

УДК 349.6

**Ю.Е. Марковский**

Донецкий национальный технический университет, г. Донецк  
кафедра электронной техники  
E-mail: [y.markovskyy@gmail.com](mailto:y.markovskyy@gmail.com)

## ИСПОЛЬЗОВАНИЕ МЕТОДОВ ОЦЕНКИ ТРЕНДА ВРЕМЕННОГО РЯДА ДЛЯ ПРОГНОЗИРОВАНИЯ ЗНАЧЕНИЙ ОБЩЕЙ МИНЕРАЛИЗАЦИИ ПРЭСНЫХ ВОД

### *Abstract*

*Markovskyy Y.E. Usage of Evaluation Methods of Time Series' Trend for Forecasting of Values of Total Mineralization of Fresh Waters. The main principles of short-range ecological forecasting are discussed. Availability of exponential smoothing (and its derivations) for short-range ecological forecasting are shown. Simulation of short-range forecasting of values of total mineralization of fresh water by Holt-Winters method is conducted.*

**Keywords:** *ecological prognostication, total mineralization, exponential smoothing out.*

### *Анотація*

*Марковський Ю.Є. Використання методів оцінки тренда часового ряду для прогнозування значень загальної мінералізації прісних вод. Розглянуто основні принципи короткострокового екологічного прогнозування. Показано можливість застосування експонентного згладжування (і похідних методів) для побудови екологічних прогнозів. Промодельовано побудову короткострокового прогнозу значень загальної мінералізації прісної води методом Хольта-Уинтерса.*

**Ключові слова:** *екологічне прогнозування, загальна мінералізація, експоненціальне згладжування.*

### *Аннотация*

*Марковский Ю.Е. Использование методов оценки тренда временного ряда для прогнозирования значений общей минерализации пресных вод. Рассмотрены основные принципы краткосрочного экологического прогнозирования. Показана возможность применения экспоненциального сглаживания (и производных методов) для построения экологических прогнозов. Промоделировано построение краткосрочного прогноза значений общей минерализации пресной воды методом Хольта-Уинтерса.*

**Ключевые слова:** *экологическое прогнозирование, общая минерализация, экспоненциальное сглаживание.*

**Общая постановка проблемы.** На современном этапе экологическое прогнозирование стало настолько значимым, что этап развития экологии характеризуется интенсивным формированием прогнозирования как самостоятельного научного направления. Это в первую очередь связано с необходимостью построения экологических прогнозов для перспективных народнохозяйственных планов и проектирования новых технологий по использованию природных ресурсов. Под экологическим прогнозированием целесообразно понимать создание таких обобщений и методов, которые обеспечивают выработку прогностических оценок по отношению к наблюдаемым параметрам экосистем. Сугубо практический аспект прогнозирования подчеркивает обязательную заблаговременность и необходимое дальнейшее действие прогноза, его тесные связи с проектированием и перспективными планами. Одной из целей экологического прогнозирования является повышение вероятности предупреждения различных аварийных ситуаций, например, резкого превышения предельно допустимых концентраций вредных веществ в воздухе или воде и т.п.

Хотя от экологического прогнозирования принципиально нельзя требовать предельно точных картин будущего и однозначного предсказания конкретных деталей изменения того или иного параметра в еще не реализованных условиях, необходимо сделать более интенсивным процесс изучения компонентов экосистем и анализировать полученные результаты в комплексе с данными других наук с целью повышения точности и дальности действия прогноза последствий от современного техногенеза.

**Временные ряды. Основные компоненты и влияющие факторы.** Основным инструментом любого прогноза является схема экстраполяции. Различают формальную и прогнозную экстраполяцию. Формальная базируется на предположении о сохранении в будущем прошлых и настоящих тенденций развития объекта прогноза. При прогнозной экстраполяции фактическое развитие увязывается с гипотезами о динамике исследуемого процесса с учетом в перспективе его физической и логической сущности. Основу экстраполяционных методов прогнозирования составляет изучение временных рядов (рядов динамики), представляющих собой упорядоченные во времени наборы измерений тех или иных характеристик исследуемого объекта, процесса.

Первым этапом при выборе стратегии и методов предварительной обработки и анализа рядов динамики является оценка тренда временного ряда. Любой временной ряд может быть разделен на три компоненты [1]:

$$x(t) = f(t) + g(t) + h, \quad (1)$$

где:

- $f(t)$  — детерминированная компонента, представляющая собой некоторую аналитическую функцию, выражающую тенденцию в ряду динамики;
- $g(t)$  — стохастическая компонента, моделирующая характер периодической и квазипериодической вариации исследуемого явления;
- $h$  — случайная компонента типа "белый шум".

Для длинных рядов выделение тренда носит обычно разведочный характер, так как часто невозможно указать подходящую параметрическую кривую для аппроксимации ряда на всей его длине. Для выделения тренда в этом случае используют различные непараметрические методы анализа временных рядов, такие как, сглаживание скользящими средними или скользящими медианами, частотную фильтрацию и т.п. [2, 3]. Получение достаточно гладкой траектории дает возможность визуально оценить наличие тенденции в условиях сильной зашумленности.

Члены временного ряда не являются статистически независимыми и не являются одинаково распределенными. Это значит, что нельзя распространять свойства и правила статистического анализа случайной выборки на временные ряды. С другой стороны, взаимозависимость членов временного ряда создает свою специфическую базу для построения прогнозных значений анализируемого показателя (т.е. для построения оценок  $\hat{x}(N+k)$  для неизвестных значений  $x(N+k)$ ) по значениям наблюдений  $x(1), x(2), \dots, x(N)$ .

На временной ряд влияют следующие факторы:

- долговременные, формирующие общую тенденцию в изменении анализируемого признака  $x(t)$ . Обычно эту тенденцию описывают с помощью той или иной неслучайной функции  $f(t)$  — выражение (1). Это функцию называют функцией тренда или трендом;
- сезонные, формирующие периодически повторяющиеся в определенное время года колебания анализируемого признака;
- циклические, формирующие изменения анализируемого признака, обусловленные действием долговременных циклов демографической, экономической и другой природы;
- случайные, не поддающиеся учету и регистрации.

При прогнозировании на основе статистического анализа временных рядов следует учитывать следующие условия:

- требуемый горизонт  $l$  прогнозирования, т.е. на сколько временных тактов ( $l$ ) вперед мы собираемся строить наш прогноз; при  $l \leq 3$  прогноз обычно называется краткосрочным, при  $3 < l \leq 6$  — среднесрочным, а при  $l > 6$  — долгосрочным (такая классификация достаточно условна: она может зависеть от конкретного смысла одного такта времени — подразумевается ли под ним день, неделя, месяц, квартал, год и т.д.);
- длина анализируемого временного ряда (условно говоря, при  $N < 50$  ряд считается коротким, а при  $N > 50$  — длинным);
- наличие или присутствие в анализируемом ряду сезонной составляющей или каких-либо «нестандартностей» (скачкообразных изменений в поведении тренда, слишком большой величины дисперсии случайных остатков и т.п.).

**Краткосрочное прогнозирование временных рядов.** Для краткосрочного прогнозирования нестационарных временных рядов, имеющих случайные изменения уровня и угла наклона, применяется метод экспоненциального сглаживания, который известен также под названием метода Брауна. Данный метод позволяет построить такое описание процесса, при котором более поздним наблюдениям приписываются большие веса по сравнению с ранними наблюдениями, причем веса наблюдений убывают по экспоненте. Выражение

$$S_t^{[k]}(y) = \alpha \sum_{i=0}^n (1 - \alpha)^i S_{t-i}^{[k]}(y) \tag{2}$$

называется экспоненциальной средней  $k$ -го порядка для ряда  $y_t$ , где  $\alpha$  — параметр сглаживания.

В расчетах для определения экспоненциальной средней пользуются рекуррентной формулой

$$S_t^{[k]}(y) = \alpha S_t^{[k-1]}(y) + (1 - \alpha) S_{t-1}^{[k]}(y). \tag{3}$$

Использование соотношения (3) предполагает задание начальных условий  $S_0^{[1]}, S_0^{[2]}, \dots, S_0^{[k]}$ . Для этого можно воспользоваться формулой Брауна-Мейера, которая связывает коэффициенты прогнозирующего полинома с экспоненциальными средними соответствующих порядков:

$$S_t^{[k]} = \sum_{p=0}^n (-1)^p \frac{\hat{a}_p}{p!} \frac{\alpha \beta}{(k-1)!} \sum_{j=0}^{\infty} j^p \beta^j \frac{(p-1+j)!}{j!}, \tag{4}$$

где  $p = 1, 2, \dots, n+1$ ;  $\hat{a}_p$  — оценки коэффициентов;  $\beta = 1 - \alpha$ .

Можно получить оценки начальных условий, в частности, для линейной модели:

$$\begin{aligned} S_0^{[1]} &= a_0 - \frac{\beta}{\alpha} a_1; \\ S_0^{[2]} &= a_0 - \frac{2\beta}{\alpha} a_1, \end{aligned} \tag{5}$$

для квадратичной модели:

$$\begin{aligned} S_0^{[1]} &= a_0 - \frac{\beta}{\alpha} a_1 + \frac{\beta(2-\alpha)}{2\alpha^2} a_2; \\ S_0^{[2]} &= a_0 - \frac{2\beta}{\alpha} a_1 + \frac{\beta(3-2\alpha)}{2\alpha^2} a_2; \\ S_0^{[3]} &= a_0 - \frac{3\beta}{\alpha} a_1 + \frac{\beta(4-3\alpha)}{2\alpha^2} a_2. \end{aligned} \tag{6}$$

Зная начальные условия  $S_0^{[k]}$  и значения параметра  $\alpha$ , можно вычислить экспоненциальные средние  $S_t^{[k]}$ .

Оценки коэффициентов прогнозирующего полинома определяются через экспоненциальные средние по фундаментальной теореме Брауна-Мейера. Для линейной модели получим:

$$\begin{aligned}\hat{a}_0 &= 2S_t^{[1]} - S_t^{[2]}, \\ \hat{a}_1 &= \frac{\alpha}{\beta}(S_t^{[1]} - S_t^{[2]}),\end{aligned}\tag{7}$$

для квадратичной:

$$\begin{aligned}\hat{a}_0 &= 3(S_t^{[1]} - S_t^{[2]}) + S_t^{[3]}, \\ \hat{a}_1 &= \frac{\alpha}{\beta^2}[(6 - 5\alpha)S_t^{[1]} - 2(5 - 4\alpha)S_t^{[2]} + (4 - 3\alpha)S_t^{[3]}], \\ \hat{a}_2 &= \frac{\alpha^2}{\beta^2}(S_t^{[1]} - 2S_t^{[2]} + S_t^{[3]}).\end{aligned}\tag{8}$$

Прогноз реализуется по выбранному многочлену. Соответственно, для линейной модели

$$\hat{y}_{t+\tau} = \hat{a}_0 + \hat{a}_1\tau, \text{ для квадратичной модели — } \hat{y}_{t+\tau} = \hat{a}_0 + \hat{a}_1\tau + \frac{\hat{a}_2}{2}\tau^2, \text{ где } \tau \text{ — период прогноза.}$$

Важную роль в методе экспоненциального сглаживания играет выбор оптимального параметра сглаживания  $\alpha$ , так как именно он определяет оценки коэффициентов модели, а, следовательно, и результаты прогноза [4]. В зависимости от величины параметра сглаживания прогнозные оценки по-разному учитывают влияние исходного ряда наблюдений: чем больше  $\alpha$ , тем больше вклад последних наблюдений в формирование тренда, а влияние начальных условий быстро убывает. При малом параметре  $\alpha$  прогнозные оценки учитывают все наблюдения, при этом уменьшение влияния более «старой» информации происходит медленно.

Известны два основных соотношения, позволяющие найти приближенную оценку параметра  $\alpha$ . Первое — соотношение Брауна, — выведенное из условия равенства скользящей и экспоненциальной средней:

$$\alpha = \frac{2}{N + 1},\tag{9}$$

где  $N$  — число точек ряда, для которых динамика считается однородной и устойчивой (период сглаживания). Вторым является соотношение Мейера:

$$\alpha \approx \frac{\sigma_n}{\sigma_\varepsilon},\tag{10}$$

где  $\sigma_n$  — среднеквадратическая ошибка модели,  $\sigma_\varepsilon$  — среднеквадратическая ошибка исходного ряда. Однако использование последнего соотношения затруднено тем, что достоверно определить  $\sigma_n$  и  $\sigma_\varepsilon$  из исходной информации крайне сложно.

Для уменьшения ошибки прогноза рекомендуется выбирать минимальный параметр сглаживания  $\alpha$ , в то же время параметр  $\alpha$  определяет начальные условия, и, чем он меньше, тем ниже точность определения начальных условий, а, следовательно, ухудшается и качество прогноза. Таким образом, при определении параметра сглаживания возникает противоречие: с уменьшением  $\alpha$  уменьшается среднеквадратическая ошибка, но возрастает ошибка в начальных условиях, что в свою очередь влияет на качество прогноза.

Также весьма существенным для практического использования является вопрос о выборе порядка прогнозирующего полинома, что во многом определяет качество прогноза. В большинстве случаев использование полинома 1-й степени дает удовлетворительные результаты. Превышение 2-го порядка прогнозирующего полинома не приводит к существенному увеличению точности прогноза, однако значительно усложняет процедуру расчета [5].

**Прогнозирование значений общей минерализации с помощью экспоненциального сглаживания.** Применим процедуру экспоненциального сглаживания для нахождения прогноза значения общей минерализации пресной воды, поступающей на донецкую фильтровальную станцию (ДФС) по данным последних лет (данные по общей минерализации пресной воды за последние

4 года предоставлены донецкой фильтровальной станцией). Промоделируем построение краткосрочного прогноза общей минерализации на 2005г. по данным за 2004г. Прогнозирование будем осуществлять на месяц вперед, после чего для нахождения прогноза на следующий месяц, заменим найденное ранее прогнозные значение общей минерализации фактическим. Для прогнозирования будем использовать полином первого порядка. Результаты помесячного прогнозирования значений общей минерализации с помощью экспоненциального сглаживания (параметр сглаживания  $\alpha = 0,4$ ) на 2005г. представлены в табл. 1.

Таблица 1 — Результаты прогнозирования значений общей минерализации методом экспоненциального сглаживания

Номер месяца 2005г	Фактическое значение общей минерализации, мг/л.	Прогноз минерализации, мг/л.	$\delta$ , %
1	675,80	660,30	2,29
2	759,00	695,98	8,30
3	784,00	687,91	12,26
4	748,00	716,34	4,23
5	668,50	743,41	11,21
6	683,50	745,24	9,03
7	658,00	714,55	8,59
8	596,50	702,13	17,71
9	610,50	684,48	12,12
10	719,00	649,29	9,70
11	784,00	633,77	19,16
12	846,00	667,86	21,06

Из результатов, приведенных в табл. 1 следует, что ошибка прогнозирования в некоторых случаях превышает 20%, что является неудовлетворительным результатом. Следовательно, требуется значительная коррекция результатов прогноза.

Фактические значения общей минерализации пресной воды и данные прогноза за период 2004–2008гг с помощью экспоненциального сглаживания приведены на рис. 1.

В табл. 2 Приведены данные о средних ошибках прогноза общей минерализации по годам, начиная с 2004.

Таблица 2 — Средние ежегодные ошибки прогноза общей минерализации по методу экспоненциального сглаживания за период 2004–2008г

№ п/п	Год	Средняя погрешность прогноза, %
1	2004	12,06
2	2005	11,30
3	2006	13,53
4	2007	9,85
8	2008	9,26

Т.к., погрешность прогнозирования при экспоненциальном сглаживании является значительной (см. табл.1), то целесообразно рассмотреть еще несколько методов, использующих идеологию экспоненциального сглаживания, которые развивают метод Брауна в различных направлениях.

**Усовершенствованные методы краткосрочного прогнозирования.** Холт [6] ослабил ограничения метода Брауна, связанные с его однопараметричностью, введением двух параметров сглаживания  $\lambda_1$  и  $\lambda_2$  ( $0 < \lambda_1 < 1$ ;  $0 < \lambda_2 < 1$ ). В его модели прогноз  $\hat{y}_t^l$  на  $l$  тактов вперед в текущий момент времени  $t$  также определяется линейной моделью вида:

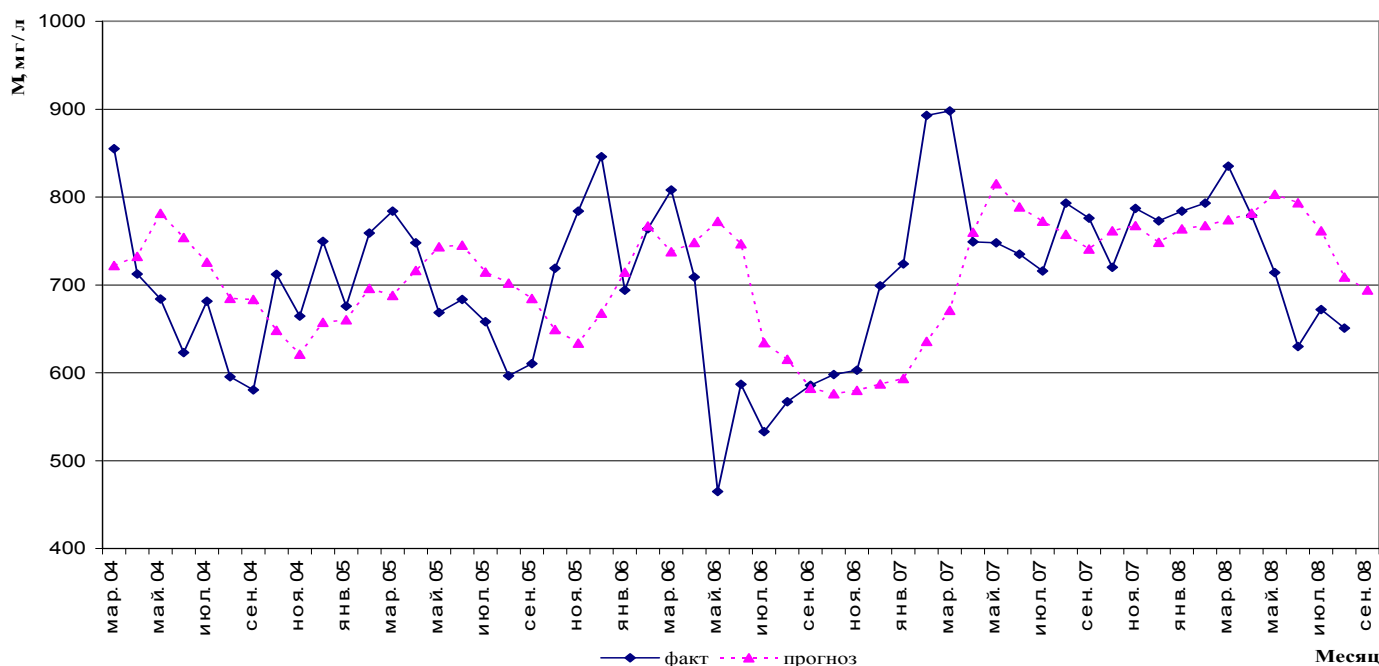


Рисунок 1 — Фактические значения общей минерализации пресной воды и значения общей минерализации, полученные с помощью прогноза по методу экспоненциального сглаживания за период 2004–2008гг.

$$\hat{y}_t^l = \hat{a}_0^t + l\hat{a}_1^t, \tag{11}$$

где обновление прогнозирующих коэффициентов производится по формулам:

$$\begin{aligned} \hat{a}_0^{t+1} &= \lambda_1 y_t + (1 - \lambda_1)(\hat{a}_0^t + \hat{a}_1^t); \\ \hat{a}_1^{t+1} &= \lambda_2 (\hat{a}_0^{t+1} - \hat{a}_1^t) + (1 - \lambda_2)\hat{a}_1^t. \end{aligned} \tag{12}$$

Уинтерс [7] развил метод Хольта так, чтобы он охватывал еще и сезонные эффекты. Прогноз, сделанный в момент времени  $t$  на  $l$  тактов вперед, равен

$$\hat{y}_t^l = (\hat{a}_0^t + l\hat{a}_1^t) \cdot \omega_{t+l-N}, \tag{13}$$

где  $\omega$  — коэффициент сезонности, а  $N$  — число временных тактов, содержащихся в полном сезонном цикле. Сезонность в этой формуле представлена мультипликативно. Метод использует три параметра сглаживания  $\lambda_1, \lambda_2, \lambda_3$  ( $0 < \lambda_j < 1, j=1, 2, 3$ ), а его формулы обновления имеют вид:

$$\begin{aligned} \hat{a}_0^{t+1} &= \lambda_1 \frac{y_{t+1}}{\omega_{t+1-N}} + (1 - \lambda_1) \cdot (\hat{a}_0^t + \hat{a}_1^t); \\ \omega_{t+1} &= \lambda_2 \frac{y_{t+1}}{\hat{a}_0^{t+1}} + (1 - \lambda_2) \cdot \omega_{t+1-N}; \\ \hat{a}_1^{t+1} &= \lambda_3 (\hat{a}_0^{t+1} - \hat{a}_0^t) + (1 - \lambda_3) \cdot \hat{a}_1^t. \end{aligned} \tag{14}$$

Прогноз строится на основании прошлых и текущих значений временного ряда, параметров адаптации  $\lambda_1, \lambda_2, \lambda_3$ , а также начальных значений  $\hat{a}_0^0, \hat{a}_1^0, \omega_0$ .

В работе использован метод Хольта-Уинтерса для прогнозирования значений общей минерализации, т.к. данный метод учитывает сезонные эффекты, которые присутствуют в наблюдаемых рядах данных (ряды значений общей минерализации за 2004–2008гг.).

Проделав аналогичное моделирование, как и при прогнозировании с помощью экспоненциального сглаживания (см. выше), получим результаты помесячного прогноза значений общей минерализации с помощью метода Хольта-Уинтерса (параметр сглаживания  $\alpha_0 = \alpha_1 = 0,3$ ) на 2005г. (см. табл. 3).

Таблица 3 — Результаты прогнозирования значений общей минерализации методом Хольта-Уинтерса

Номер месяца	Фактическое значение общей минерализации	Прогноз минерализации, мг/л.	$\delta$ , %
1	675,80	715,04	5,81%
2	759,00	723,70	4,65%
3	784,00	790,87	0,88%
4	748,00	768,36	2,72%
5	668,50	687,66	2,87%
6	683,50	660,90	3,31%
7	658,00	663,37	0,82%
8	596,50	608,91	2,08%
9	610,50	591,57	3,10%
10	719,00	684,97	4,73%
11	784,00	786,84	0,36%
12	846,00	847,75	0,21%

Из результатов табл. 3 видно, что ошибка прогнозирования в данном случае не превышает 6%, что является вполне удовлетворительным результатом.

Данные о фактическом изменении общей минерализации пресной воды и данные прогноза по методу Хольта-Уинтерса приведены на рис. 2.

Из сравнения результатов, приведенных на рис.1 и рис.2, следует, что прогноз по методу Хольта-Уинтерса дает более точные результаты. В табл. 4 приведены данные о средних ошибках прогноза общей минерализации по годам, начиная с 2004.

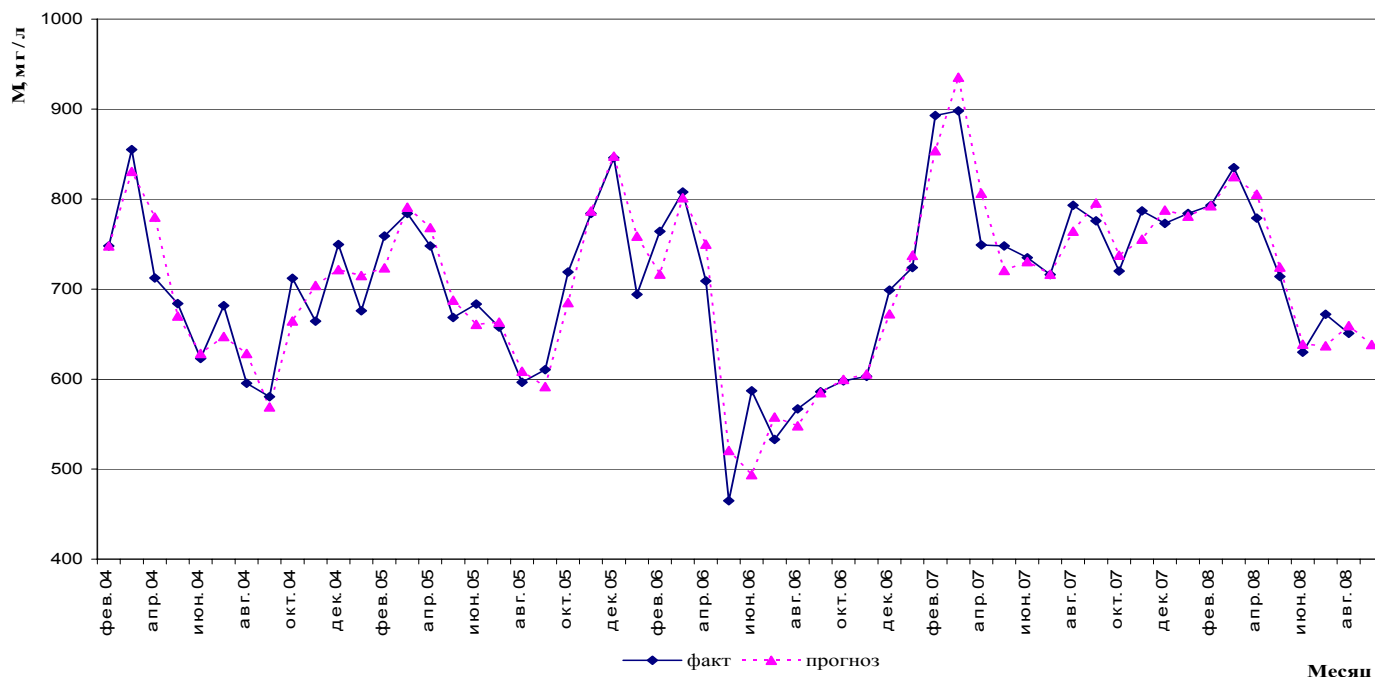


Рисунок 2 — Фактические значения общей минерализации пресной воды и значения общей минерализации, полученные с помощью прогноза по методу Хольта-Уинтерса за период 2004–2008гг

Сравнение средних ошибок прогноза, полученного методом экспоненциального сглаживания (табл.2), и прогноза, полученного методом Хольта-Уинтерса (табл.4), показывает, что использование метода Хольта-Уинтерса для прогнозирования общей минерализации пресной воды позволяет получить значительно более точные результаты (погрешность прогноза уменьшается практически в 2 раза).

Таким образом, из проведенных исследований можно сделать вывод о приемлемости метода Хольта-Уинтерса для построения краткосрочных (на месяц вперед) значений общей минерализации пресной воды.

Таблица 4 — Средние ежегодные ошибки прогноза общей минерализации методом Хольта-Уинтерса за период 2004–2008г.

№ п/п	Год	Средняя погрешность прогноза, %
1	2004	4,41
2	2005	2,63
3	2006	5,22
4	2007	3,07
8	2008	2,11

#### Выводы:

1. Результаты исследований показали, что, в отличие от простого экспоненциального сглаживания, применение метода Хольта-Уинтерса для построения краткосрочных прогнозов (с учетом сезонных эффектов) значений общей минерализации пресной воды дает значительно более точные результаты прогнозирования. Точность прогноза возрастает практически в 2 раза, погрешность прогноза в этом случае составляет около 5%.

2. Использование более коротких горизонтов прогнозирования (неделя, день, несколько часов) для построения прогнозов общей минерализации позволит повысить вероятность предупреждения аварийных ситуаций, связанных с выбросами вредных веществ в воду рек, каналов и др. источников пресной воды, и повысить оперативность реагирования, в случае возникновения таковых ситуаций.

#### Литература

1. Экологическое прогнозирование. (Функциональные предикторы временных рядов) / Г.С. Розенберг, В.К.Шитиков, П.М. Брусилковский. — Тольятти: ИЭВБ РАН, 1994. — 182 с.
2. Вайну, Я.Я.-Ф. Корреляция рядов динамики / Я.Я.-Ф. Вайну. — М.: Статистика, 1977. — 119 с.
3. Построение динамических стохастических моделей по экспериментальным данным / Р.Л. Кашьяр, А.Р. Рао. — М.: Наука, 1983. — 384 с.
4. Тихонов, Э.Е. Методы прогнозирования в условиях рынка: учебное пособие / Э.Е. Тихонов. — Невинномысск, 2006 — 221 с.
5. Ивахненко, А.Г. Долгосрочное прогнозирование и управление сложными системами / А.Г. Ивахненко. — Киев: Наукова думка, 1975. — 340 с.
6. Holt C.C. (1957) "Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages", Carnegie Inst. Tech. Res. Mem., 52.
7. Winters P.R. (1960) "Forecasting Sales by Exponentially Weighted Moving Averages", Mgmt. Sci., 6, 324.

Здано в редакцію:  
12.03.2009р.

Рекомендовано до друку:  
д.т.н, проф. Зорі А.А.