

СТАТИСТИЧЕСКИЙ АНАЛИЗ И МОДЕЛИРОВАНИЕ ИЗМЕНЧИВОСТИ КАЧЕСТВА СТОЧНЫХ ВОД В СИСТЕМЕ ПРОИЗВОДСТВЕННОГО ВОДООТВЕДЕНИЯ

Е.Д. Копнова,
Л.А. Родионова

Представлены результаты исследования по решению задачи экономико-статистического обоснования совершенствования водно-экологического менеджмента промышленного предприятия. В качестве основного инструментария в работе использовалась методика моделирования временных рядов с помощью стационарных случайных процессов. Базовыми моделями являлись модели интегрированных процессов авторегрессии и скользящего среднего с учетом сезонности. Для отражения длинной памяти во временных рядах показателей тестировались модели дробно интегрированных процессов, а также модели авторегрессии с условной гетероскедастичностью. Анализ динамических связей производился на основе модели векторной авторегрессии.

Обнаружено, что для всех анализируемых показателей загрязнения, наряду с явным отсутствием тренда среднего уровня, имеет место большая вариабельность значений, проявляющаяся как в годовой, так и несезонной цикличности. Выявлены долговременные связи между отдельными показателями, в частности для большинства из них затухание влияния единичного сверхнормативного сброса любого другого показателя длилось не менее года.

Продемонстрированная в работе адекватность применяемого эконометрического инструментария определила возможность достоверного прогнозирования качества сточных вод, а также оптимизации превентивных мер по предотвращению сверхнормативных сбросов. Оценка характера периодичности контролируемых показателей с учетом сезонности, синергетического эффекта загрязнения указала на возможность повышения эффективности процесса водоочистки с выбором оптимального режима затрат. Выявление инерционности временных рядов отдельных показателей свидетельствует о возможной агрегации загрязнения из различных источников сбросов и, как следствие, о необходимости усиления контроля над сбросами сточных вод как по каждому антропогенному источнику, так и по фоновому природному загрязнению. Определение динамических связей между отдельными загрязнителями обосновало возможность повышения очищающей способности пруда с учетом структуры и длительности этих связей.

Ключевые слова: статистический анализ, качество сточных вод, временные ряды, многомерный анализ, кластеризация волатильности, длинная память, показатель Херста, причинность по Гранжеру.

JEL: C22, C32, Q53, Q57.

Сточные воды промышленных предприятий содержат большое количество загрязняющих веществ, которые наносят значительный ущерб принимающим их водоемам. В первую очередь они попадают в буферные пруды¹, которые, с одной стороны, играют роль очистителей промышленных сточных вод, а с другой - являются полноценными водными объектами, со своей флорой и фауной и взаимосвязями с подземными и грунтовыми водами региона. Нарушение экологических нормативов² при использовании буферного пруда ведет к нарушению его гидрологических характеристик, деградации видового биоразнообразия и в конечном итоге представ-

ляет потенциальную опасность для окружающих экосистем.

Уникальность буферного пруда как экосистемы обусловлена еще и тем, что он обычно принимает сточные воды от различных предприятий, так что его загрязнение формируется в условиях, приближенных к неопределенности. Для управления такой экосистемой необходимо наличие развитой системы мониторинга, способной обеспечивать возможность прогнозирования ее состояния и обоснованно планировать превентивные мероприятия по предотвращению сверхнормативных сбросов. На практике же оценку загрязнения буферного пруда дает отдельное предприятие,

Копнова Елена Дмитриевна (ekopnova@hse.ru) - канд. техн. наук, доцент департамента статистики и анализа данных, Национальный исследовательский университет «Высшая школа экономики» (г. Москва, Россия).

Родионова Лилия Анатольевна (lrodionova@hse.ru) - канд. экон. наук, доцент департамента статистики и анализа данных, Национальный исследовательский университет «Высшая школа экономики» (г. Москва, Россия).

¹ Буферный пруд - это гидросооружение для усреднения состава и отстаивания промышленных сточных вод в процессе их очистки.

² СанПиН 2.1.5.980-00. Водоотведение населенных мест, санитарная охрана водных объектов. Гигиенические требования к охране поверхностных вод.

отчитывающееся по форме № 2ТП³, заполняя ее лишь на основе средних значений локально измеренных концентраций без корректного учета их возможных вариаций и характера динамики.

Задача совершенствования методологии моделирования биогеоценоза с использованием математико-статистического инструментария и информационно-вычислительного обеспечения сформулирована на международном уровне в официальных документах по экологической политике⁴. Она также декларируется ведущими российскими специалистами в сфере водопользования [4, с. 311]. Однако в настоящее время значительные успехи при ее решении достигнуты лишь в разработке методов и моделей состояния гидробиоты [9], лежащих в основе метода биоиндикации качества воды [5]. При этом следует отметить, что наряду с несомненными преимуществами этих результатов, специфика существенной многомерности данных, включающих характеристики свыше 200 видов отдельных популяций, не позволяет на практике использовать классические методы анализа временных рядов. Кроме того, требование ведения обязательной статистической отчетности по водопользованию диктует создание методик, опирающихся на включенные в нее данные гидрохимических показателей, тем более что современная практика внедрения аппаратных средств в процесс измерения показателей в режиме онлайн [8, 12] создает реальную возможность расширения информационной базы мониторинга. Поэтому так актуальна задача совершенствования существующей системы экологического учета водоотведения в части расширения статистической методологии анализа временных рядов.

Цель работы состояла в развитии статистической методики исследования качества воды в буферном пруду по данным наблюдений за контролируемыми загрязнителями. К задачам исследования относилось изучение характера изменчивости отдельных показателей сбросов, а также динамических связей между ними с использованием моделей стационарных случайных процессов.

Описание и предварительный анализ данных

Исследование проводилось на примере буферного пруда, входящего в систему водоотведения на кемеровском заводе «Азот». В составе принимаемых им сбросов примерно третью часть составляют частично очищенные сточные воды с завода «Азот», не менее 65% - с городской водопроводной сети и 5% - с теплоэлектростанции. Пройдя очистку в пруду через осаждение и разбавление, стоки сливаются в реку Томь. По данным кемеровского филиала ФГБУ «Верхнеобьрыбвод», этот пруд - один из самых загрязненных в регионе водоемов с высоким содержанием мутирующей ихтиофауны.

Кемеровский завод «Азот» является крупнейшим предприятием химической отрасли, широко известным своей продукцией как в России, так и за рубежом. Внушительному объему его производства соответствует и масштабность природоохранных проектов. Так, в мае 2014 г. на выходе из буферного пруда в реку Томь была введена в эксплуатацию установка ультрафиолетовой очистки сточных вод стоимостью 223,2 млн рублей. В связи с вводом в действие этого дорогостоящего очистного сооружения совершенствование системы мониторинга качества сточных вод в буферном пруду позволит решить еще и задачу сохранения эколого-экономического баланса.

В исследовании использовались ежемесячные данные федеральной статистической отчетности по качеству воды в буферном пруду, представленные экологической службой завода «Азот» по форме № 2ТП (водхоз) за 2001-2011 гг. В качестве изучаемых показателей были взяты наиболее полно отображенные в отчетах данные о концентрации контролируемых в пруду гидрохимических элементов (в мг/куб. дм). В их перечень (см. таблицу 1) вошли такие интегральные показатели качества воды, как ХПК (химическое потребление кислорода) и БПК (биохимическое потребление кислорода), характеризующие содержание кислорода в воде и отражающие результат реакции экосистемы на сбросы⁵, а также взвеси и сухой остаток, характеризующие механическую состав-

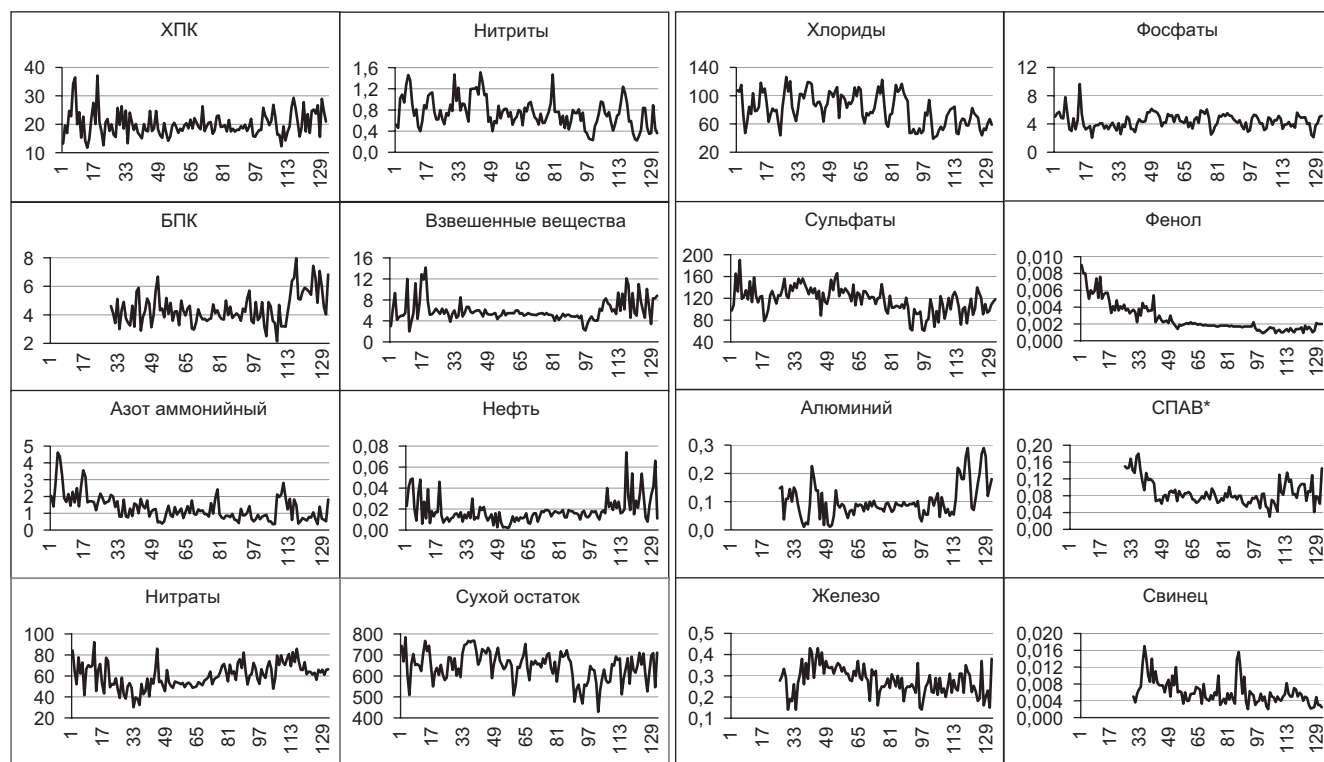
³ Приказ Росстата от 19.10.2009 № 230 «Об утверждении статистического инструментария для организации Росводресурсами федерального статистического наблюдения об использовании воды».

⁴ Establishing a dynamic system of surface water quality regulation: Guidance for countries of Eastern Europe, Caucasus and Central Asia. OECD, March, 2011 (draft); Guidelines for Developing National Strategies to Use Air and Water Quality Monitoring as Environmental Policy Tools Eastern Europe, the Caucasus, Central Asia and South-Eastern Europe / United Nations, Economic Commission For Europe, New York and Geneva, 2012.

⁵ ГОСТ Р 52708-2007 Метод определения химического потребления кислорода.

ляющую в составе загрязнения. Из химических элементов в выборку вошли загрязнители трех классов опасности (2, 3, 4), представляющие как

органическую, так и неорганическую компоненту стоков. Графики временных рядов этих показателей приведены на рис. 1.



* Синтетические поверхностно-активные вещества.

Примечание: горизонтальная ось - временной период в месяцах; вертикальная ось - концентрация загрязнителей.

Рис. 1. Динамика показателей качества воды (концентрация в мг/куб. дм) в период с января 2001 по ноябрь 2011 г.

Визуальный анализ графиков показывает, что наряду с незначительностью тренда, имеет место существенная вариабельность значений, проявляющаяся как в годовой (хлориды, свинец), так и во внегодовой цикличности (нитраты, сульфаты). Кроме того, для некоторых показателей заметно наличие структурных сдвигов (БПК, фосфаты, СПАВ) и кластеризация волатильности (взвешенные вещества, нефтепродукты, алюминий), характеризующаяся наличием отдельных перио-

дов с относительно низкой и высокой волатильностью.

Для статистического анализа данные наблюдений были приведены к сопоставимому виду путем соотнесения их с предельно допустимыми концентрациями (ПДК)⁶.

Для корректной спецификации модели было проведено предварительное исследование каждого временного ряда. Его основные результаты представлены в таблице 1.

Таблица 1

Описательные статистики показателей
(в единицах ПДК)*

	Класс опасности	Среднее значение	Максимум	СКО**	Асимметрия	Куртозис	J-B-проб***
ХПК	-	0,67	1,24	0,15	1,19	5,43	0,00
БПК	-	0,52	1,18	0,31	0,83	3,65	0,00
Азот	4	0,88	3,08	0,51	1,60	6,66	0,00
Нитраты	3	1,34	2,05	0,26	0,11	2,82	0,80
Нитриты	2	0,23	0,46	0,09	0,53	3,07	0,06
Взвеси	-	0,47	1,12	0,16	1,66	6,16	0,00

⁶ ГН 2.1.5.689-98. Предельно допустимые концентрации (ПДК) химических веществ в воде водных объектов хозяйственно-питьевого и культурно-бытового водопользования.

	Класс опасности	Среднее значение	Максимум	СКО**	Асимметрия	Куртозис	J-V-prob***
Нефть	4	0,06	0,25	0,04	1,96	7,52	0,00
Сухой остаток	-	0,80	0,98	0,09	-0,51	3,12	0,07
Хлориды	4	0,23	0,36	0,06	0,03	1,98	0,07
Сульфаты	4	0,23	0,38	0,05	-0,09	3,26	0,73
Алюминий	3	0,42	1,45	0,33	1,24	4,53	0,00
Железо	3	0,75	1,43	0,41	-0,02	2,60	0,69
Фосфаты	3	3,86	8,47	0,92	0,97	7,13	0,00
Фенол	4	2,64	9,00	1,69	1,62	5,21	0,00
СПАВ	4	0,14	0,36	0,09	1,05	3,90	0,00
Свинец	2	0,48	1,70	0,36	1,47	5,22	0,00

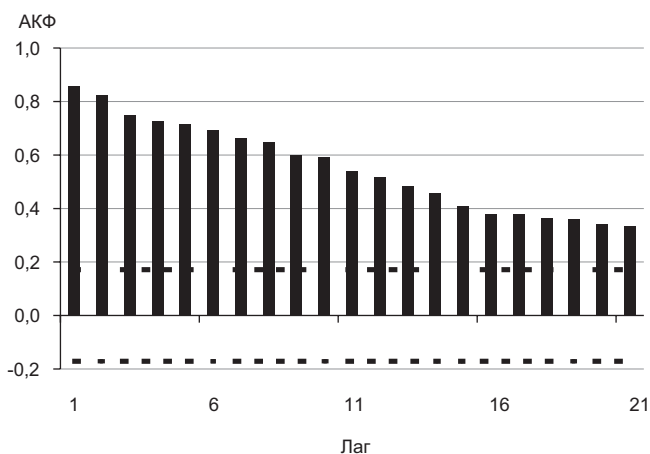
* Жирным шрифтом в этой таблице и таблице 2 выделены проблемные значения, комментируемые в тексте.

** Среднее квадратическое отклонение.

*** Обозначение J-V-prob означает вероятность превышения тестовой статистики J-V расчетного уровня. Аналогичное обозначение используется далее в тексте для других тестов.

В частности, данные таблицы 1 указывают на то, что даже средние значения отдельных показателей существенно превышают ПДК, достаточно велики максимумы, имеет место правосторонняя асимметрия. Также заметны повышенные относительно нормального распределения значения куртозиса, которые указывают на возможность «толстых» хвостов распределения (выбросов), кластеризацию волатильности, а также на длинную память временных рядов [17, р. 304].

По каждому ряду были построены графики оценок автокорреляционной функции (АКФ) и спектральной плотности (СП). Они в основном соответствовали выводам о наличии тренда и сезонности в представленных рядах данных, а также о внегодовой цикличности.



Примечание: пунктиром показаны границы критической области значений АКФ, значимых на уровне 0,05.

Рис. 2. Коррелограмма для концентрации фенола

Характерно, что для большинства рядов убывание выборочной АКФ происходило относительно медленно (см. рис. 2), что свидетельствует о наличии в них длинной памяти, инерционности процессов загрязнения. Пики на низких частотах графиков СП свидетельствовали о наличии несезонной цикличности (см. рис. 3), указывающей на способность экосистемы продолжительно аккумулировать загрязнение.

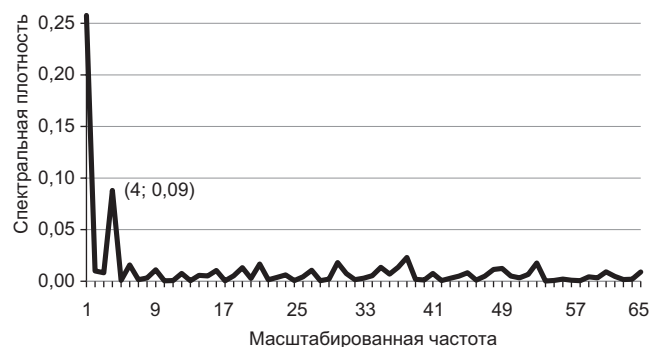


Рис. 3. Периодограмма для концентрации нитратов

Для каждого временного ряда изучалась проблема его стационарности [2]. Использовался расширенный ADF-тест Дики-Фуллера, PP-тест Филлипса-Перрона, а также KPSS-тест (тест Квятковского-Филлипса-Шмидта-Шина). Результаты тестирования подавляющего большинства рядов свидетельствовали об отсутствии стохастического тренда, однако некоторая противоречивость выводов тестов отдельных показателей указывала на то, что они могут иметь длинную,

⁷ Стационарный процесс является процессом с длинной памятью, если существуют такие числа α , с ($0 < \alpha < 1$, $c > 0$), что для его корреляционной функции ρ_k выполняется предельное соотношение: $\lim_{k \rightarrow \infty} \frac{\rho_k}{ck^{-\alpha}} = 1$.

но не бесконечную память⁷. Это подтверждало и лептокуртическое распределение вероятности с завышенным значением куртозиса, наличие несезонной цикличности, медленное убывание АКФ.

Для статистически корректного выявления длинной памяти использовался метод R/S -анализа, предложенный Херстом [13] и развитый Мандельбротом [15]. В качестве инструмента обнаружения длинной памяти во временных рядах применяется так называемая R/S -статистика, равная отношению размаха частичных сумм центрированных значений ряда к его стандартному отклонению. Идея R/S -анализа состоит в том, что в предположении полного отсутствия памяти размах ряда увеличивается пропорционально корню из величины временного промежутка, то есть $R/S = ct^H$, $H = 0,5$. В противном случае показатель степени H , называемый показателем Херста, имеет другое значение, которое и характеризует наличие памяти в ряду. При $0,5 < H < 1$ процесс является персистентным, то есть монотонно следует вдоль детерминированного тренда. При $0 < H < 0,5$ - говорят об антиперсистентности, то есть высокочастотной изменчивости значений относительно тренда. Для проверки гипотезы об отсутствии длинной памяти в ряду используется величина, рассчитываемая на основе R/S -статистики, предложенная А. Ло [14].

Для каждого анализируемого временного ряда были определены оценки показателя Херста, которые указали на то, что для всех рядов характерно свойство персистентности. При этом гипотеза об отсутствии длинной памяти отвергалась на уровне значимости 0,05 для подавляющего большинства рядов. Обнаруженный феномен длинной памяти отражает неоднородность воздействия некоторых водопользователей на данную экосистему, поскольку обычно возникает при агрегировании отдельных стационарных схожих процессов с различной лаговой структурой [6]. Он также характеризует особенность биоценоза физически существовать в условиях антропогенной нагрузки. В частности, известно, что происходит накопление загрязнения в донных отложениях, реагирует на сбросы, разрастаясь, гидробиота [3, 10, 11]. В результате происходит снижение эффективности функционирования буферного пруда как средства водоочистки сбросов. Все это указывает на необходимость обеспечения согласованности действий водопользователей и всестороннего учета биоразнообразия экосистемы.

Моделирование процессов загрязнения

Многофакторность антропогенного воздействия на качество воды вряд ли поддается детальному описанию. Однако остается возможность определить общие закономерности его формирования и использовать полученные выводы для управления. Рассматриваемые далее процедуры моделирования временных рядов были применены с целью статистически корректного обоснования характера их изменчивости, определения возможности прогнозирования, а также изучения динамических связей между отдельными показателями загрязнения.

В качестве базовой модели каждого процесса загрязнения была выбрана модель $ARIMA(p, d, q)$ (*Integrated Auto-regression and Moving Average*) [2], которая представляет собой линейную комбинацию, соответственно, p и q лаговых величин стационарного случайного процесса, полученного из исходного процесса Y_t взятием d разностей и ошибок ε_t^* :

$$\begin{aligned} \varphi(L)(1-L)^d Y_t &= \psi(L)\varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2), \quad d \in \mathbb{N}, \\ E[\varepsilon_t] &= 0, V[\varepsilon_t] = \sigma^2, \quad cov[\varepsilon_t, \varepsilon_{t-k}] = 0 \quad (k \neq 0)** \\ L^k Y_t &= Y_{t-k}, \quad \varphi(L) = 1 - \delta_1 L + \dots + \delta_p L^p, \\ \psi(L) &= 1 + \theta_1 L + \dots + \theta_q L^q. \end{aligned}$$

* Модель может также включать константу и детерминированный тренд.

** Последовательность $\{\varepsilon_t\}$ с указанными свойствами называется белым шумом (*White Noise*) и обозначается $\varepsilon_t \sim WN(0, \sigma^2)$.

Для отражения сезонности использовалась модификация этой модели $SARIMA(p, d, q)(D, P, Q)_s$, (*Seasonal ARIMA*):

$$\begin{aligned} \varphi_s(L)\varphi(L)(1-L)^d(1-L)^D Y_t &= \psi_s(L)\psi(L)\varepsilon_t \\ \varphi_s(L) &= 1 - \delta_s L + \dots + \delta_s P L^P, \\ \psi_s(L) &= 1 + \theta_s L + \dots + \theta_s Q L^Q. \end{aligned}$$

Параметры D, P, Q, s означают соответственно порядок сезонной разности, число сезонных лагов в ряду и в ошибках, период сезонности.

Модель $ARIMA$ применялась в сочетании с моделью условной гетероскедастичности $GARCH$ (*Generalized Auto-Regressive, Conditional Heteroscedastic*), поскольку графики как исходных рядов, так и остатков $ARIMA$ явно указывали на

кластеризацию волатильности, а значение куртосиса для них было достаточно велико. При этом гипотеза об отсутствии условной гетероскедастичности отвергалась на уровне значимости 0,05 для 12 из 16 рассматриваемых рядов.

Модель ARMA-GARCH рассматривалась в виде:

$$\varphi(L)Y_t = \psi(L)U_t, \quad U_t = \varepsilon_t \sigma_t, \quad \varepsilon_t \sim SWN(0,1)^*$$

$$\sigma_t^2 = Var[U_t | u_{t-1}, u_{t-2}, \dots, u_{t-q}] = \omega + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{i=1}^p \gamma_i \sigma_{t-i}^2,$$

$$\omega > 0, \quad \alpha_i \geq 0 \quad (i=1, 2, \dots, q), \quad \gamma_i \geq 0 \quad (i=1, 2, \dots, p).$$

* $\varepsilon_t \sim SWN(0, \sigma)$ означает, что ε_t - так называемый *сильный белый шум (Strong White Noise)*, который отличается от белого шума тем, что составляющие его величины независимы, а не некоррелированы.

Тестировалась также и модель дробно интегрированных процессов ARFIMA(p, d, q) (*Fractionally Integrated Autoregression and Moving Average*) [6], модификация модели ARIMA для стационарных процессов с длинной памятью, в которой параметр d может принимать дробные значения:

$$\varphi(L)(1-L)^d Y_t = \psi(L)\varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2), \quad d > -1.$$

При этом параметр d и показатель Херста H связаны соотношением: $H = d + 0,5$.

Результаты идентификации модели для всех рядов показали значимое на уровне 0,05 значение параметра d, которое в основном соответствовало оценке показателя Херста, полученной в R/S-анализе. Это позволило статистически корректно подтвердить выводы предварительного анализа о наличии долговременной памяти в рассматриваемых процессах и характере их изменчивости.

Оптимизация параметров моделей и выбор наиболее адекватной из них осуществлялись на основе минимизации информационных критериев [2] и результатов тестирования остатков на соответствие процессу белого шума. Все модели оценивались с константой, лишь для двух рядов в модель был включен тренд. В таблице 2 приведены результаты выбора моделей с указанием их параметров и значений вероятностей (χ^2 -prob), соответствующих оценкам тестовой статистики Жарка-Бера. Видно, что лишь для двух рассматриваемых рядов гипотеза о том, что остатки распределены по нормальному закону, отвергается на уровне значимости 0,01.

Параметры модели в таблице 2 представлены в форме перечисления включаемых в нее лагов. Например, модель для концентрации азота $Trend(1,4,7;0;10)(1;0;1)_{12}$ имеет вид:

$$(1 - \delta_{s1} L^{12})(1 - \delta_1 L - \delta_4 L^4 - \delta_7 L^7) Y_t = \beta_0 + \beta_1 t + (1 + \theta_{s1} L^{12})(1 + \theta_{10} L^{10}) \varepsilon_t.$$

Из таблицы 2 видно, что для большинства моделей число лагов достаточно велико, что соответствует выявленному феномену длинной памяти. Заметим также, что для двух рядов с противоречивыми выводами о стационарности оптимальными моделями оказались модели в разностях (взвеси, фенол), еще одна из таких моделей построена для сезонной разности (хлориды). Для большинства рядов были определены GARCH-модели, отражающие длинную память ряда о вариациях. Значимость этих моделей указала на существование наиболее критических и наиболее спокойных периодов биоценоза. Поэтому результаты мониторинга качества сбросов могли бы быть полезны в общей системе водно-экологического менеджмента.

Построение адекватной модели для каждого процесса позволило не только описать его поведение с учетом его предыстории, но и построить достаточно достоверный динамический прогноз. На рис. 4 приведен пример результатов прогнозирования по модели в сопоставлении с наблюдаемыми значениями в 2011 г. для концентрации хлоридов. Видно, что наблюдаемые значения не выходят из 95%-ного доверительного интервала прогнозных оценок в течение всего прогнозного периода и особенно близки к ним первые полгода.

Таблица 2

Результаты выбора моделей и тестирования на причинность по Гранжеру

Переменная	ARIMA, SARIMA	GARCH		J-B	G
		q	p		
ХПК	(10;0;18)(1,0,1) ₁₂	1	1	0,52	0,00
БПК	(1;0;1)	1	0	0,40	0,12
Азот	<i>Trend</i> (1,4,7;0;10)(1;0;1) ₁₂	-	-	0,06	0,00
Нитраты	(2;0;1,2,4,12)	1	1	0,99	0,05
Нитриты	(0;0;1,2)(1;0;1) ₁₂	-	-	0,18	0,07
Взвеси	(1,2,3,33;1;2,3)	0	1	0,68	0,01
Нефть	(2,8;0;1,2,6)	0	1	0,00	0,00
Сухой остаток	(1,6;0;12)	1	1	0,38	0,01
Хлориды	(1;0;0)(1;1;1)₁₂	1	1	0,35	0,49

Окончание таблицы 2

Переменная	ARIMA, SARIMA	GARCH		J-B	G
		q	p	χ^2 -prob	χ^2 -prob
Сульфаты	Trend (2;0;1)	-	-	0,49	0,04
Алюминий	(2;0;1)	1	1	0,60	0,00
Железо	(1;0;1,7)	-	-	0,07	0,47
Фосфаты	(1,20;0;0)	1	1	0,81	0,05
Фенол	(1,7,12;1;1)(0;0;1)₁₂	1	1	0,00	0,00
СПАВ	(0;0;1)(0;0;1) ₁₂	0	1	0,41	0,00
Свинец	(1,3;0;2,11,13,18)	1	0	0,37	0,74

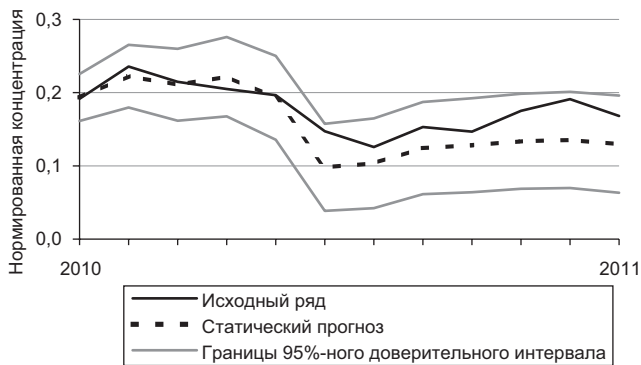


Рис. 4. Модельный прогноз концентрации хлоридов на год

Для выявления динамических связей рядов использовалось понятие причинности⁸ по Гранжеру, которое для двух переменных означает, что изменение одной из них предшествует изменению другой. Для многомерного случая для каждого временного ряда в качестве потенциальной причины выступает вектор остальных переменных. Идея анализа состоит в сравнении длинной и короткой авторегрессии, отличающихся числом лагов компонент этого вектора в рамках модели векторной авторегрессии - VAR (*Vector autoregression*) [2].

Для каждого j -го ряда ($j = 1, 2, \dots, 16$) в VAR рассматривалось уравнение, включающее лаги всех переменных:

$$Y_{jt} = \alpha_0 + \sum_{k=1}^p \alpha_j Y_{jt-k} + \sum_{i=1}^{16} \sum_{k=1}^p \beta_{ik} Y_{it-k} + \varepsilon_{jt},$$

$i \neq j.$

Число лагов p в VAR было выбрано при минимальном значении информационного критерия AIC (*Akaike information criterion*) и равнялось трем

($p = 3$). Тестировалась гипотеза о том, что коэффициенты при лагах всех других переменных равны нулю:

$$H_0 : \beta_{ik} = 0, \quad i = 1, 2, \dots, 16 \quad (i \neq j), \quad k = 1, 2, 3.$$

Если гипотеза отвергалась, то делался вывод о том, что совокупность других рядов является причиной по Гранжеру для данного ряда.

Результаты тестирования приводятся в таблице 2 (графа G) в форме значений вероятности (χ^2 -prob), соответствующей оценке тестовой статистики. Из таблицы видно, что на уровне значимости 0,01 подтверждается естественная причинная связь общей совокупности рассматриваемых выбросов на интегральный индикатор качества воды (ХПК), а также и на общие загрязнители (сухой остаток, взвеси). Примечательно, что на уровне значимости 0,1 не отвергается гипотеза о том, что совокупное загрязнение способствует накоплению в экосистеме всех соединений азота, в том числе особо опасных нитритов, сбросы которых осуществляет основной водопользователь. Данный вывод указывает на то, что для поддержания полноценного функционирования буферного пруда в системе водоотведения необходимо проведение превентивных мероприятий для его комплексной очистки.

Адекватность построенной модели VAR, в частности ее стабильность и отсутствие автокорреляции остатков (на уровне 0,05), позволила не только провести исследование причинности по Гранжеру для отдельных пар показателей и по совокупности данных, но и обосновать долговременность этих связей. Для этого использовались функции откликов одних загрязнителей на импульсы других - IRF (*Impulse response function*). Для каждой переменной y_j , входящей в VAR, IRF характеризует изменение y_i в течение определенного периода времени при воздействии единичного изменения ошибки ε_i (шоке) в начале этого периода. Для большинства рассматриваемых показателей затухание IRF наблюдалось не ранее чем через год. Для примера на рис. 5 приведен график функции отклика концентрации нитритов на единичный шок сбросов азота⁹. Кривыми серого цвета указаны границы 95%-ного асимптотического доверительного интервала.

⁸ Предполагалось, что все ряды стационарны.

⁹ Единичный импульс рассматривался в форме стандартного отклонения ошибки авторегрессии концентрации азота в модели VAR с учетом декомпозиции Холески матрицы ковариации остатков [16].

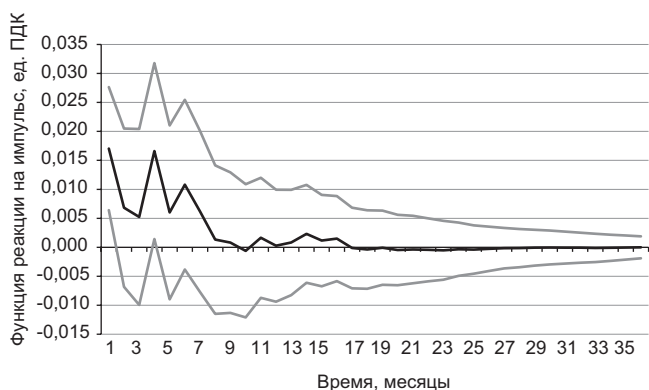


Рис. 5. Реакция концентрации нитритов на мгновенное повышение сбросов азота

Из рис. 5 видно, что время затухания реакции составляет более двух лет. Поэтому экологической службе на заводе при возникновении единичных случаев превышения нормативов при сбросе азота следует предусмотреть и возможное долгосрочное превышение концентрации опасных нитритов.

* *
*

Применение методики статистического анализа временных рядов к данным наблюдений за сбросами в буферном пруду завода «Азот» дало возможность решить следующие задачи.

Проанализирован характер изменчивости отдельных загрязнителей. Построены модели, позволяющие адекватно описывать процессы загрязнения и прогнозировать развитие процессов загрязнения в буферном пруду. Показано, что лишь некоторым из них присуща сезонность. Определены показатели с выраженной несезонной цикличностью. Выявлены показатели с кластеризацией волатильности. Выявление инерционности процессов загрязнения водоема отдельными веществами, наличие в них длинной памяти указало на возможную агрегацию из различных источников и, как следствие, на необходимость усиления контроля за сбросами сточных вод как по каждому антропогенному источнику, так и по фоновому природному загрязнению.

Были сделаны выводы о совершенствовании системы мониторинга качества сточных вод в части учета характера изменчивости отдельных загрязнителей. Ясно, что при наличии высокой волатильности измерять показатель и контролировать процесс очистки надо чаще с учетом существенного превышения среднего уровня. При

наличии кластеризации волатильности необходимо регулирование процесса очистки с выбором оптимального режима затрат. Учет периодичности сбросов сточных вод, несомненно, также должен способствовать эффективности проведения превентивных мер как по предотвращению высоких концентраций сбросов, так и экономии средств при ожидаемых низких уровнях загрязнения. Контроль динамических связей отдельных загрязнителей позволит обоснованно повысить функциональную способность буферного пруда в системе водоотведения.

Следует также отметить, что внедрение продемонстрированных в работе статистических методов анализа временных рядов будет более эффективно при совершенствовании информационной базы, в частности перехода к измерениям качества воды в реальном режиме времени и расширении системы контролируемых показателей.

Литература

1. Айвазян С.А. Прикладная статистика. Основы эконометрики: Учебник для вузов: В 2 т. 2-е изд., испр. - Т. 1: Айвазян С.А., Мхитарян В.С. Теория вероятностей и прикладная статистика. - М.: ЮНИТИ-ДАНА, 2001. - 656 с.
2. Вербик М. Путеводитель по современной эконометрике. Пер. с англ. В.А. Банникова. Науч. ред. и предисл. С.А. Айвазяна. - М.: Научная книга, 2008. - 616 с.
3. Гелашвили Д.Б., Якимов В.Н., Иудин Д.И. и др. Фрактальные аспекты таксономического разнообразия // Журнал общей биологии. 2010. Т. 71. № 2. С. 115-130.
4. Данилов-Данильян В.И., Пряжинская В.Г., Храмович И.Л. и др. Экономические и территориальные аспекты управления водохозяйственным комплексом России / Под ред. В.И. Данилова-Данильяна и В.Г. Пряжинской. - М.: РАСХН, 2013.
5. Левич А.П., Булгаков Н.Г. и др. Экологический контроль окружающей среды по данным биологического и физико-химического мониторинга природных объектов // Компьютерные исследования и моделирование. 2010. Т. 2. № 2. С. 199-207.
6. Обрезков О. Долгосрочная связь временных рядов и паритет покупательной способности // Квантиль. 2007. № 2. С. 131-140.
7. Росси Э. Одномерные GARCH-модели: обзор // Квантиль. 2010. № 8. С. 1-67.
8. Семенова И.В., Булгаков В.Г. и др. Автоматические станции контроля качества воды: первые результаты опытной эксплуатации // Фундаментальные проблемы воды и водных ресурсов: Труды IV Меж-

дународной научной конференции. - М.: ИВП РАН, 2015. С. 203-205.

9. **Шитиков В.К., Зинченко Т.Д., Розенберг Г.С.** Макроэкология речных сообществ: концепции, методы, модели. - Тольятти: Кассандра, 2011. - 255 с.

10. **Austin M.P.** Species distribution models and ecological theory: a critical assessment and some possible new approaches // *Ecological Modelling*. 2007. Vol. 200. P. 1-19.

11. **Azaele S., Pigolotti S., Banavar J.R., Maritan A.** Dynamical evolution of ecosystems // *Nature*. 2006. Vol. 444. P. 926-928.

12. **Dongxian He, Weifen Du, Juanxiu Hu.** Water quality dynamic monitoring technology and application based on ion selective electrodes // *Recent developments in mobile communications - A multidisciplinary approach*. Edited by

Dr Juan P. Maicas. Rijeka, Croatia: Publisher InTech, 2011. P. 251-272.

13. **Hurst H.E.** Long-term storage capacity of reservoirs // *Transactions of the American Society of Civil Engineers*. 1951. Vol. 116. P. 770-799.

14. **Lo A.W.** Long-term memory in stock market prices // *Econometrica*. 1991. Vol. 59. No. 5. P. 1279-1313.

15. **Mandelbrot B.** Statistical methodology for non-periodic cycles: From the covariance to R/S analysis // *Annals of Economic Social Measurement*. 1972. Vol. 1. No. 3. P. 259-290.

16. **Mills T.** The econometric modeling of financial time series. Cambridge University Press, 2008. - 456 p.

17. **Peters E.** Fractal market analysis: Applying chaos theory to investment and economics. New York, J. Wiley & Sons, 1994.

STATISTICAL ANALYSIS AND MODELING OF THE VARIABILITY OF WASTEWATER QUALITY IN THE SYSTEM OF INDUSTRIAL WASTEWATER

Elena D. Kopnova

Author affiliation: National Research University Higher School of Economics (Moscow, Russia). E-mail: ekopnova@hse.ru.

Lilia A. Rodionova

Author affiliation: National Research University Higher School of Economics (Moscow, Russia). E-mail: lrodionova@hse.ru.

This article presents results of the study on economic and statistical justification for improvement of water and environmental management of an industrial enterprise. As a main tool the authors applied - was the method for modeling time series using stationary stochastic processes. The models of the integrated auto-regression and moving average, seasonally adjusted were used as the base. The models of fractionally integrated processes and models of autoregressive conditional heteroskedasticity were tested to reflect the long memory in time series of indicators. Analysis of dynamic links was based on vector autoregression model.

The authors demonstrated that for all the analyzed indicators of pollution, along with the apparent lack of mid-level trend, there is a considerable variability of values, which manifested in both annual and non-seasonal cyclical and structural changes. The longstanding interrelations between the individual indicators were revealed - for most of them the damping effect of a single excess discharge of any other indicator lasted for at least a year.

The article proves sufficiency of the applied econometric tools which have determined the possibility for reliable forecasting the wastewater quality along with optimization of the measures for preventing excessive discharges. Identifying the character of the periodicity of the discharges with account to seasonality, as well as the synergistic effect of contamination indicated the possibility of increasing the efficiency of water treatment process by selecting the optimum costs. The identification of the inertia of the processes of pollution of individual indicators, testified to their possible aggregation from different sources to the necessity of strengthening of control over wastewater discharges for each anthropogenic source and the natural background contamination. Determining the dynamic interrelations between the individual polluters justified a reasonable opportunity to improve the pool cleanability with regard to the structure and duration of those relations.

Keywords: statistical analysis, quality of wastewater, time series, multidimensional analysis, volatility clustering, long memory, Hurst exponent, Granger causality.

JEL: C22, C32, Q53, Q57.

References

1. **Ayvazyan S.A.** [Applied statistics. The basics of econometrics: the Textbook for higher education institutions in 2 vol., 2nd ed., rev. Vol. 1] In: Ayvazyan S.A., Mkhitarian V.S. *Teoriya veroyatnostey i prikladnaya statistika* [Probability theory and applied statistics]. Moscow, Unity-Dana Publ., 2001. 656 p. (In Russ.).

2. **Verbeek M.** A guide to modern econometrics, 2nd ed. John Wiley & Sons, Ltd., 2004. 429 p. (Russ. ed.: Verbik M. *Putevoditel po sovremennoj ekonometrike*. Moscow, Science book Publ., 2008. 616 p.).

3. **Gelashvili D.B., Yakimov, V.N., Iudin D.I., et al.** Fraktalnye aspekty taksonomicheskogo raznoobraziya [Fractal aspects of the taxic diversity]. *Zhurnal Obshchei Biologii (Journal of General Biology)*, 2010, vol. 71, no. 2, pp. 115-130. (In Russ.).

4. **Danilov-Danilyan V.I., Pryazhinskaya V.G., Khranovich I.L., et al.** *Ekonomicheskiye i territorialnyye aspekty upravleniya vodokhozyaystvennym kompleksom Rossii* [Economic and territorial aspects of water management complex of Russia]. Moscow, RAAS Publ., 2013. (In Russ.).
5. **Levich A.P., Bulgakov N.G., et al.** Ekologicheskiy kontrol okruzhayushchey sredy po dannym biologicheskogo i fiziko-khimicheskogo monitoringa prirodnykh obyektov [Biological and physico-chemical data from natural objects for ecological environmental monitoring]. *Computer Research and Modeling*, 2010, vol. 2, no. 2, pp. 199-207. (In Russ.).
6. **Obrezkov O.** Dolgosrochnaya svyaz vremennykh ryadov i paritet pokupatelnoy sposobnosti [Long range dependence and the purchasing power parity]. *Quantile*, 2007, no. 2, pp. 131-140. (In Russ.).
7. **Rossi E.** Odnomernyye GARCH-modeli: obzor [Univariate GARCH models: A survey]. *Quantile*, 2010, no. 8, pp. 1-67. (In Russ.).
8. **Semenova I.V., Bulgakov V.G., et al.** [Automatic station water quality monitoring: first results of pilot operation]. *Trudy IV Mezhdunarodnoy nauchnoy konferentsii «Fundamentalnyye problemy vody i vodnykh resursov»* [Proc. IV Int. Sci. Conf. «Fundamental problems of water and water resources»]. Moscow, WPI RAS Publ., 2015, pp. 203-205. (In Russ.).
9. **Shitikov V.K., Zinchenko T.D., Rozenberg G.S.** *Makroekologiya rechnykh soobshchestv: kontseptsii, metody, modeli* [Macroecology river communities: concepts, methods, models]. Tolyatti, Cassandra Publ., 2011. 255 p. (In Russ.).
10. **Austin M.P.** Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*, 2007, vol. 200, pp. 1-19.
11. **Azaele S., Pigolotti S., Banavar J.R., Maritan A.** Dynamical evolution of ecosystems. *Nature*, 2006, vol. 444, pp. 926-928.
12. **Dongxian He, Weifen Du, Juanxiu Hu.** Water quality dynamic monitoring technology and application based on ion selective electrodes. In: Juan P. Maicas (ed.) *Recent developments in mobile communications - A multidisciplinary approach*. Rijeka, Croatia, InTech Publ., 2011. P. 251-272.
13. **Hurst H.E.** Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 1951, vol. 116, pp. 770-799.
14. **Lo A.W.** Long-term memory in stock market prices. *Econometrica*, 1991, vol. 59, no. 5, pp. 1279-1313.
15. **Mandelbrot B.** Statistical methodology for non-periodic cycles: From the covariance to R/S analysis. *Annals of Economic Social Measurement*, 1972, vol. 1, no. 3, pp. 259-290.
16. **Mills T.** *The econometric modeling of financial time series*. Cambridge University Press, 2008. 456 p.
17. **Peters E.** *Fractal market analysis: Applying chaos theory to investment and economics*. New York, J. Wiley & Sons, 1994.