

**УДК 556.048**

## **О статистическом прогнозировании характеристик речного стока**

**И.А. Старков, Р.А. Рангулов**

**Пермский государственный университет**

Статистическое прогнозирование является динамично развивающимся направлением прикладной статистики. В тех областях знаний, где анализируемой информацией являются временные ряды, методы статистического моделирования и прогнозирования позволяют получать наиболее вероятные сценарии развития случайных процессов. По этой причине они широко применяются в научных исследованиях и практической деятельности, в частности, в экономической сфере для прогноза курсов акций, валют, цен на драгоценные металлы.

Среди разнообразных методов прогнозирования особенно важное значение имеют адаптивные методы (модели). Их отличие от других прогностических моделей состоит в том, что они, отражая текущие свойства ряда, способны непрерывно учитывать эволюцию динамических характеристик изучаемых статистических (стохастических) процессов, вызываемую изменяющимися во времени условиями. Адаптивные модели предназначаются прежде всего для краткосрочного прогнозирования и при всей своей простоте могут давать более надежные результаты, чем сложные эконометрические системы уравнений [3].

Как известно, стохастические процессы делят на стационарные и нестационарные. Отличительным свойством стационарной последовательности является неизменность во времени ее свойств, в частности, постоянство математического ожидания. Ковариация же между членами ряда (автоковариация) зависит лишь от разности их номеров (лага). Среди стохастических моделей стационарных временных рядов наиболее распространены модели авторегрессии (АР), скользящего среднего (СС) и смешанные (АРСС). Все они являются частными случаями общей линейной модели временного ряда и в практических целях обычно применяются с малым числом параметров (не более 2).

Многие временные ряды характеризуются нестационарностью в различных ее проявлениях, затрудняющей прогнозирование. Однако, как показывают исследования, большую часть таких рядов, в том числе и рядов с тренд-циклической, сезонной или нерегулярной составляющей, можно описывать моделью авторегрессии – проинтегрированного скользящего среднего (АРПСС), или моделью Бокса – Дженкинса, являющейся модифицированной формой модели АРСС. Эта модель достаточно гибкая и охватывает широкий класс как стационарных, так и нестационарных процессов [1].

В настоящей статье приведены результаты анализа временных рядов и прогноза некоторых характеристик водного режима (экстремальных и среднегодовых расходов воды, наивысших и наименьших уровней воды), выполненного по данным гидрометрических наблюдений за стоком в нескольких пунктах бассейна р.Камы с продолжительными рядами наблюдений (более 60 лет). Исследование временных рядов и статистическое прогнозирование процессов с использованием модели АПРСС, а также других стохастических моделей реализовано в специализированной системе STATISTICA, являющейся универсальной системой анализа случайных процессов. Задача заключалась в оценке возможности ее применения для прогноза гидрологических характеристик. Практические приемы и методы статистического прогнозирования в этой системе, а также его теоретические основы изложены в [2].

Задачи прогнозирования в системе STATISTICA сводятся к следующему:

- первичный анализ «исходных» данных,
- идентификация модели, подходящей для описания имеющегося временного ряда,
- оценивание параметров модели,
- исследование адекватности модели,
- прогноз.

Анализ «исходных» данных – это прежде всего их визуализация и определение некоторых числовых характеристик (параметров). Система STATISTICA содержит мощный графический и математический аппарат для такого анализа.

Идентификация модели заключается в определении количества ее параметров. В несезонной модели АПРСС, например, содержится три типа параметров:  $d$  – порядок разности,  $p$  – порядок авторегрессии,  $q$  – порядок скользящего среднего. В сезонную модель добавляются еще три типа: сезонный параметр авторегрессии –  $P_s$ , сезонная разность –  $d_s$ , сезонный параметр скользящего среднего –  $Q_s$ . Основным критерием идентификации модели является характер выборочной автокорреляционной и частной автокорреляционной функций ряда. В системе STATISTICA предусмотрен быстрый и удобный интерфейс для получения графиков и численных значений той и другой функции анализируемого ряда.

Численные значения неизвестных параметров идентифицированной модели АПРСС оцениваются методом итераций. При этом необходимо задать начальное значение, с которого начнется поиск. Точность оценок понижается с уменьшением длины ряда и увеличением количества параметров.

Адекватность применяемой модели определяется путем исследования остатков (ошибок прогноза), представляющих собой разности между прогнозными и фактическими значениями. В системе STATISTICA имеются специализированные средства анализа остатков, в частности, оценка и

графическое изображение их автокорреляционной и частной автокорреляционной функций.

Наконец, процедура прогноза, являющаяся конечным этапом исследования, осуществляется также быстро после указания системе количества прогнозируемых значений и величины доверительного интервала. При необходимости визуализации прогнозы добавляются к траектории исходного ряда.

В качестве примера на рис.1 приведены прогнозы годового стока р.Сылва у с.Подкаменное на 1981 и 1982 гг., выполненные методом АРПСС по временному ряду продолжительностью 100 лет (за 1881–1980 гг.). Стохастическая модель определена как авторегрессия первого порядка АР(1), описываемая уравнением

$$x(t) = 0,986 x(t-1) + a(t), t = 1,2,\dots, \\ x(0) = 0$$

где  $a(t)$  – белый шум (некоррелированные одинаково распределенные случайные величины).

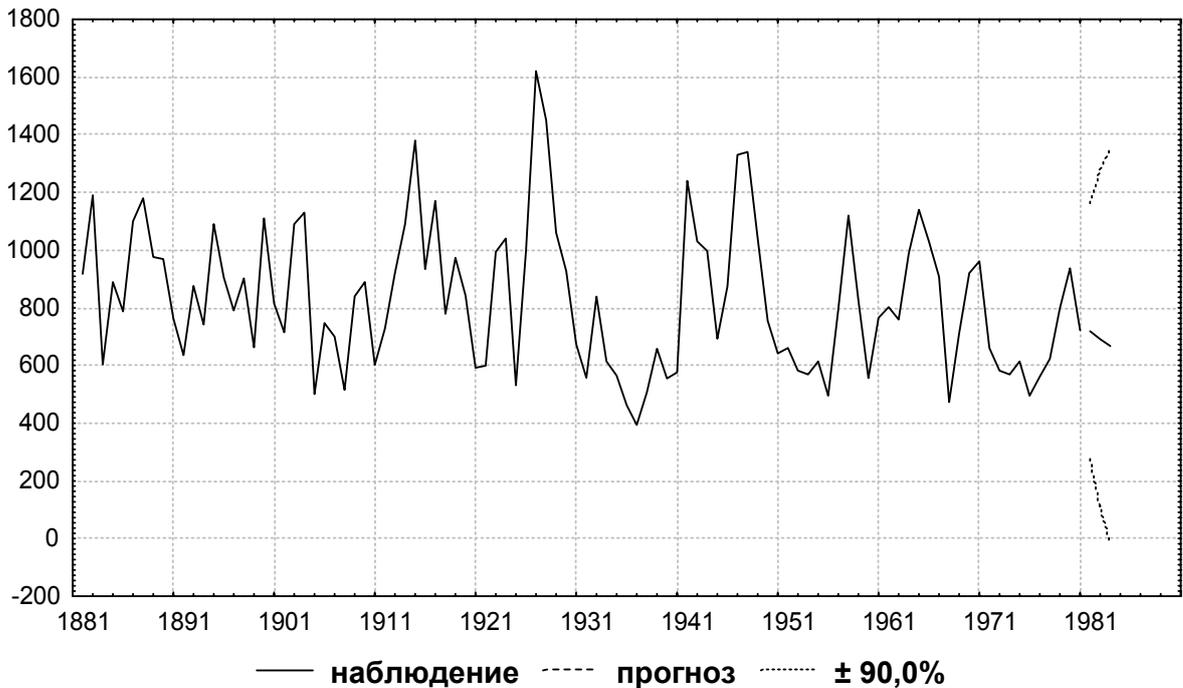
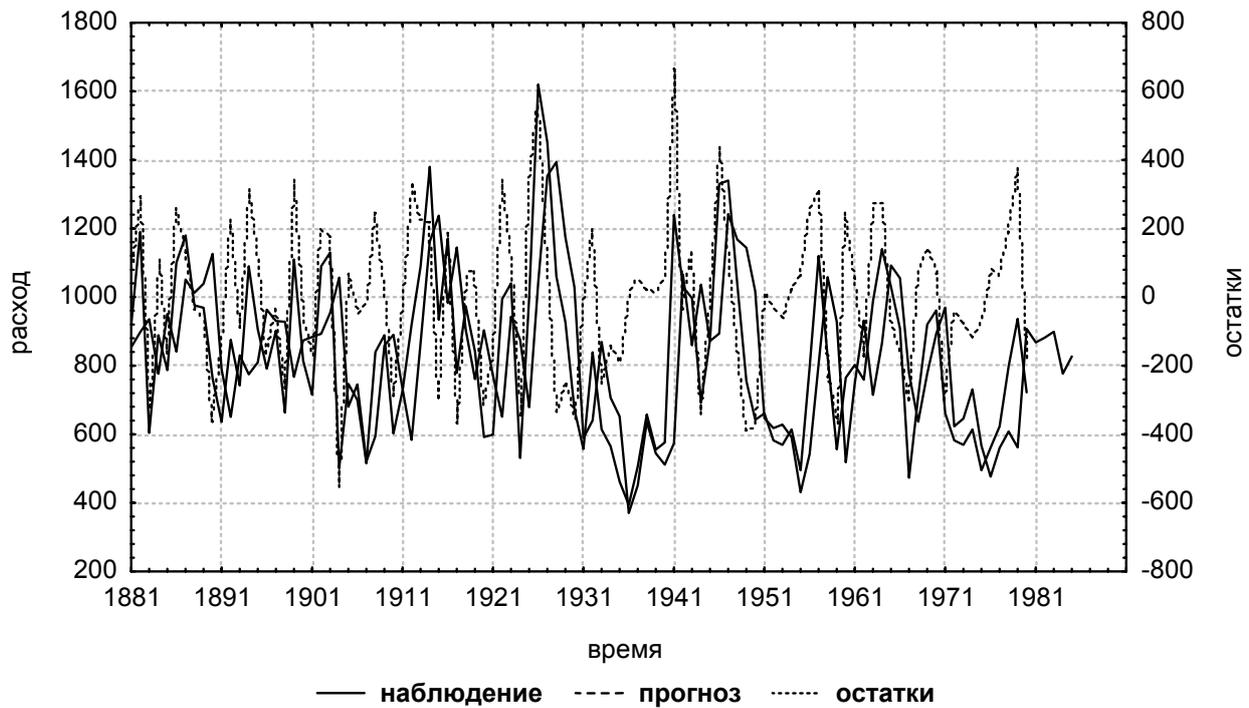


Рис. 1. Прогноз годового стока р.Белая у г.Бирск на 1981 и 1982 гг., выполненный методом АРПСС

Вторым методом прогнозирования является метод экспоненциального сглаживания, позволяющий в ряде случаев также строить прогнозы временных рядов. Суть метода в том, что исходный ряд сглаживается с некоторыми экспоненциальными весами, образуя новый временной ряд  $S(t)$  с меньшим уровнем шума, поведение которого можно спрогнозировать. Данный метод является более простым по сравнению с АРПСС, но обладает большими возможностями, он позволяет учесть сезонный фактор и тренд. Точная формула простого экспоненциального сглаживания имеет следующий вид:



$$S_t = \alpha * X_t + (1-\alpha) * S_{t-1} .$$

Когда эта формула применяется рекуррентно, то каждое новое сглаженное значение, которое также является прогнозом, вычисляется как взвешенное среднее текущего наблюдения и сглаженного ряда. Очевидно, результат сглаживания зависит от параметра  $\alpha$  (альфа). Если  $\alpha$  равно 1, то предыдущие наблюдения полностью игнорируются. Если  $\alpha$  равно 0, то игнорируются текущие наблюдения. Значения  $\alpha$  между 0 и 1 дают промежуточные результаты.

На рис.2 представлены результаты прогноза годового стока р.Сылва у с.Подкаменное на 1981 и 1982 гг., выполненные экспоненциальным методом.

Рис. 2. Прогноз годового стока р.Белая у г.Бирск на 1981 и 1982 гг., выполненный экспоненциальным методом

Наконец, в таблице приведены численные значения прогнозов, полученных двумя методами. Как видно из таблицы, в целом прогнозные значения годового стока незначительно отличаются от фактических. Ошибки прогноза с заблаговременностью 1 год показали достаточную адекватность выбранных статистических моделей.

### Сравнительная характеристика результатов прогнозов среднегодовых расходов воды р.Белая –г.Бирск

Метод	1981 г.		Ошибка		1982 г.		Ошибка	
	Набл.	Прогн.	м <sup>3</sup> /с	%	Набл.	Прогн.	м <sup>3</sup> /с	%
АРПСС	829	739	90	9.8	553	650	97	17.54
Экспоненциально го сглаживания	829	873	44	5.3	553	844	291	52.6

Задачей дальнейших исследований в этом направлении является проверка на примере гидрологических временных рядов других методов прогноза, реализованных в системе STATISTICA.

### **Библиографический список**

1. *Бокс Дж.* Анализ временных рядов. Прогноз и управление / Дж. Бокс, Г.Дженкинс. М.: Мир, 1974. Вып. 1.288 с.; Вып. 2. 254 с.
2. *Боровиков В.П.* Прогнозирование в системе STATISTICA в среде Windows. Основы теории и интенсивная практика на компьютере: учеб. пособие / В.П. Боровиков, Г.И. Ивченко. М.: Финансы и статистика, 1999. 384 с.
3. *Лукашин Ю.П.* Адаптивные методы прогнозирования временных рядов/ Ю.П. Лукашин. М.: Финансы и статистика, 2003. 416 с.