

УДК 519.24:001

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

## SVD-ЭНТРОПИЙНЫЙ МЕТОД АНАЛИЗА ПРИРОДНЫХ ВРЕМЕННЫХ РЯДОВ

**Анотація.** Наведені основні положення методу *SVD*-ентропії, що базується на сингулярному розкладенні матриць.

Розглянуті питання кількісної оцінки *SVD*-ентропії та характер зміни її поведінки з часом відповідно до якості гідрологічних прогнозів.

**Ключові слова:** часові природні ряди, *SVD*-ентропія, ентропійний метод аналізу гідрологічних рядів, якість гідрологічного прогнозу.

**Аннотация.** Приведены основные положения метода *SVD*-энтропии, который базируется на сингулярном разложении матриц.

Рассмотрены вопросы количественной оценки *SVD*-энтропии и характер изменения ее поведения во времени применительно к качеству гидрологических прогнозов.

**Ключевые слова:** временные природные ряды, *SVD*-энтропия, энтропийный метод анализа гидрологических рядов, качество гидрологического прогноза.

**Annotation.** In given article is in detail considered method of SVD entropy.

It is considered practical application of the method of the SVD entropy to analysis hydrological time series.

It is considered the inter-relations between entropy of hydrological time series and quality of hydrological prediction.

**Keywords:** natural time series, *SVD*-entropy, quality of hydrological prediction.

### Введение

В технике анализа временных рядов важное место занимают методы, основанные на расчетах энтропий. Энтропии являются достаточно

информативными параметрами, количественно характеризующими состояние различных систем.

Энтропии характеризуют либо сложность поведения (отдельных процессов), либо сложность строения, либо другие сложности в изучаемой системе [1,2].

В упрощенной трактовке энтропия указывает на степень непредсказуемости (хаотичности) в поведении анализируемой системы: чем больше величина энтропии, тем более непредсказуемое поведение система демонстрирует, тем труднее ее прогнозировать.

Заметим, что данное положение относится именно к природным временным рядам.

В статье рассматриваются основные принципы *SVD*-энтропии временных рядов и анализируются тенденции в ее поведении применительно к гидрологическим рядам.

### **Метод анализа**

В настоящее время предложено большое число различных энтропий. При этом в итоговых расчетах применяется в основном классическая формула для информационной энтропии К.Шеннона [3].

В конце 90-х годов прошлого века был введен в практику вид энтропии, основанный на использовании сингулярного разложения матриц (*SVD*).

Для расчета данной энтропии необходимо знать только сингулярные числа для *SVD*-разложения исходной матрицы.

Однако в нашем случае имеется обычный временной ряд, то есть вектор, поэтому исходный ряд должен быть вначале преобразован в матрицу.

Такая матрица в хаотической динамике носит название траекторной матрицы, или реконструированного аттрактора.

В настоящее время существует целый ряд методов, позволяющих по одномерному ряду восстановить (реконструировать) аттрактор.

Нами в этой связи был использован наиболее простой метод реконструкции аттрактора – метод Такенса временных задержек [4].

Как известно, этот метод требует знания двух параметров: величины временной задержки (*TAY*) и размерности пространства реконструированного аттрактора (*DIMENSION*, или сокращенно *DIM*).

Обычно при реконструкции аттрактора используются конкретные величины *TAY*, которые определяются при помощи специальных методов. В

данном случае значение  $TAY$  принудительно устанавливалось равным единице ( $TAY=1$ ).

Это означает, что каждая точка исходного одномерного ряда (вектора) будет обязательно присутствовать в реконструированном аттракторе (матрице реконструированного аттрактора).

Другими словами, любая информация при формировании такой матрицы (аттрактора) не будет утеряна.

Что касается параметра  $DIM$  (числа столбцов полученной траекторной матрицы), то для большинства исследуемых нами природных рядов эта величина может быть принята равной  $DIM=14$ .

В настоящий момент в хаотической динамике хорошо изученными остаются только маломерные системы.

Тем не менее, большинство природных рядов – это именно ряды с большой размерностью фазового пространства.

Понятно, что для таких рядов мы не можем гарантировать в полной мере низкую размерность соответствующих динамических систем.

В этой связи величина  $DIM$ , возможно, должна была бы иметь значительно большее значение, чем  $DIM=14$ . Однако даже при  $DIM=182$  мы получили не очень большое отличие в значениях  $SVD$ -энтропии для всех анализируемых нами рядов.

При этом качественная картина процесса оставалась практически такой же.

Выбрав соответствующее значение  $DIM=14$ , далее будем придерживаться этого значения во всех проводимых нами численных экспериментах.

Как известно, количество значимых (наиболее больших по величине) сингулярных чисел прямым образом зависит от количества динамических компонент системы.

То есть по энтропии, определяемой на основании значений сингулярных чисел, можно характеризовать (измерять) сложность исследуемого сигнала.

Однако в начале сингулярные значения должны быть нормированы путем деления каждого сингулярного значения на общую сумму всех сингулярных значений:

$$\bar{\sigma}(i) = \frac{\sigma(i)}{\sum_{i=1}^n \sigma(i)}, \quad (1)$$

где  $\bar{\sigma}(j)$  – нормализованное  $j$ - сингулярное значение;

$\sigma(j)$  – исходное  $j$ - сингулярное значение;

$N$  – общее число сингулярных значений (сингулярных чисел).

Далее рассчитывается собственно энтропия:

$$H_{SVD} = -\sum_{j=1}^{j=N} (\bar{\sigma}(j) * \text{LOG}(\bar{\sigma}(j))). \quad (2)$$

Как видно, при вычислении  $SVD$ -энтропии использовался натуральный логарифм (логарифм по основанию  $EXP(1) \approx 2,7183$ ).

Обычно в литературе выражение (2) считают полученным значением  $SVD$ -энтропии (т. е. вещественным числом).

В нашей работе мы использовали также нормировку энтропии, заданной выражением (2).

После такой нормировки значение энтропии будет находиться в диапазоне значений  $[0; 1]$ , что существенно облегчает интерпретацию полученных результатов. Сама нормировка выполняется путем деления (2) на  $LOG(N)$ .

Пакет процедур для вычисления  $SVD$ -энтропии природных временных рядов был разработан и реализован на языке программирования *MATLAB*.

Код без существенных модификаций работает также в таких бесплатных аналогах *MATLAB* как *OCTAVE* и *FREEMAT* (как в *WINDOWS*, так и *LINUX*).

Программная реализация следует расчетным формулам (1) и (2), однако имеет и ряд алгоритмических тонкостей, которые хорошо прослеживаются в приведенном ниже программном коде.

```
function [ E ] = SVD_ENTROPY( INPUT_SERIES , DIMENSION , TAY );  
E = [ ];  
if IS_SERIES( INPUT_SERIES ) == false  
return  
end  
if IS_POSITIVE_INTEGER_NUMBER( DIMENSION ) == false  
return  
end
```

```

if IS_POSITIVE_INTEGER_NUMBER( TAY ) == false
return
end
if DIMENSION < 2
return
end
INPUT_SERIES = INPUT_SERIES( : );
ATTRACTOR = ATTRACTOR_RECONSTRUCTION( INPUT_SERIES ,
DIMENSION , TAY );
if isempty( ATTRACTOR ) == true
return
end
SZ1 = size( ATTRACTOR , 1 );
SZ2 = size( ATTRACTOR , 2 );
if ( SZ1 < 2 ) | ( SZ2 < 2 )
return
end
ATTRACTOR = ATTRACTOR - mean( ATTRACTOR( : ) );
S = svd( ATTRACTOR );
L = length( S );
S( S <= eps ) = eps;
SS = S ./ sum( S );
E = - sum( SS .* log( SS ) ) ./ log( L );
if std( ATTRACTOR( : ) ) == 0
E = 0;
end
end
function [ ATTRACTOR ] = ATTRACTOR_RECONSTRUCTION(
INPUT_SERIES , DIMENSION , TAY ;
ATTRACTOR = [ ];
if IS_SERIES( INPUT_SERIES ) == false
return
end
if IS_POSITIVE_INTEGER_NUMBER( DIMENSION ) == false
return

```

```

end
if IS_POSITIVE_INTEGER_NUMBER( TAY ) == false
return
end
INPUT_SERIES = INPUT_SERIES( : );
SZ = length( INPUT_SERIES );
1
SZ1 = SZ - ( DIMENSION - 1 ) .* TAY;
if SZ1 < 1
return
end
SZ2 = DIMENSION;
ATTRACTOR = zeros( SZ1 , SZ2 );
INDX1 = 0;
INDX2 = 0;
for J = 1 : SZ2
INDX1 = 1 + ( J - 1 ) .* TAY;
INDX2 = SZ - ( SZ2 - J ) .* TAY;
ATTRACTOR( : , J ) = INPUT_SERIES( INDX1 : INDX2 );
end
end
function [ ANS ] = IS_POSITIVE_INTEGER_NUMBER( N );
ANS = false;
if isempty( N ) == true
return
end
if isnumeric( N ) == false
return
end
if isscalar( N ) == false
return
end
if isnan( N ) == true
return
end

```

```

if isinf( N ) == true
return
end
if isreal( N ) == false
return
end
if N ~= fix( N )
return
end
if N < 1
return
end
ANS = true;
end
function [ ANS ] = IS_SERIES( S );
ANS = false;
if isempty( S ) == true
return
end
if isnumeric( S ) == false
return
end
2
if isreal( S ) == false
return
end
if isvector( S ) == false
return
end
if any( isnan( S ) ) == true
return
end
if any( isinf( S ) ) == true
return
end
end

```

ANS = true;

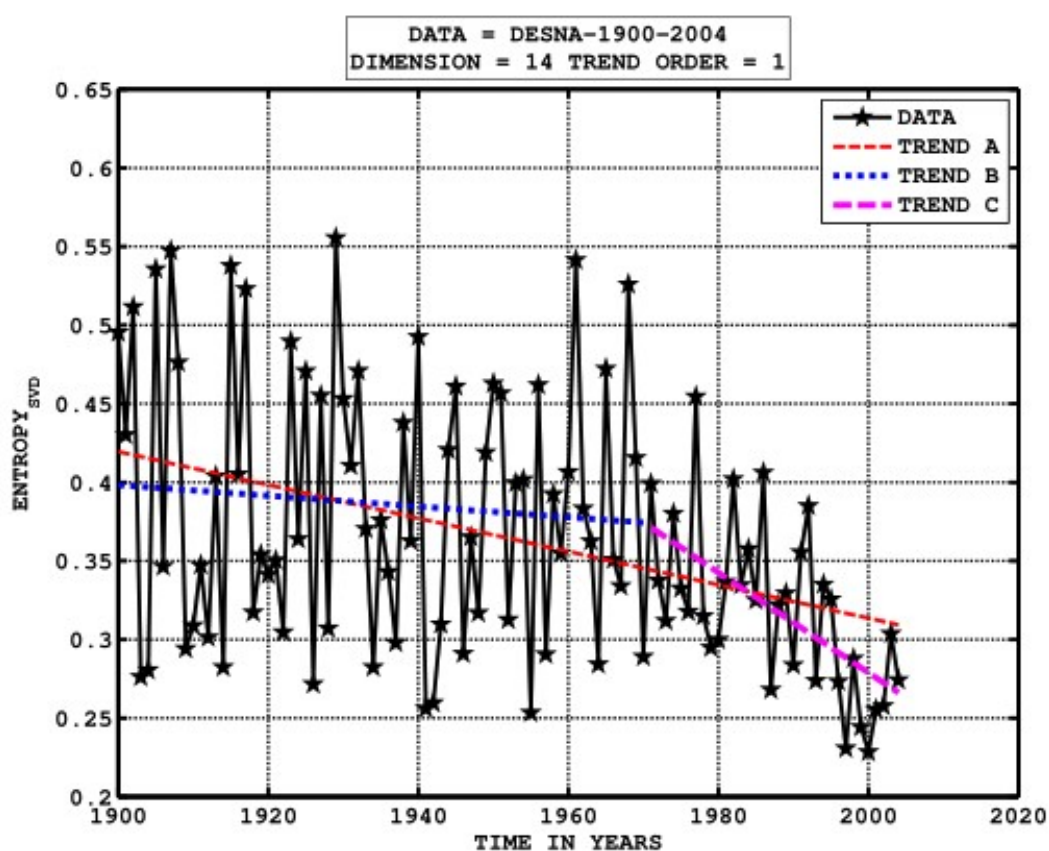
end

### SVD-энтропия гидрологических временных рядов

Для оценки *SVD*-энтропии гидрологических рядов были использованы среднесуточные данные расходов, содержащие 365 значений в году.

Данные за 29 февраля високосных годов из рассмотрения исключались, что было сделано для упрощения процедуры обработки данных.

На рисунке 1 приведен рассчитанный график *SVD*-энтропии для р. Десна.



**Рисунок 1** – График *SVD*-энтропии для р. Десна

Энтропия рассчитывалась для каждого года отдельно, начиная с 1900 года и заканчивая 2004 годом включительно (всего 105 точек ряда).

Для большей наглядности графика были проведены тренды в виде прямых линий.



Тренд *A* демонстрирует заметное уменьшение энтропии за весь анализируемый период. При этом максимальное ее значение не превосходило величины 0,55.

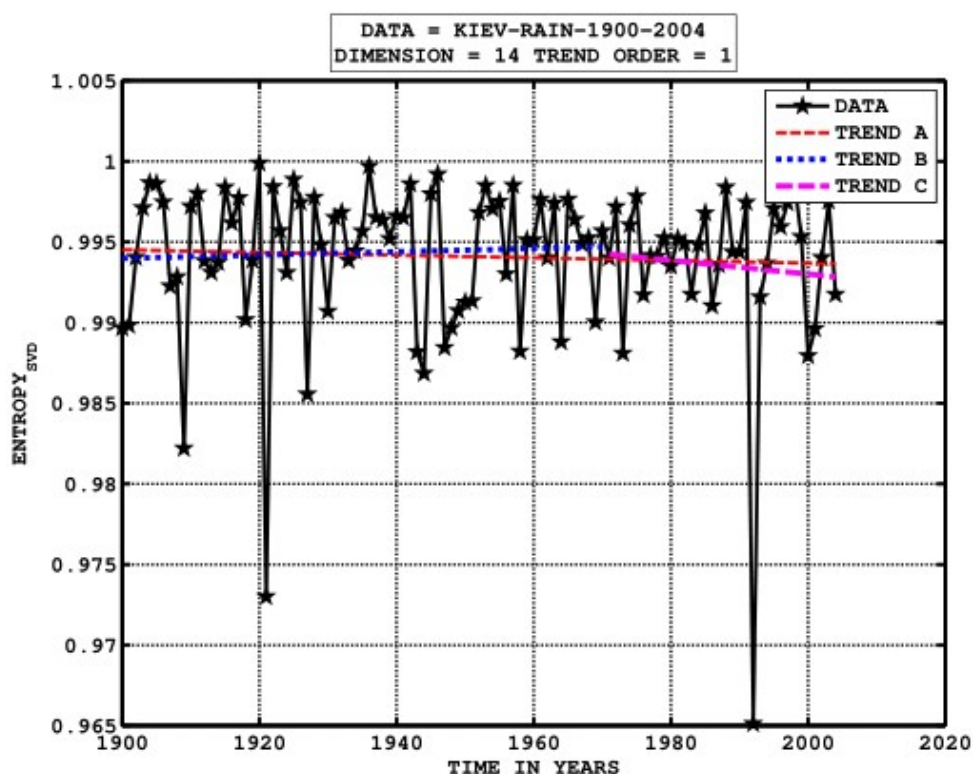
Тренд *B* был проведен через данные по энтропии, начиная с 1900 г. и заканчивая 1970 годом.

С позиций этого линейного тренда видно, что за этот период энтропия также уменьшалась, хотя и крайне незначительно.

Что касается тренда *C*, то график энтропии за период 1970...2004 годы имеет существенный наклон, величина ее резко снижается.

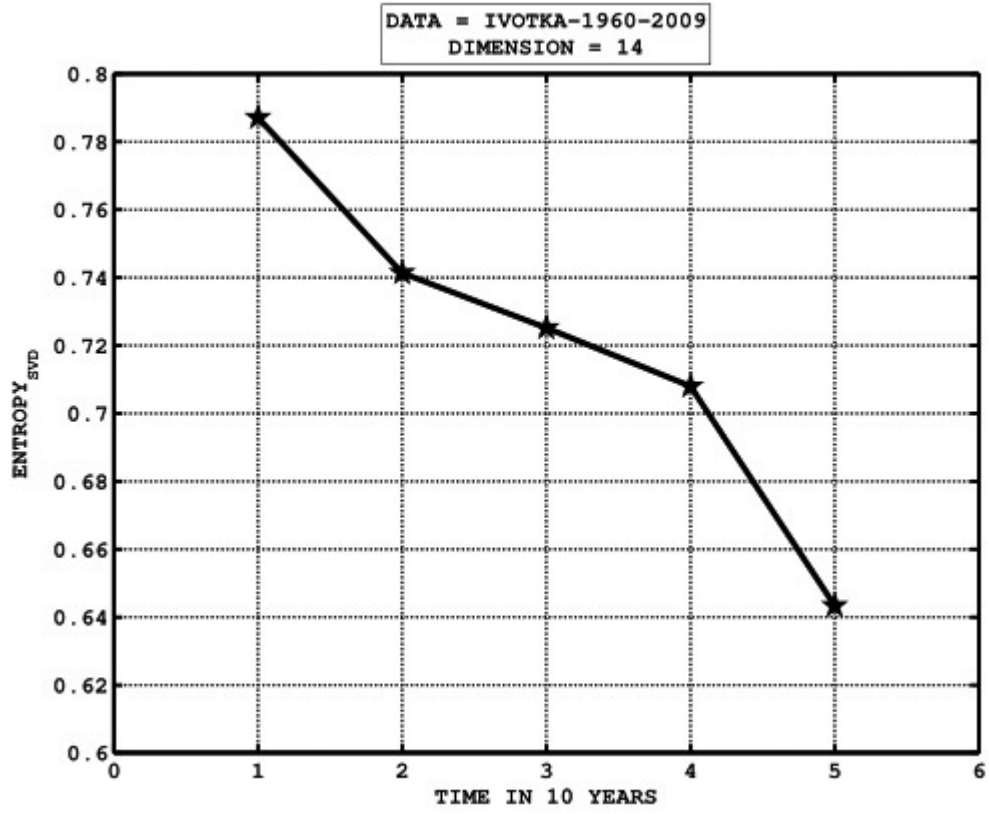
Таким образом, можно говорить о том, что анализируемый ряд среднесуточных расходов в целом по своему поведению достаточно упорядочен и предсказуем.

Для сравнения, рассчитанная *SVD*-энтропия для ряда осадков демонстрирует близость ее численных значений к единице (см. рис. 2), что говорит о высокой степени непредсказуемости.

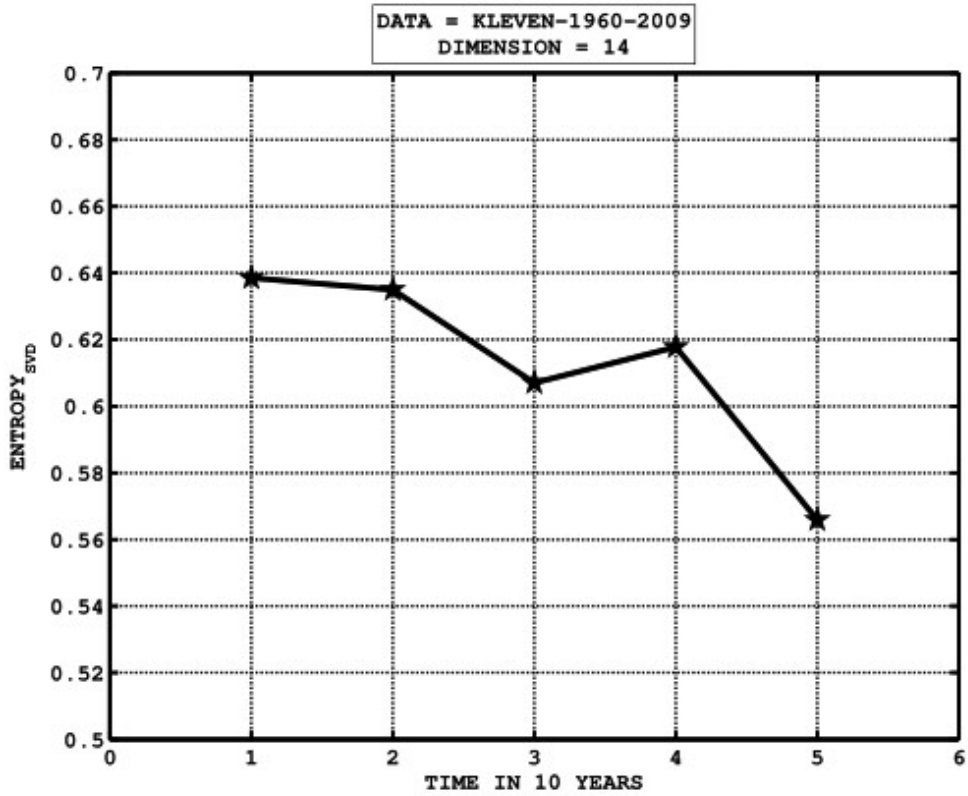


**Рисунок 2** – *SVD*-энтропия временного ряда осадков

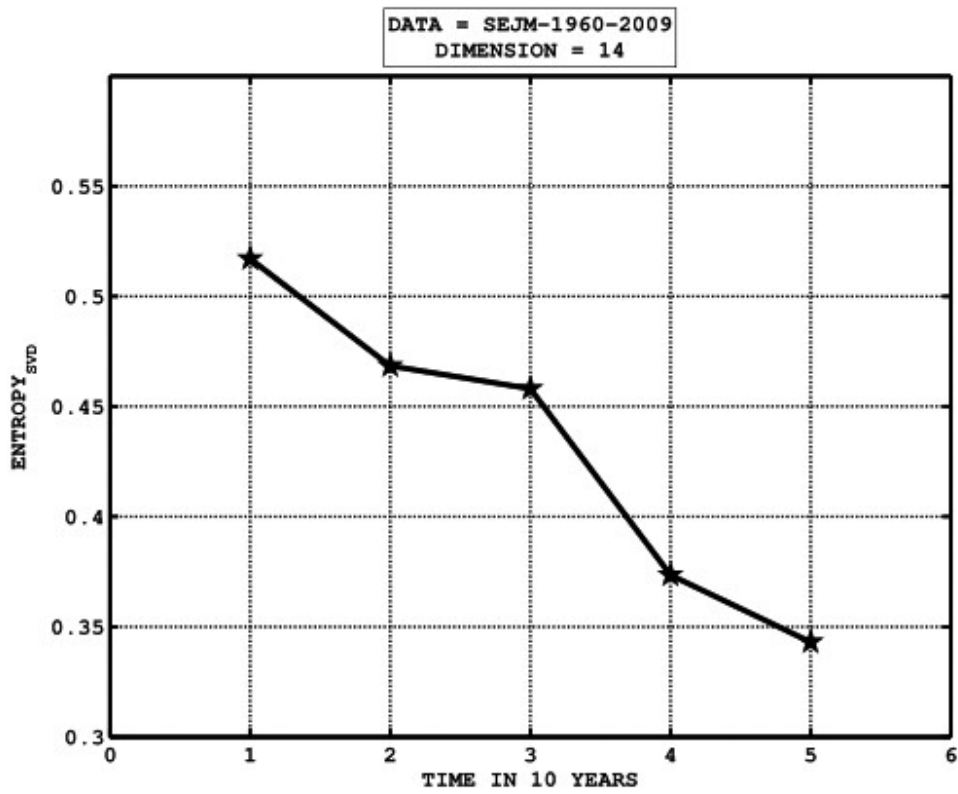
Проанализируем далее результаты расчетов *SVD*-энтропии для четырех равнинных рек, принадлежащих бассейну р. Десна и отличающихся по величине среднесуточных расходов (рис.3).



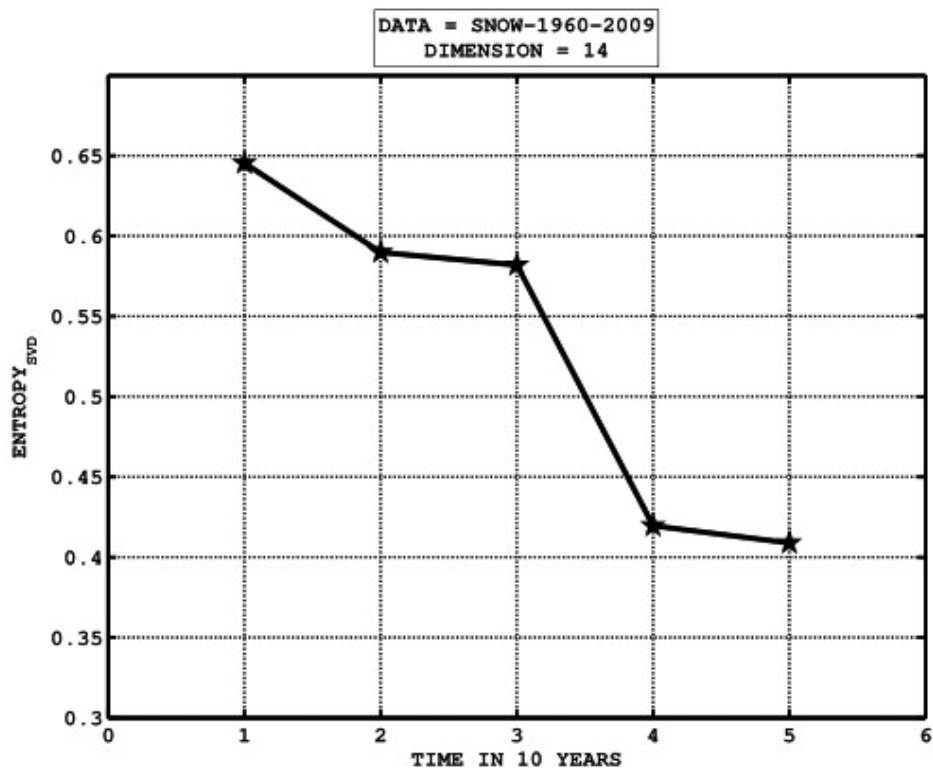
a)



b)



В)



Г)

**Рисунок 3** – Результаты расчетов *SVD*-энтропии для рек Ивотка (а), Клевен (б), Сейм (в) и Снов (г)

В данном случае также четко прослеживается существенное уменьшение энтропии за рассматриваемый период наблюдений для всех четырех рек.

Если такая тенденция в ближайшем будущем сохранится, следует ожидать дальнейшего улучшения качества гидрологических прогнозов.

Однако необходимо отметить, что речь идет о среднем качестве прогнозов. При этом нет полной уверенности в том, что аномальные значения расходов будут также предсказываться лучше.

Понятно, что значение энтропии природных рядов не может все время уменьшаться, поскольку ее долговременное изменение на самом деле носит сложный колебательный характер.

### **Выводы**

Изложены свойства *SVD*-энтропии применительно к анализу гидрологических временных рядов.

Как показали проведенные исследования, для разных рек одного бассейна наблюдается согласованное поведение их *SVD*-энтропий.

При этом четко прослеживается уменьшение энтропий по ходу времени.

С точки зрения практики это означает, что качество гидрологических прогнозов, полученных для периодов последних десятилетий с помощью адекватных методов, будет заведомо лучше, чем для прогнозов первых десятилетий всего рассмотренного временного интервала.

Если такая тенденция в поведении *SVD*-энтропий сохранится, можно ожидать дальнейшего улучшения качества прогнозов.

### **Література**

1. Вильсон А. Дж. Энтропийные методы моделирования сложных систем. – М.: Наука, 1978. – 248 с.
2. Чумак О.В. Энтропии и фракталы в анализе данных. – М.: –Ижевск: НИЦ «Регулярная и хаотическая динамика», Институт компьютерных исследований, 2011. – 164 с.
3. Шеннон К.Э. Работы по теории информации и кибернетике. – М.: Изд-во иностр. лит., 1963. – 829 с.
4. Безручко Б.П. Математическое моделирование и хаотические временные ряды/ Б.П. Безручко, Д.А. Смирнов. – Саратов: ГосУНЦ «Колледж», 2005. – 320 с.

### **Рецензенти**

Кузло М.Т., д-р техн. наук, НУВГП (Рівне)

Гончаренко Ф.П., канд. техн. наук, ДП “Укрдіпроддор” (Київ)

### **Reviewers**

Kuzlo M.T., Dr.Tech.Sci., NUWM (Rivne)

Honcharenko F.P., Ph.D., “Ukrdiproddor” (Kyiv)