а использование мер, связанных с абсолютными, относительными и другими ошибками, оказалось менее информативным.

Приведенный выше график четко свидетельствует о детерминируемости гидрологического ряда, поскольку для «случайного» ряда, сгенерированного с помощью «Генератора случайных чисел», не удалось получить адекватного прогноза даже на одну точку вперед.

В этой связи можно говорить о том, что метод локальной аппроксимации также является методом анализа временных гидрологических рядов на детерминируемость, т.е. с помощью данного метода можно легко отличить детерминируемый ряд от ряда «случайного».

Однако для решения задачи прогнозирования ряда на 1 месяц (30 суток) вперед в данном случае необходимо за один такт процедуры прогнозирования получить 6 прогнозных значений (точек).

Расчеты показали, что и в этом случае качество прогноза оказывается неудовлетворительным.

Поэтому далее использовались ряды с большими интервалами усреднения (ряды с усреднением по 10 суткам).

Как оказалось, адекватно прогнозировать величину расхода для ряда **DESNA-010** (см.рис.2) также возможно только на одну точку (10 суток) вперед, но с достаточно большой точностью ($C=0,9405$ при значениях **LS**= 3960 точек, **NFS**=1, **DIM**=1, **NNV**=50 и **N**=160).

Для прогноза участка ряда на 30 суток вперед необходимо за один такт процедуры осуществить прогнозирование сразу на 3 точки вперед.

Однако качество прогноза оказалось при этом также невысоким.

На рис.3 представлены результаты одношагового прогнозирования для ряда **DE**

Полученное значение коэффициента корреляции составило $C=0,8418$.

Заметим, что только изменением одного из параметров прогнозирования (**NNV**) удалось повысить значение коэффициента корреляции до величины $C=0,8625$ (использован параметр **NNV**=300).

Получение прогноза сразу на 30 суток вперёд возможно в данном случае при значении $NFS=2$ (двухшаговом прогнозировании, т.е. прогнозировании на 2 точки за 1 такт процедуры).

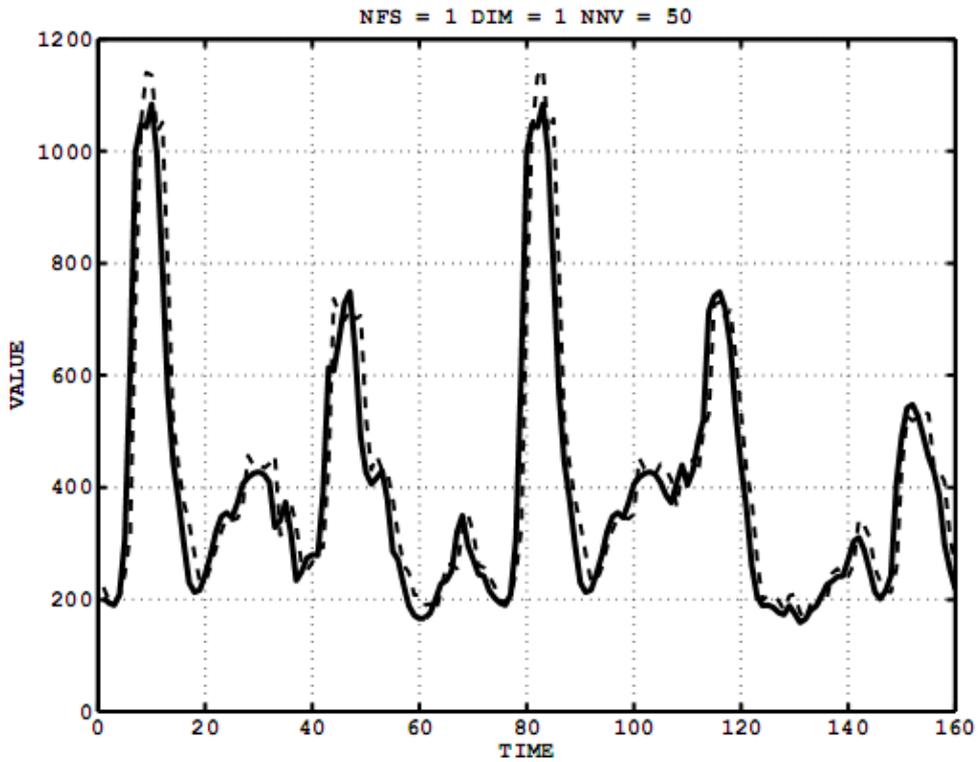


Рисунок 2

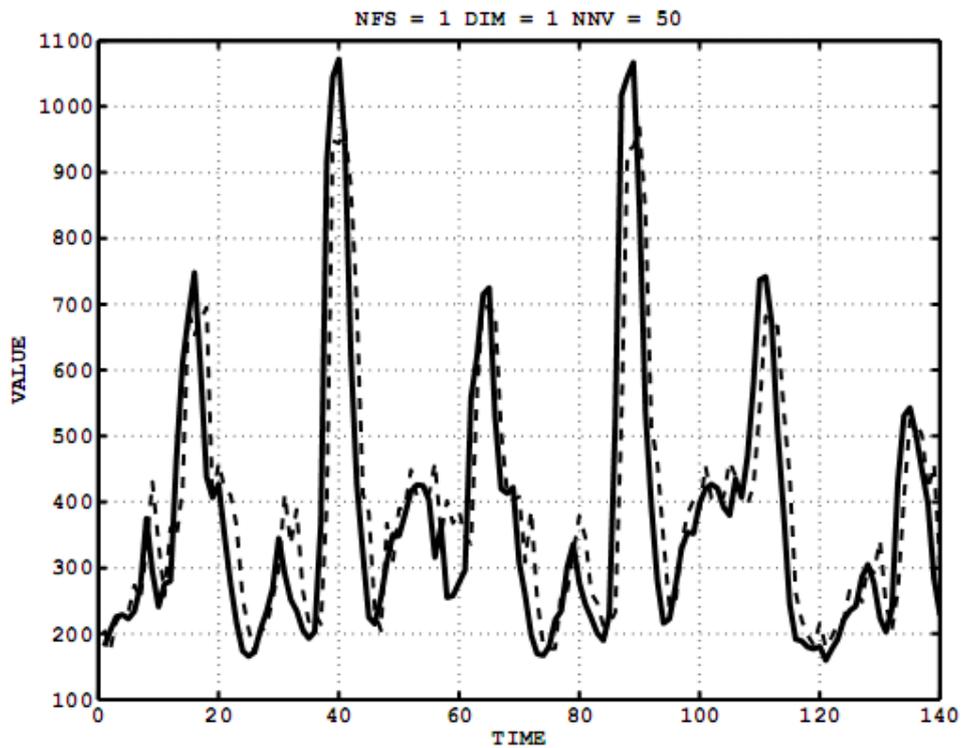


Рисунок 3

На рис.4 представлен один из лучших полученных результатов прогноза расходов воды на 30суток для участка ряда **DESNA-015** (**LS=2640**, **NFS=2**, **DIM=12**, **NNV=12** и **N=140**).

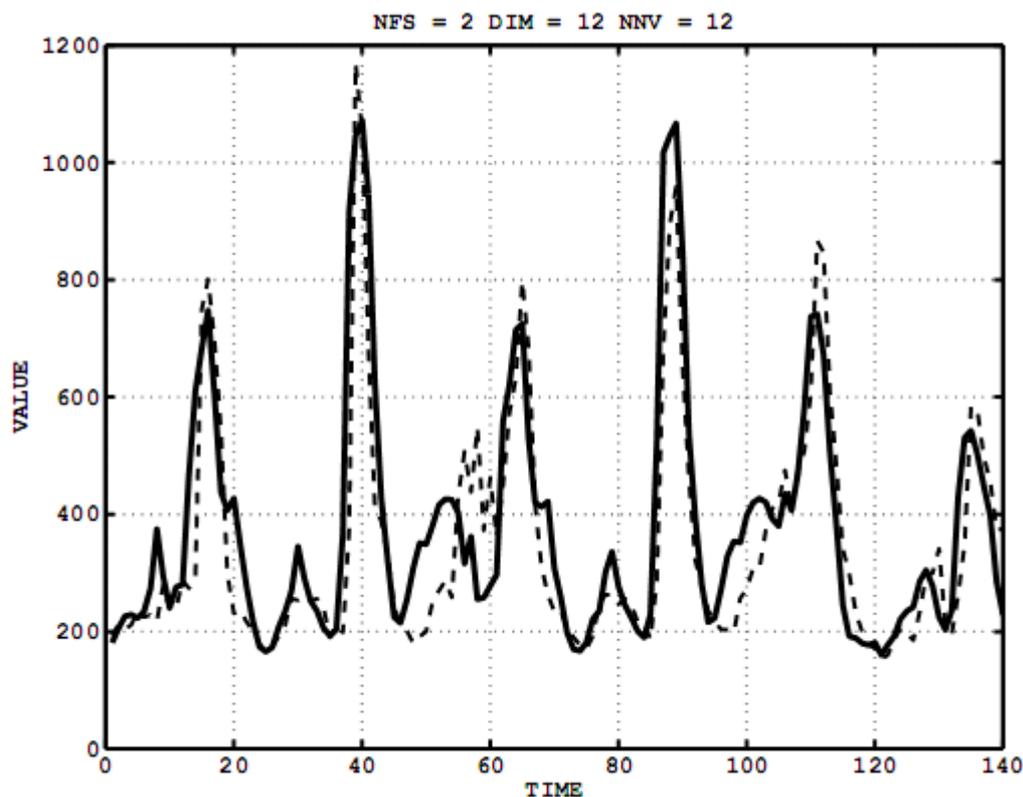


Рисунок 4

Коэффициент корреляции составил $C=0,8602$.

Таким образом, получен вполне удовлетворительный результат прогнозирования.

И, наконец, на рис.5 представлен результат одношагового прогнозирования ряда **DESNA-030** (ряда, усредненного по 30 суткам) при значениях параметров соответственно **LS=1320**, **NFS=1**, **DIM=4**, **NNV=48** и **N=120**.

SNA-015 при значениях **LS= 2640**, **NFS=1**, **DIM=1**, **NNV=50** и **N=140**.

Значение коэффициента корреляции составило $C=0,8028$.

Заметим, что для практики представляет интерес даже такой прогноз с усреднением по 30 суткам, поскольку позволяет предварительно оценить величину расхода.

В итоге имеем два альтернативных варианта решения задачи прогнозирования расходов воды на 1 месяц (30 суток) вперёд, - ряд **DESNA-015** при значении **NFS=2** и ряд **DESNA-030** при значении **NFS=1**.

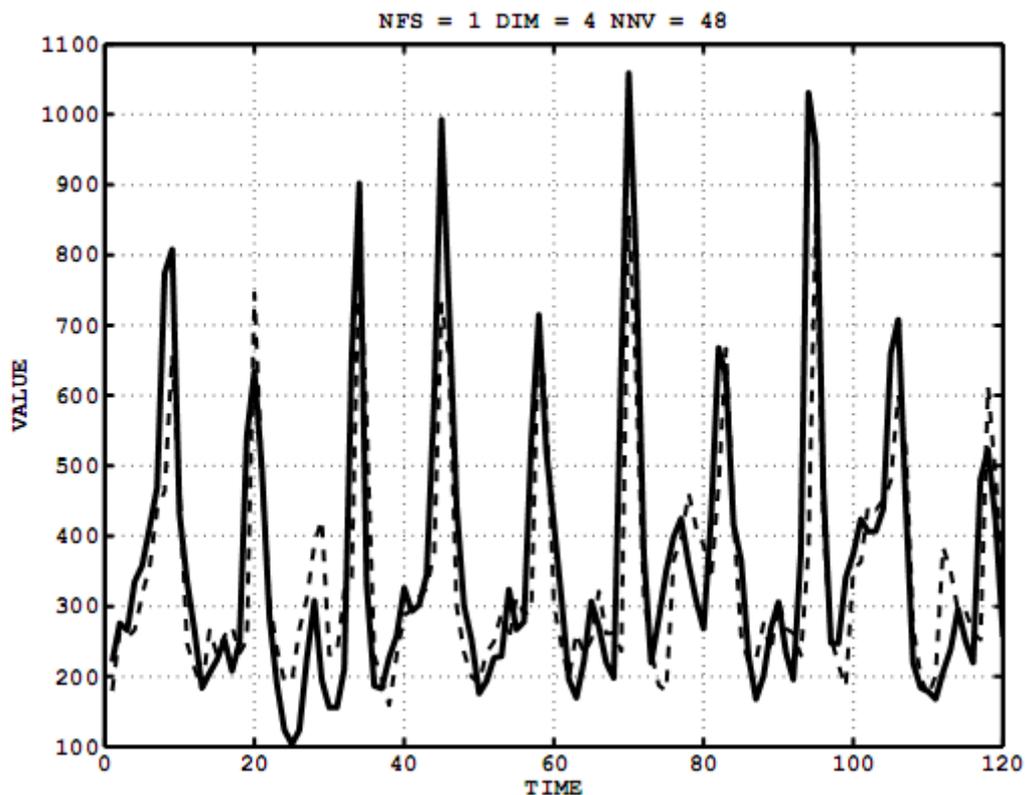


Рисунок 5

Сравнение полученных результатов показало, что лучшие показатели достигаются для ряда, где проводится усреднение по 15 суткам и используется двухшаговое прогнозирование.

Как видно, по мере возрастания интервала усреднения диапазон значений ряда уменьшается. Однако с некоторой погрешностью при необходимости можно вернуться к диапазону значений исходного ряда путем введения специальной пересчетной процедуры (в статье не обсуждается).

После выбора усредненного ряда, который подходит для прогнозирования наилучшим образом, а также тактики прогнозирования, остается лишь подобрать оптимальные параметры для достижения прогноза наиболее высокого качества.

Таким образом, даже для сравнительно коротких гидрологических рядов получены с точки зрения практики прогнозирования достаточно приемлемые результаты.

Использование предварительной SVD-фильтрации исходного ряда для повышения качества прогноза

Обычно процедура **LA**- прогнозирования сама выявляет закономерности в поведении ряда (т.е. осуществляется поиск, по которому можно определить, - каким образом последующие составляющие ряда зависят от предыдущих), а далее уже эти закономерности используются непосредственно для целей прогнозирования [1].

Однако часто данная процедура не может в полной мере выявить эти закономерности, что связано прежде всего со сложностью законов, которые формируют природные нерегулярные ряды.

В этом случае полезно выполнить предварительную обработку временных рядов, исключив из них незначительные детали и оставив только детали главные.

В работе [2] анализировался ряд в виде матрицы, где сама матрица интерпретировалась как своего рода «изображение».

Интуитивно ясно, что наличие многих мелких, незначительных деталей затрудняет распознавание образов на таких «изображениях», особенно при использовании достаточно простых процедур распознавания. Если устранить эти незначительные детали, можно существенно улучшить работу процедур распознавания.

Устранение незначительных деталей - это, по сути, фильтрация (в нашем случае фильтрация исходных данных).

Какие именно детали следует считать незначительными и устранять из «изображения» для облегчения процесса распознавания, зависит от особенностей самого «изображения».

В этом случае можно говорить о фильтрации, которая управляется характером самих фильтруемых данных, т.е. об интеллектуальной фильтрации.

В настоящее время большинство методов интеллектуальной фильтрации касаются только матричных данных («изображений»).

Один из наиболее распространенных методов интеллектуальной фильтрации временных рядов базируется на **SVD** - разложении (Singular Value Decomposition).

Однако выполнение интеллектуальной **SVD**-фильтрации одномерных временных рядов имеет свои особенности, которые в данном случае будут сводиться к следующему.

Пусть имеется исходный одномерный (векторный) ряд

$$S = [S(1); S(2); S(3); \dots \dots S(8)], \quad (1)$$

где $S(1); S(2); S(3); \dots \dots S(8)$ -составляющие ряда (числа).

Составим из ряда S , например, траекторную матрицу \mathbf{M} согласно [1] с $\mathbf{DIM}=5$ вида

$$\mathbf{M} = \begin{bmatrix} S(1) & S(2) & S(3) & S(4) & S(5) \\ S(2) & S(3) & S(4) & S(5) & S(6) \\ S(3) & S(4) & S(5) & S(6) & S(7) \\ S(4) & S(5) & S(6) & S(7) & S(8) \end{bmatrix} \quad (2)$$

выполнив таким образом переход от одномерного ряда (1) к ряду многомерному, т.е. матрице (2).

Такую матрицу уже можно трактовать как своего рода «изображение» и, соответственно, применять к ней методы интеллектуальной фильтрации.

По идейному смыслу такая фильтрация очень близка к методу главных компонент [3].

На практике на вход процедуры **SVD** подается не сама матрица \mathbf{M} , а матрица $\bar{\mathbf{M}}$ (центрированная матрица), полученная поэлементным вычитанием среднего значения по всей исходной матрице \mathbf{M} (среднее значение находилось по отдельной процедуре).

Это связано прежде всего с особенностью работы самой процедуры **SVD**

При сингулярном разложении из центрированной матрицы получаем соответственно матрицы \mathbf{U} , \mathbf{S} и \mathbf{V} , что выглядит как

$$\bar{\mathbf{M}} = \mathbf{U} * \mathbf{S} * \mathbf{V}^T, \quad (3)$$

где \mathbf{T} -процедура транспонирования матрицы.

Понятно, что в этом случае фильтрация отсутствует.

Но если при восстановлении (композиции) матрицы \bar{M} использовать только несколько первых столбцов матриц U и V , например,

$$\begin{aligned}M_1 &= U(:,1) * S(1,1) * V(:,1)^T \\M_2 &= U(:,2) * S(2,2) * V(:,2)^T \\M_3 &= U(:,3) * S(3,3) * V(:,3)^T,\end{aligned}\tag{4}$$

а затем сложить (поэлементно) эти матрицы, то в результате будем иметь матрицу MM , равную

$$MM=M_1+M_2+M_3.\tag{5}$$

Добавив к матрице MM среднее значение, которое мы ранее предварительно вычитали из исходной матрицы M , получим в результате фильтрованную матрицу $MM_{(F)}$.

Очевидно, что матрицы M и $MM_{(F)}$ будут одного и того же размера.

В данном примере траекторная матрица (2) имеет 5 столбцов. Однако при восстановлении использовали сумму только из трех первых составляющих (можно, по-видимому, применить далее термин «гармоники», если не забывать при этом о смысле выполняемых преобразований). В этом случае обозначение фильтра для данного примера будет «**SVD-5-3**», т.е. у сформированной из ряда (1) траекторной матрицы 5 столбцов (максимальное число «гармоник», которым мы располагаем, равно 5), из которых взято только 3 «первых гармоники».

Далее остается только снова превратить матрицу $MM_{(F)}$ в одномерный ряд (произвести обратное преобразование по отношению к тому, что выполнялось нами ранее).

Пример расчета

В качестве оценки эффективности использования предварительной **SVD**-фильтрации рассмотрим результаты прогнозирования достаточно короткого гидрологического ряда расходов воды, м³/с, усреднённого по 15 суткам (см. выше).

Исходя из того, что прогнозирование гидрологических временных рядов рассматривались нами ранее из детерминистических позиций, логично предварительную **SVD**-фильтрацию исходного ряда также рассматривать с тех же позиций.

Параметры используемого в примере **SVD** - фильтра позволили выделить 24 гармоники, из которых в дальнейшем оставляли только 12 «первых гармоник» (соответствующих первым наибольшим 12-ти сингулярным числам).

После проведенной **SVD**-фильтрации ряда DESNA-015 получили соответственно ряд

$$\mathbf{SVD} -24-12-\mathbf{DESNA}-015. \quad (6)$$

При указанных параметрах **SVD**-фильтра форма исходного ряда после фильтрации изменилась незначительно вследствие небольшой «глубины» фильтрации (понятно, что чем более короткий ряд, тем больше следует задавать «глубину» фильтрации).

Поскольку методом **LA** прогнозировался не фактический исходный ряд, а ряд фильтрованный, результаты прогнозирования сравнивались соответственно с фильтрованным рядом.

При этом параметры прогнозирования составляли соответственно :

DIM=2, **NNV=100**, **N=140** и **LS=2640**.

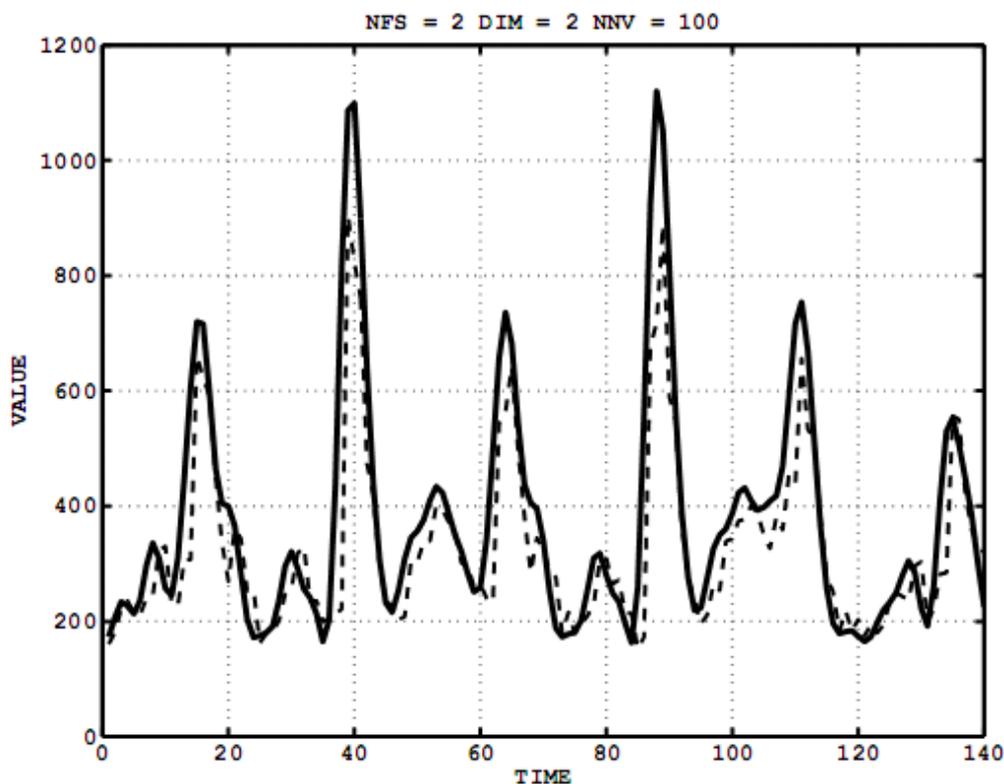


Рисунок 6

На рис.6 представлен результат двухшагового **LA** - прогнозирования (**NFS=2**) участка фильтрованного ряда (6), м³/с.

Значение коэффициента корреляции в данном случае составило $C = 0,8912$ (без предварительной фильтрации ряда $C=0,8602$).

Приведенный подход позволит принимать более обоснованные инженерные решения, направленные на повышение качества прогнозирования гидрологических временных рядов.

Выводы

1. Рассмотрены особенности прогнозирования коротких гидрологических временных рядов методом локальной аппроксимации (**LA**-методом).

Представлены результаты прогнозирования гидрологических рядов с различными интервалами усреднения и выбран оптимальный вариант прогноза на один месяц вперед.

2. Как показали выполненные расчеты, метод **LA** является также методом анализа гидрологических временных рядов на детерминируемость, т.е. с помощью этого метода можно легко отличить детерминируемый ряд от ряда "случайного".

3. Предложен простой способ предварительной **SVD**-фильтрации нерегулярных временных рядов для повышения качества их прогноза (учитывается, что "неглубокая" **SVD**-фильтрация не отражается глобально на форме исходного временного ряда).

4. Приведены результаты **LA**-прогнозирования короткого гидрологического ряда с использованием процедуры предварительной **SVD**-фильтрации исходного ряда

Литература

1. Артеменко В.А., Петрович В.В. Прогнозування нерегулярних часових рядів методом локальної аппроксимації // Автомобільні дороги і дорожнє будівництво. - Вип. 86. - К.: Вид-во НТУ. - 2012. - С. 176-195.

2. Петрович В.В., Артеменко В.А. Прогнозування часових природних рядів на основі методу **SVD**-розкладення // Автомобільні дороги і дорожнє будівництво. - Вип. 84. - К.: Вид-во НТУ. - 2012. - С. 13-29.

3. Коханер Д., Моулер К., Неш С. Численные методы и программное обеспечение. - М.: Мир, 2001. – 575 с.