

## ДЕКОМПОЗИЦИЯ НА ЭМПИРИЧЕСКИЕ МОДЫ В ЗАДАЧЕ КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ ГЕОФИЗИЧЕСКОЙ ПРИРОДЫ

*Запорожцев И.Ф., аспирант Мурманского государственного технического университета,  
e-mail: zaporozhtsev.if@gmail.com;*

*Серёда А.-В.И., д.т.н., профессор Мурманского государственного технического университета,  
e-mail: seredavi@mstu.edu.ru*

**Ключевые слова:** краткосрочный прогноз, пространственное распределение физической характеристики, кластеризация, аномалии уровня моря, Арктика.

### Введение

Моделирование природных динамических процессов в большинстве случаев приводит к построению многопараметрических математических моделей, например, в виде систем дифференциальных уравнений [1-2]. Зачастую параметры таких моделей не могут быть измерены или оценены с приемлемой точностью, что существенно снижает практическую ценность таких построений для прикладных исследований. В качестве альтернативы используется подход, основанный на формировании и анализе многомерных временных рядов по регулярно обновляемым данным естественно-научной природы, накапливаемым в заданной пространственной области и отражающих пространственно-временную изменчивость той или иной характеристики исследуемого динамического процесса. Такой подход особенно эффективен в случаях, когда природная динамическая система (область атмосферы, акватория океана и т.д.) испытывает большое количество трудно фиксируемых в отдельности внешних воздействий, а среди многочисленных характеристик состояния этой системы практический интерес представляют лишь некоторые из них. Известные методы анализа многомерных временных рядов, как правило, представляют собой обобщения методик статистического анализа их одномерных аналогов. В результате, к каждому ряду применяются ограничения, связанные, прежде всего, с требованием стационарности, так как корректность статистических моделей (например, *ARIMA* [3]), обоснована именно для рядов этого класса. Основным инструментом прогностической деятельности являются на практике линейные регрессионные модели [3]. В рамках моделей, основанных на идее максимального подобия [4], ряды рассматриваются либо в значительной мере изолированно друг от друга, либо в системе, но без использования информации о специфике пространственного распределения исследуемой характеристики. Методики анализа пространственно-временной изменчивости с помощью генетических алгоритмов [5] или эмпирических ортогональных функций (по сути, метода главных компонент) [6] также не предполагают учёта про-

*Рассматриваются две аддитивные модели многомерного временного ряда: модель, полученная в результате выполнения многомерного сингулярного спектрального анализа (MSSA), и модель, построенная как декомпозиция ряда на эмпирические моды (EMD). Первая модель предполагает также экстраполяцию рядов по линейным рекуррентным формулам, а вторая может быть использована лишь как элемент предобработки при решении задачи прогноза. Авторами делается попытка совместного использования моделей класса EMD и MSSA при решении задачи краткосрочного прогноза, что является новизной публикации, а также предлагается сравнительный анализ результатов экстраполяции многомерных временных рядов, предобработанных в соответствии с различными модификациями EMD. Целевой характеристикой для экспериментов выбраны аномалии уровня моря, исследуемая область – некоторая акватория Арктики.*

странственного расположения пунктов, проводимые измерения в которых и образуют эти ряды. Таким образом, важной задачей является корректное формирование многомерного ряда, предполагающее, как минимум, что объединяемые в систему одномерные временные ряды должны удовлетворять определённым количественным критериям, а сама система должна быть физически интерпретируемой. В данной работе будем полагать, что временные ряды в пространственном отношении соответствуют узлам некоторой регулярной сетки, описание которой, а также способ построения многомерного ряда представлены в следующем разделе.

### Формирование многомерного временного ряда

В прямоугольной пространственной области  $\Omega = [a, b] \times [c, d] \subset R^2$  задана регулярная сетка  $\Delta = \Delta x \times \Delta y$ , где:

$$\Delta_x = \{x_k \mid a = x_0 < \dots < x_n = b, x_k = x_{k-1} + h_x, k=1, 2, \dots, n\};$$

$$\Delta_y = \{y_k \mid c = y_0 < \dots < y_m = d, y_k = y_{k-1} + h_y, k=1, 2, \dots, m\}.$$

Так как значения любой характеристики, измеренные в узлах этой сетки в выбранный момент времени, удобно представлять в виде матрицы, а географические координаты узлов не важны для дальнейших расчётов, регулярная сетка может быть определена как множество пар индексов (строчного и столбцового) соответствующих узлов:

$$G = [1, d_x] \times [1, d_y] \subset N^2, \text{ где } d_x = m + 1 \text{ и } d_y = n + 1.$$

Очевидно, что при таком определении сетки  $G$  её шаг по обеим осям равен единице.

Для каждого узла сетки  $G$  известен временной ряд, отражающий временную изменчивость целевой харак-

теристики исследуемого динамического процесса, причём временные ряды для каждого из узлов сформированы единообразно: совпадают стартовый временной отсчёт, шаг по времени между измерениями и анализируемая длина временного ряда (временной интервал или количество отсчетов). Такой подход к формированию временных рядов позволяет говорить, что в области  $\Omega$  задан многомерный временной ряд  $F_T = (F_T^1, F_T^2, \dots, F_T^s)$  длины  $T$ , где  $s = d_r \cdot d_c$ . Понятно, что в силу упомянутых ограничений, ряд  $F_T$  может быть сформирован по измеренным данным, вероятно, лишь в результате некоторой их аппроксимации.

### Методика прогноза

Основным прогностическим средством в данной работе является метод  $K$ -продолжения многомерного временного ряда, предполагающий проведение многомерного сингулярного спектрального анализа (далее метод *MSSA* [7]). Известно, что метод *MSSA* имеет преимущество перед своим одномерным аналогом из класса *SSA* только при работе с системой согласованных одномерных временных рядов. Под согласованностью понимается сходство собственных троек сингулярного разложения траекторной матрицы каждого из анализируемых рядов в отдельности, из чего следует существенная близость слагаемых аддитивной модели каждого ряда. Кроме того, желательно обеспечить стационарность исследуемых временных рядов, хотя это требование не является обязательным.

В этой связи предлагаемая вниманию методика формирования прогноза предполагает предварительную подготовку многомерного временного ряда к использованию метода *MSSA*. Эта подготовка включает в себя разбиение (кластеризацию) исследуемого многомерного временного ряда на конечное число подмножеств временных рядов, удовлетворяющих заданному критерию согласованности и необходимую предобработку временных рядов (удаление выбросов, сглаживание, взятие разностей, разложение на аддитивные составляющие и т.д.). Рассмотрим указанные составляющие подготовки более детально.

### Кластеризация

Кластеризация предполагает разбиение множества временных рядов, образующих исходный многомерный ряд, на конечное число групп, соответствующих так называемым кластерам. Под кластером понимается такая односвязная пространственная область, что для временных рядов, соответствующих принадлежащим этой области узлам пространственной сетки  $G$ , количественное значение некоторого критерия согласованности (сходства) превышает наперед заданное значение. Каждому выделенному кластеру соответствует многомерный (кластерный) временной ряд, который из соображений удобства также будет иметь обозначение  $F_T = (F_T^1, F_T^2, \dots, F_T^s)$ , где  $s$  равно числу принадлежащих кластеру узлов прямоугольной сетки  $G$ .

Введенное понятие кластера позволяет при необходимости использовать этот термин, как для множества принадлежащих ему узлов, так и для соответствующего

ему множества временных рядов – кластерного многомерного временного ряда. Каждый из сформированных в результате кластеризации кластерных многомерных временных рядов, объединяет в определенном смысле более согласованные одномерные ряды, чем соответствующий всей области  $G$  многомерный временной ряд. Это позволяет ожидать большей эффективности использования для их последующей обработки метода *MSSA*.

При группировке (кластеризации) одномерных временных рядов в качестве критерия степени их согласованности будем использовать значение коэффициентов кросс-корреляции [8-9]. Кроме того, будем руководствоваться следующими правилами:

1) можно группировать узлы, находящиеся внутри прямоугольных областей, так как исходная регулярная сетка является прямоугольной;

2) коэффициент кросс-корреляции, вычисленный для любой пары узлов (то есть для соответствующей пары временных рядов) одного кластера, не должен быть меньше наперед заданного порогового значения;

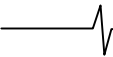
3) для любой пары узлов одного кластера наибольшее значение коэффициента корреляции для заданных фрагментов соответствующих им рядов должно приходиться на нулевой лаг;

4) кластеры можно укрупнять, объединяя прямоугольные кластеры в соответствии с заданным алгоритмом: объединяются пары имеющих общую пространственную границу кластеров с наибольшим значением наименьшего коэффициента кросс-корреляции, который вычисляется для пар узлов (то есть для соответствующих пар временных рядов), взятых из разных кластеров.

### Предобработка временных рядов

Для повышения качества решения задачи прогнозирования значений временного ряда на заданный горизонт целесообразно, как уже отмечалось, осуществить предобработку ряда: удаление выбросов, сглаживание, взятие разностей, разложение на аддитивные составляющие. Кроме того, линейные регрессионные модели (в *MSSA* используется модель этого класса) наиболее эффективны для стационарных временных рядов. В этой связи предлагаемая в работе схема предобработки имеет в своей основе использование метода эмпирической модовой декомпозиции (*EMD* [10]). Важно отметить, что этот метод разложения на аддитивные составляющие сам по себе не позволяет выполнять прогноз, однако зачастую применяется на подготовительном этапе в том числе для прогноза климатических, метеорологических и океанологических характеристик [11]. Суть *EMD* сводится к систематическому построению верхней и нижней огибающих одномерного временного ряда: кривых, проходящих через локальные минимумы и максимумы. При этом аддитивные компоненты ряда (эмпирические моды) формируются в результате вычитания полусумм огибающих из ряда. Построение огибающих осуществляется посредством интерполяции, например, естественными кубическими сплайнами.

Общим недостатком всех существующих методов, основанных на интерполяции, в рамках *EMD* является краевой эффект [12]: необходимо либо использовать первый и



последний отсчеты ряда и включать их как граничные точки на интерполяционной кривой, либо дополнять ряд экстремумами, выходя за пределы ряда. В последнем случае придется использовать временные отсчеты, для которых данных нет, что не приемлемо при решении задачи прогнозирования. В этой работе предлагается использовать граничные точки ряда как точки на огибающих. На обеих границах при сплайн-интерполяции естественно требовать наличие нулевой второй производной. Это требование можно поставить под сомнение из соображений качества прогноза: для сравнения результатов можно использовать, например, условие нулевой первой производной.

*EMD* направлен на выделение именно ортогональных компонент в разложении ряда. Процесс построения таких аддитивных составляющих получил название просеивания.

Последовательно генерируемые разности ряда и полусуммы его огибающих будут иметь большую степень стационарности, чем исходный ряд, поэтому можно предположить, что использование линейной регрессионной модели *MSSA* к аддитивным составляющим, выделяемым в рамках идей *EMD*, будет более эффективным. Компоненты каждого одномерного ряда с одинаковыми номерами формируют новые многомерные временные ряды, к каждому из которых применяется *MSSA*, а результаты прогноза для каждого из них суммируются с целью получения итогового значения.

Важно отметить, что метод *EMD* был разработан для одномерных временных рядов. В процессе просеивания определяется количество аддитивных компонент каждого ряда в отдельности. Очевидно, что сильная положительная корреляционная связь не может быть гарантией совпадения аддитивных моделей всех рядов многомерного временного ряда по структуре (числу компонент). Структурная несогласованность моделей для одномерных рядов системы приводит либо к невозможности применения *MSSA* для анализа многомерного временного ряда, либо к необходимости отбрасывания некоторых компонент разложения для части рядов, что не может быть физически интерпретируемо.

Для адаптации *EMD* к задаче многомерного прогноза можно использовать следующие модификации:

1) отказаться от процесса просеивания (и ортогональности компонент, как следствие), введя критерий останова, гарантирующий одно и то же число компонент в разложении всех рядов;

2) строить огибающие так, чтобы они проходили через точки, у которых номера временных отсчетов были бы одинаковы для всех рядов.

Далее будут представлены эксперименты для следующих модификаций первого типа:

1) разложение каждого временного ряда, входящего в исследуемый многомерный ряд, на 4 компоненты: три из них являются полусуммами огибающих, а четвертая – остаток;

2) разложение каждого ряда на 4 компоненты: первые три формируются в результате сплайн-интерполяции через точки, полученные при вычислении полусуммы кусочно-линейных огибающих, а последняя – остаток.

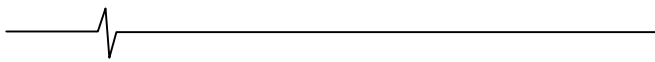
В настоящее время можно выделить две основные работы [13-14] по адаптации идей *EMD* для многомерного случая (в общем виде). Если в процессе декомпозиции стремиться выделять в достаточной степени ортогональные и стационарные компоненты, то можно принять подход из статьи [13]. Рассмотрим вектор-функцию  $F(t)$ , которая вводится для интерполяции значений многомерного временного ряда  $FT$  и является непрерывно дифференцируемой,  $F(t) \in C^1$ :

Так как понятие экстремума для вектор-функций в общем случае не определено, авторы [13] предлагают понятие экстремумов осцилляции (oscillation extrema) как множества локальных минимумов скалярной функции  $\beta(t) : \beta(t) = \left\| \frac{dF}{dt} \right\|^2$ , где  $\|\cdot\|$  используется для обозначения евклидовой нормы. Легко видеть, что множество экстремумов осцилляции в случае скалярной функции  $F(t)$  включает в себя как экстремумы в классическом смысле, так и седловые точки этой функции. Для каждого временного ряда в отдельности строится пара огибающих, одна из которых проходит через экстремумы осцилляции с четными номерами, а другая – с нечетными. Граничные точки каждого ряда также используются для вычисления огибающих. Важно, что все огибающие в качестве узловых точек используют одни и те же временные отсчеты.

### Вычислительный эксперимент

Для оценки практической эффективности предлагаемой методики анализа и прогнозирования многомерного временного ряда была проведена серия вычислительных экспериментов для различных сочетаний и модификаций конкретных реализаций основных ее составляющих. Результаты представлены в таблице:

Обозначение	Характеристика
MSSA	Многомерное $K$ -продолжение, инструмент многомерного сингулярного спектрального анализа
MEMD <sub>1d</sub> MEMD <sub>2d</sub>	Многомерный вариант декомпозиции на эмпирические моды, использующий множество экстремумов осцилляции для интерполяции; огибающие строятся с помощью сплайнов, условие на границе (нижний индекс в названии) – нулевая первая производная (1d) или вторая (2d); прогноз для каждой многомерной эмпирической моды проводится по MSSA
E <sub>1d</sub> E <sub>2d</sub>	Аддитивными компонентами объявляются три последовательно взятые полусуммы огибающих, построенных путём сплайн-интерполяции, а также остаток ряда, полученный в результате серии вычитаний; прогноз компонент по MSSA; нижний индекс интерпретируется аналогично предыдущему пункту
LE <sub>1d</sub> LE <sub>2d</sub>	Аддитивными компонентами объявляются три последовательно взятые полусуммы огибающих и остаток; в отличие от предыдущего пункта огибающие являются кусочно-линейными функциями, а их полусумма вычисляется только в точках локальных экстремумов, а затем интерполируется сплайнами; прогноз компонент по MSSA; нижний индекс интерпретируется аналогично предыдущему пункту



Прогноз для каждого временного отсчёта в кластере выполняется по фрагменту ряда длиной  $N$  на заданный горизонт  $h$  с помощью рекурсивного оператора  $MSSA$ . Прогнозные значения также образуют многомерный ряд с количеством отсчетов, равным  $h$ :

$$\tilde{F}_h(t) = (\tilde{F}_h^1(t), \dots, \tilde{F}_h^s(t)) = MSSA(F_N(t_0), \dots, F_N(t-1))$$

$$\begin{aligned} \dots \\ \tilde{F}_h(t+h-1) &= MSSA^h(F_N(t_0), \dots, F_N(t-1)) = \\ &= MSSA^{h-1}(F_N(t_0+1), \dots, F_N(t-1), \\ &MSSA(F_N(t_0), \dots, F_N(t-1))) \end{aligned}$$

О точности прогноза для кластерного многомерного временного ряда можно судить по нормированной относительной ошибке: абсолютной ошибке, делённой на разность  $D(T)$  наибольшего и наименьшего значений всех временных рядов длины  $T$ , формирующих исследуемый многомерный временной ряд:

$$D(T) = \max_{i=1, \dots, s} \left\{ \max_{t \leq T} F_T^i(t) - \min_{t \leq T} F_T^i(t) \right\}.$$

Для оценки точности прогноза будем использовать следующую формулу, в основе которой лежит использование нормированной среднеквадратичной ошибки ( $NRMSE$ ) в %:

$$\varepsilon_i^h(F_T, \tilde{F}_h) = 100 \cdot \max_{i=1, \dots, s} \frac{1}{D(T)} \sqrt{\frac{\sum_{j=0}^{h-1} (F_T^i(t+j) - \tilde{F}_h^i(t+j))^2}{h}}.$$

**Исходные данные для экспериментов**

Для проведения вычислительных экспериментов использовались данные о значениях аномалий уровня моря [1-2, 15] (альтиметрические данные) для выбранной акватории Баренцева моря (рис. 1): 25 градусов восточной долготы – 44,7 градусов восточной долготы (отсчёты в системе Aviso [15]: 75-134), 71 градус северной широты– 76,4 градуса северной широты (отсчёты в системе Aviso: 764-823). Регулярная сетка для этой акватории имеет размеры 60×60 узлов. Ежедневные данные о значениях целевой характеристики представлены за годовой период, оканчивающийся 12 августа 2013 года.

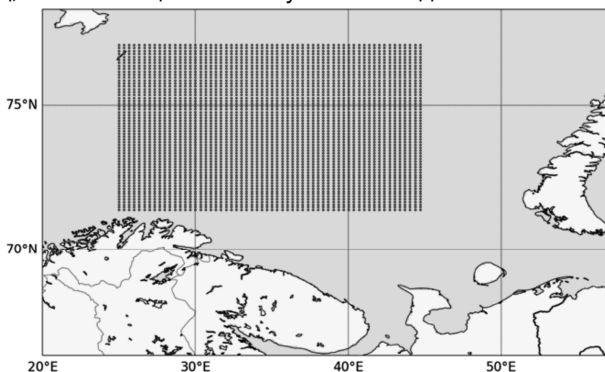


Рис. 1. Изображение узлов сетки  $G$  на цилиндрической проекции Миллера

Данные для временных отсчётов, не превышающих 150 отсчетов, будут использоваться при вычислениях, а последующие (также 150 отсчетов) – для сравнения результатов прогноза, полученных различными методами, с фактическими данными. Таким образом, в расчётах будем

использовать данные за период продолжительностью 300 дней.

Разобьём сетку  $G$  на четыре равные подобласти (рис. 2). Вычислительные эксперименты выполним для кластеров из разных подобластей – по одному из каждой подобласти

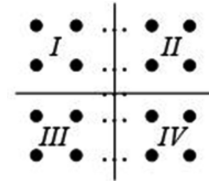


Рис. 2. Разбиение сетки  $G$  на подобласти

Очевидно, что результаты кластеризации зависят от длины фрагмента временного ряда, выбранного для каждого узла. Выберем те кластеры, структура которых не меняется при использовании данных за различные временные промежутки постоянной длины. Выполним кластеризацию отдельно для каждого промежутка времени, как указано на рис. 3.

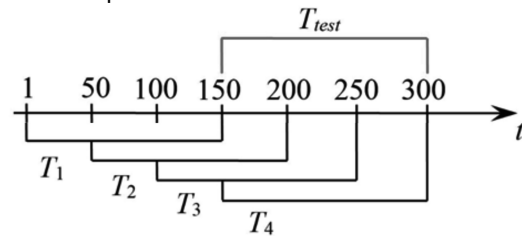


Рис. 3. Промежутки времени, для которых выполнялась кластеризация узлов сетки  $G$

При начальном разбиении области  $\Omega$  на 36 прямоугольников и пороговом значении коэффициента корреляции, равном 0,9, выполнены четыре варианта кластеризации. Для промежутка  $T_1$  получено 236 кластеров, для  $T_2$  имеем 223 кластера, в случае  $T_3$  – 195 кластеров, для  $T_4$  – 232 кластера. Кластеры в каждой из четырёх областей, сохраняющие свою структуру, а также имеющие наибольшее количество узлов, отметим цветом на карте (рис. 4), состоящей из 3600 прямоугольников (прямоугольник используется для изображения узла сетки  $G$ ).

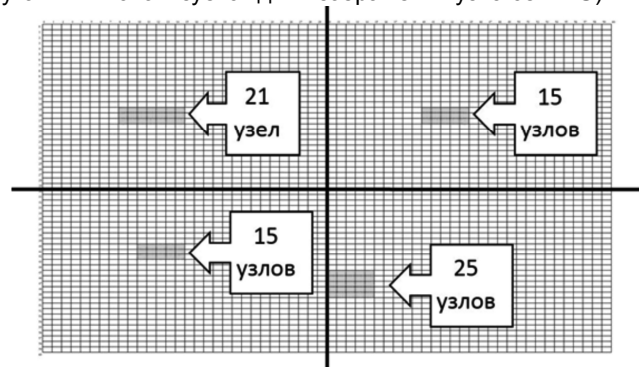
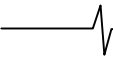


Рис. 4. Кластеры, сохраняющие свою структуру с течением времени и имеющие наибольшее число узлов

**Результаты экспериментов**

Для кластера из каждой области реализован следующий план вычислительных экспериментов:

1. Сформирован многомерный временной ряд из ря-



дов для каждого узла кластера длиной  $T = 365$ .

2. Для каждого ряда отдельно выполнено сглаживание посредством осреднения в скользящем окне длины 5.

3. Назначим горизонт прогноза  $h$  равным 5;

4. Для каждого временного отсчёта  $t$ , начиная со 150 (всего также 150), формируются два фрагмента ряда длиной 100 точек (параметр  $N$ ): обучающий

$$G_N = (F_T(t - N - h), \dots, F_T(t - h - 1)) \quad \text{и} \quad \text{тестовый}$$

$$F_N = (F_T(t - N), \dots, F_T(t - 1)).$$

5. Для каждого фрагмента ряда длиной  $N$  выберем длину окна  $MSSA$  (число строк траекторной матрицы)  $L = 50$ .

6. Для фрагмента  $GN$  строится серия прогнозов на горизонт  $h$ , в которых размерность линейной рекуррентной формулы  $MSSA$  (число значимых сингулярных чисел  $c$ ) равна 1, 2, ...,  $L$ . По результатам этой серии выбирается оптимальный вариант значения  $c_{opt}$ , который доставляет минимум  $NRMSE$ .

7. Для фрагмента  $FN$  выполняется единственный эксперимент ( $c = c_{opt}$ ). Значение  $NRMSE$  для этого фрагмента будем считать итоговым значением  $\varepsilon_i^h(F_T, \tilde{F}_h)$ .

Подобные вычислительные эксперименты осуществляются для каждого из семи упомянутых методов, упорядоченных следующим образом:  $MSSA$ ,  $MEMD_{1d}$ ,  $MEMD_{2d}$ ,  $E_{1d}$ ,  $E_{2d}$ ,  $LE_{1d}$ ,  $LE_{2d}$ .

Для выбранных кластеров из каждой подобласти сетки  $G$  определяется, сколько раз каждый из методов дал результат с ошибкой  $\varepsilon_i^h(F_T, \tilde{F}_h)$ , не превосходящей

10% (в этом случае считаем, что прогноз оправдался), а также количество временных отсчётов, для которых она была наименьшей. Определим также общую результативность (оправдываемость) исследуемых методов – количество отсчётов, для которых хотя бы один из методов обеспечил попадание ошибки в указанный интервал. Полученные значения выразим в процентах и запишем в табл. 1.

По данным из последнего столбца табл. 2 делаем вывод, что удаление трёх методов понизило общую результативность не более чем на 8%. Следует отметить, что на практике использование ансамбля моделей (системы методов в данном случае) затруднено необходимостью эмпирического подбора коэффициентов доверия результату каждой из них. Чем меньше таких моделей в ансамбле (случай единственной модели не рассматривается), тем проще подобрать коэффициенты доверия в линейной комбинации их результатов. Таким образом, исключение трёх методов имеет практический смысл.

Таблицы результатов составлены по изображениям, приведённым на рис. 5. Для каждого кластера, представляющего одну из подобластей области  $\Omega$ , было сформировано изображение в виде таблицы размером  $(150 \times 7)$ . Число строк этой таблицы равно числу временных отсчётов для тестирования, а число столбцов – количеству методов. Номер строки интерпретируется как  $\Delta t$  относительно отсчёта с номером 151 ( $\Delta t = 0, \dots, 149$ ). Методы индексируются, начиная с нуля, в соответствии с их порядком в табл. 1. Ячейка таблицы закрашена, если при выбранном  $\Delta t$  и методе прогноза ошибка  $\varepsilon_i^h(F_T, \tilde{F}_h)$  не превышает 10%, т.е. прогноз оправдался.

Таблица 1. Оправдываемость прогнозов для семи методов

Подобласть сетки	Аббревиатура метода прогноза							$\Sigma$
	MSSA	MEMD <sub>1d</sub>	MEMD <sub>2d</sub>	E <sub>1d</sub>	E <sub>2d</sub>	LE <sub>1d</sub>	LE <sub>2d</sub>	
I	42   17	16   7	34   13	25   7	27   9	38   17	24   8	78
II	40   20	14   3	37   17	33   8	42   7	42   8	42   11	74
III	51   21	28   5	52   17	50   9	47   9	50   14	52   12	87
IV	28   18	15   5	34   18	14   3	27   7	31   14	33   12	77

Таблица 2. Оправдываемость прогнозов для четырёх методов

Подобласть сетки	Аббревиатура метода прогноза				$\Sigma$
	MSSA	MEMD <sub>2d</sub>	LE <sub>1d</sub>	LE <sub>2d</sub>	
I	42   23	34   17	38   19	24   11	70
II	40   22	37   20	42   10	42   19	71
III	51   23	52   25	50   19	52   17	84
IV	28   18	34   19	31   17	33   15	69

## Заключение

В работе предложена концептуальная схема формирования прогноза для многомерных временных рядов достаточно общего вида. На основании приведенных результатов вычислительных экспериментов показано, что повысить оправдываемость прогнозов по  $MSSA$  можно при использовании методики, в основе которой лежит принцип формирования итогового значения прогноза по результатам ансамбля моделей. Эти модели представ-

ляют собой аддитивную совокупность временных рядов, которые могут быть получены при реализации подходов, базирующихся на идеях  $EMD$ . К определенным недостаткам предлагаемого подхода в практическом отношении можно отнести необходимость экспертного подбора коэффициентов доверия для результатов, полученных различными методами. Совершенствование рассмотренной методики прогнозирования может быть достигнуто, например, в результате разработки эффективной схемы коррекции получаемых результатов.

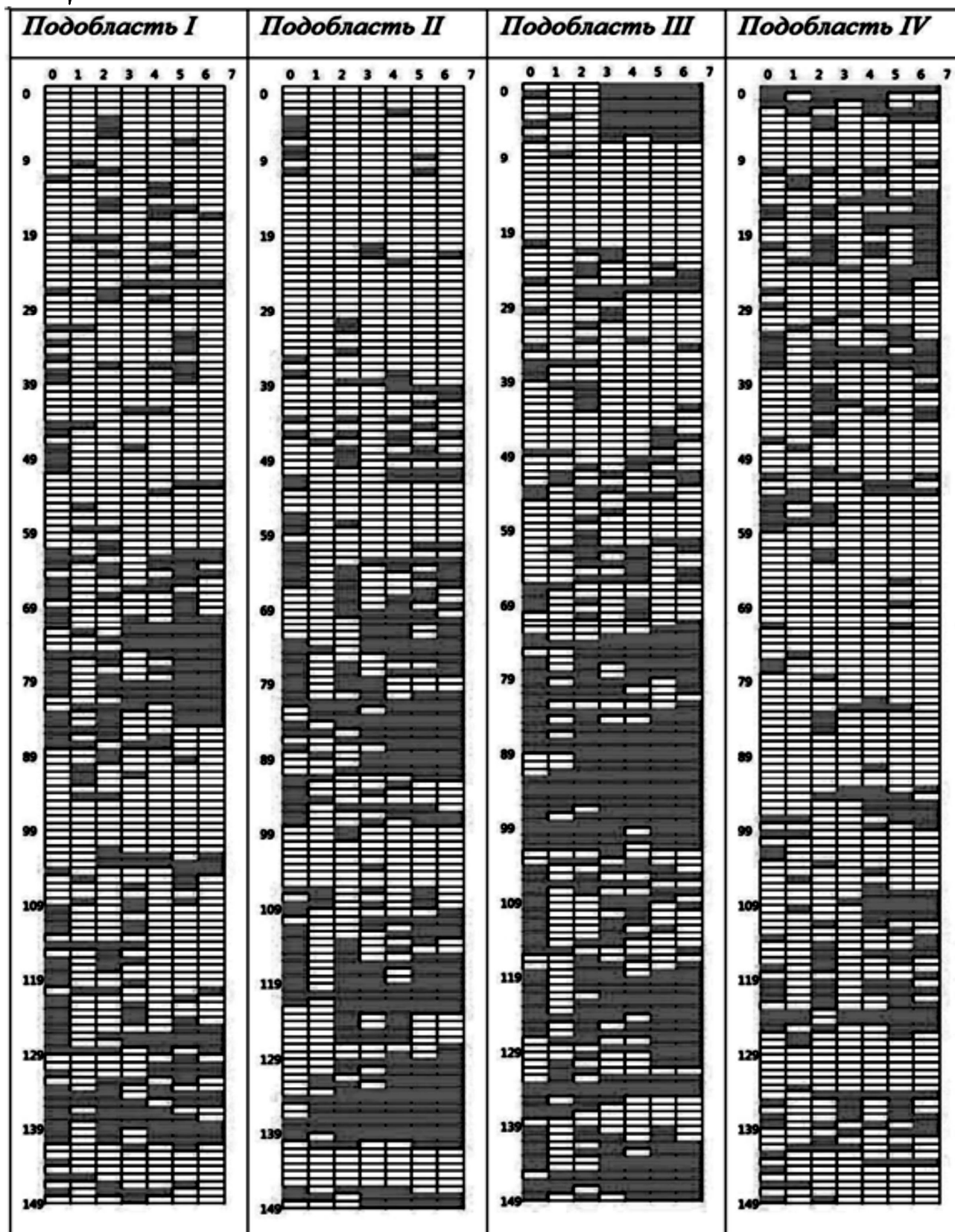


Рис. 5. Таблицы оправдываемости прогнозов, полученных для каждого момента времени с помощью одного из семи методов

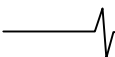
#### Литература

1. Ашик И.М. Численный гидродинамический метод прогноза колебаний уровня в юго-восточной части Баренцева и юго-западной части Карского морей [Электронный ресурс] / Режим доступа: <http://method.hydro->

[met.ru/publ/sb/sb31/sb31.html](http://met.ru/publ/sb/sb31/sb31.html), свободный. – 2005.

2. Вербицкая О.Н. Гидродинамический метод прогноза синоптических колебаний уровня и течений Каспийского моря: дис. канд. физ.-мат. наук [Электронный ресурс] // МГОУНБ. Диссертации. – 2009. – 175 с.

3. Орлов Ю.Н., Осминин К.П. Нестационарные вре-



менные ряды: Методы прогнозирования с примерами анализа финансовых и сырьевых рынков. – М.: Либроком, 2011. – 384 с.

4. Чучуева И.А. Модель прогнозирования временных рядов по выборке максимального подобия: дис. канд. техн. наук [Электронный ресурс] // МГОУНБ. Диссертации. – 2012. – 153 с.

5. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. – Новосибирск: Изд-во Ин-та математики, 1999. – 270 с.

6. Нгонголо Х.К. Статистическое прогнозирование осадков в тропиках по данным о температуре поверхности океана и квазидвухлетних колебаниях зонального потока на примере восточной африки: дис. канд. физ.-мат. наук [Электронный ресурс] // МГОУНБ. Диссертации. – 2011. – 156 с.

7. Степанов Д.А. Варианты метода "Гусеница"-SSA для прогноза многомерных временных рядов / Д.А. Степанов, Н.Э. Голяндина // Труды IV Международной конференции «Идентификация систем и задачи управления» SICPRO'05. – Москва, 2005.

8. Höppner F., Klawonn F. Compensation of Translational Displacement in Time Series Clustering Using Cross Correlation // Lecture notes in computer Science // Advances in Intelligent Data Analysis, vol. 5772, pp.71-82, 2009.

9. Liao T.W. Clustering of Time Series Data - a Survey // Pattern Recognition, vol. 38, pp. 1857-1874, 2005.

10. Huang N. E. "The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Nonstationary Time Series Analysis" [Электронный ресурс] / Режим доступа: [http://keck.ucsf.edu/~schenk/Huang\\_et al98.pdf](http://keck.ucsf.edu/~schenk/Huang_et al98.pdf), свободный. – 1998.

11. Yang P. at al. The Prediction of Non-Stationary Climate Time Series Based on Empirical Mode Decomposition // Advances in Atmospheric Sciences, vol. 27, pp. 845-854, 2010.

12. Давыдов В.А., Давыдов А.В. Уменьшение краевых эффектов при выполнении эмпирической модовой декомпозиции сигналов преобразования Гильберта-Хуанга [Электронный ресурс] / Режим доступа: [www.actualresearch.ru/nn/2011\\_1/Article/physics.../davydov20111.pdf](http://www.actualresearch.ru/nn/2011_1/Article/physics.../davydov20111.pdf), свободный. – 2011.

13. Fleureau J. at al. Multivariate Empirical Mode Decomposition and Application to Multichannel Filtering // Signal Processing, vol. 91, pp. 2783-2792, 2011.

14. Rehman N., Mandic D.P. Multivariate Empirical Mode Decomposition // Proceedings of the Royal Society A., vol. 466, no. 2117, pp. 1291-1302, 2010.

15. Aviso. Satellite Altimetry Data [Электронный ресурс] /Режим доступа: <http://www.aviso.altimetry.fr/en/home.html>, свободный.

## EMPIRICAL MODE DECOMPOSITION IN GEOPHYSICAL MULTIVARIATE TIME SERIES SHORT-TERM FORECASTING

*Zaporozhtsev I. F., Sereda A.-V.I.*

Two multivariate time series additive models are considered in the article. The first model is obtained as a result of multivariate singular spectrum analysis (MSSA) and the second one is the decomposition of time series into empirical modes (EMD). MSSA model also provides the time series extrapolation with linear recurrent formula. EMD model can only be used as preprocessing tool before forecast. Authors attempt to use EMD and MSSA together for short-term forecasting task that is a novelty element of article and propose comparative analysis of multivariate time series extrapolation results preprocessed with different modifications of EMD. Target characteristic of experiments is sea level anomaly. Studied area is located in the Arctic.

### НАУЧНОЕ НАПРАВЛЕНИЕ:

**Теория и методы цифровой обработки сигналов. Техническая реализация систем обработки сигналов**

#### **Секция 2: Теория и методы цифровой обработки сигналов**

В работах, проводимых по данному базовому направлению теоретических исследований, закладываются фундаментальные основы построения систем ЦОС. Поэтому работа секции вызывает повышенный интерес всех участников конференции. Как показывает анализ представленных докладов, исследования российских ученых ведутся по всем основным направлениям развития мировой науки в области информационных технологий реального времени, а именно: цифровой фильтрации, многоскоростной обработки и вейвлет-преобразования, частотно-временной обработки и спектрального анализа, теории быстрых алгоритмов ЦОС, оптимальной и адаптивной обработки сигналов и нелинейной фильтрации.

На секции обсуждались следующие актуальные научные проблемы:

- многоскоростная адаптивная фильтрация;
- моделирование в MATLAB алгоритмов адаптивной фильтрации;
- адаптивная фильтрация прерывистых сигналов;
- минимально-фазовая БИХ-фильтрация;
- спектральный анализ узкополосного сигнала с повышенным разрешением;
- оценивание формантных частот;

- идентификация звуковых сигналов;
- анализ биомедицинских сигналов с использованием преобразования Гильберта;
- анализ электроэнцефалограмм с использованием вейвлет-преобразования;
- оценивание мощности узкополосного сигнала на фоне близко расположенных помех.