

ГИДРОЛОГИЯ**В.Н. Малинин****О ДОЛГОСРОЧНОМ ПРОГНОЗЕ МЕСТНОГО СТОКА
В ЦЕНТРАЛЬНОМ ФЕДЕРАЛЬНОМ ОКРУГЕ***V.N. Malinin***LONG-TERM PREDICTION LOCAL RUNOFF
IN THE CENTRAL FEDERAL DISTRICT**

Предлагается физико-статистический метод долгосрочного прогноза стока в Центральном федеральном округе (ЦФО). Суть его состоит в построении многоступенчатой прогностической модели, состоящей из нескольких локальных, причем каждая новая модель учитывает результаты предыдущей. Представлены результаты трехступенчатой модели, предикторами которой послужили значения осадков на территории ЦФО за 2 предшествующих года, и температура поверхности океана в Северной Атлантике, а также оценки оправдываемости опытных прогнозов.

Physical-statistic method of long-term prediction of runoff in the Central federal district (CFD) is given. The main point is the formation of multi-step prediction model which consists of several local ones. Each new model takes into account the results of previous one. Presented are the results of three-step model for which predictors served the meanings of precipitation over the area of CFD during 2 previous years, the ocean surface temperature in the North Atlantic and the accuracy meanings of prediction on independent data.

В общем случае из существующих в настоящее время методов долгосрочного прогноза стока крупных рек на Европейской территории России (ЕТР) наиболее перспективным, на наш взгляд, является физико-статистический метод [Малинин; 1994, 2003; Смирнов, 1986]. Суть его состоит в том, что вначале устанавливаются физические связи функции отклика (речного стока) с определяющими факторами, а затем уже на статистической основе строится прогностическая модель, заблаговременность которой зависит от инерционности воздействующих на изучаемый процесс факторов. При этом статистическим аппаратом для решения указанной задачи обычно служит классический метод множественной регрессии.

Определяющими факторами межгодовых колебаний стока крупных рек ЕТР являются:

– запасы влаги в снежном покрове перед началом снеготаяния;

- предшествующее осеннее увлажнение почвы;
- предшествующее летнее увлажнение в апреле – сентябре.

Заметим, чем больше площадь водосборного бассейна, тем больший вклад этих факторов в колебания стока. Так как временные реализации годового стока для крупных водосборных бассейнов достаточно хорошо описываются моделью в виде марковского процесса первого порядка, то необходим учет более длительной предыстории стокоформирующих факторов. Это означает, что сток в i -й год зависит не только от увлажнения в $i-1$ год, но и частично в $i-2$ год.

Итак, основная рабочая гипотеза может быть сформулирована следующим образом: накопление влаги (общее увлажнение) в бассейне за два предшествующих началу половодья года практически полностью определяет речной сток в его замыкающем створе до начала следующего половодья. Естественно, главное влияние на сток оказывает первый предшествующий год. Влияние второго года сказывается главным образом в аномальные по характеру увлажнения годы.

В результате прогностическая модель стока в i -й год может быть записана как

$$Q_i = f \left\{ \sum^{k_1} P_{i-1}, \sum^{k_2} (P - E)_{i-1}, \sum^{k_1} P_{i-2}, \sum^{k_2} (P - E)_{i-2} \right\}, \quad (1)$$

где P – осадки; E – испарение; k_1, k_2 – соответственно сумма месяцев за холодный (с октября по март) и теплый (с апреля по сентябрь) периоды года.

Для Европейской территории России, особенно для зон избыточного и достаточного увлажнения, к которым относится преимущественно и Центральный федеральный округ (ЦФО), характерно, что межгодовая изменчивость осадков значительно превышает аналогичную изменчивость суммарного испарения, в том числе и в теплый период года [Малинин, 1994]. Поэтому формула (1) существенно упрощается

$$Q_i = f \left\{ \sum^{k_1} P_{i-1}, \sum^{k_2} P_{i-1}, \sum^{k_1} P_{i-2}, \sum^{k_2} P_{i-2} \right\} \quad (2)$$

Итак, осадки, выпавшие на территории округа за предшествующие 2 года, а если точнее, то за 21 месяц, являются предикторами при построении прогностической модели местного стока. При этом стокоформирующие факторы в зависимости (2) можно рассматривать как «внутренние», поскольку они входят в уравнение водного баланса водосборного бассейна и, следовательно, непосредственно влияют на изменчивость речного стока. Однако осадки сами являются продуктом крупномасштабного влагообмена в системе океан–атмосфера–материк, характеристики которого через «внутренние» факторы косвенно воздействуют на межгодовые колебания речного стока. Поэтому назовем их условно «внешними» стокоформирующими факторами. К числу

«внешних» стокоформирующих факторов относятся горизонтальный перенос атмосферной влаги, циклоническая активность атмосферы, эффективное испарение над Северной Атлантикой, температура поверхности океана и др. [Малинин, 2003, 2003а].

Исходя из этого, имеет смысл в построении многоступенчатой (иерархической) прогностической модели местного стока для ЦФО, состоящей из нескольких локальных (частных), учитывать эти факторы, причем каждая новая модель учитывает результаты предыдущей. В частности, в качестве предикторов ее первой ступени должны быть использованы внутренние стокоформирующие факторы, а на последующих ступенях – уже внешние факторы. Такая многоступенчатость, на наш взгляд, наилучшим образом обеспечивает учет сложной многофакторной системы причинно-следственных связей формирования межгодовых колебаний речного стока. Отметим, что данные по стоку для ЦФО были любезно предоставлены В.И. Бабкиным.

Наиболее эффективным методом построения статистической модели, на наш взгляд, служит пошаговый алгоритм множественной линейной регрессии (МЛР). Будем в данной работе использовать метод включения переменных, который позволяет отбирать наиболее эффективные (значимые) предикторы. Суть его заключается в том, что на первом шаге выбирается наиболее значимый фактор, описывающий наибольшую долю дисперсии функции отклика, на втором шаге – фактор, описывающий максимальную остаточную дисперсию и т. д. Проблема выбора оптимального шага довольно подробно рассмотрена в работе [Малинин, 2003]. Весьма важно, что метод включения переменных особенно эффективен при большом числе исходных предикторов. Отметим, что именно такой подход был использован в работах [Малинин, 1995; Бабкин, 1995] по долгосрочному прогнозу стока Волги, результаты которого свидетельствуют о его довольно высокой эффективности.

Исходными данными по осадкам послужила система гидрометеорологических данных NOAA NCEP/NCAR Reanalysis [[http://ingrid.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/.MONTHLY/?help+ datasetdatasetselection](http://ingrid.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/.MONTHLY/?help+datasetdatasetselection)], содержащая глобальный архив разнообразных среднемесячных данных с 1949 г., который оперативно пополняется с очень небольшим запаздыванием во времени и находится в свободном доступе в сети Интернет. Среднемесячные осадки содержатся в архиве CDAS (Climate Data Assimilation System) этой системы и могут быть получены в узлах равномерной географической сетки с шагом $1,875^\circ \times 1,875^\circ$. На территории ЦФО выделено 17 узлов (рис. 1), для которых создан специальный архив среднемесячных данных по осадкам за период 1949–2002 гг. Данные по осадкам усреднялись за холодный и теплый период года. Таким образом, получаем 34 предиктора.

Прежде всего осуществлялся анализ изменчивости значений осадков. Учитывая, что последние десятилетия XX столетия характеризуются явно выраженным потеплением климата, нами выполнен расчет линейных трендов, результаты которого приведены в табл. 1.

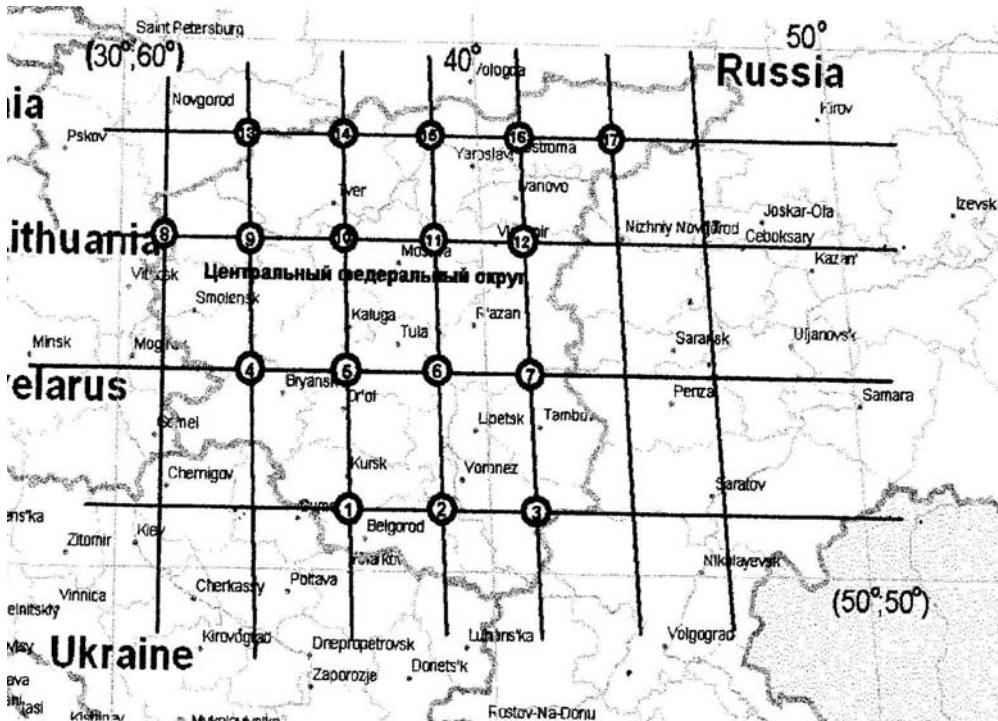


Рис. 1. Расположение узлов географической сетки с шагом $1,875^{\circ} \times 1,875^{\circ}$ на территории Центрального федерального округа

В этой таблице даны:

- коэффициент детерминации R^2 , показывающий долю дисперсии исходного ряда, объясненную уравнением линейного тренда,
- величина тренда T_r , характеризующая изменение осадков за рассматриваемый промежуток времени.

Таблица 1

**Статистические характеристики линейного тренда осадков
за холодный и теплый период года**

Узел сетки	Холодное полугодие		Теплое полугодие	
	R^2	T_r , мм/10 лет	R^2	T_r , мм/10 лет
1	0,001	1,02	0,046	8,12
2	0,005	2,49	0,058	8,43
3	0,018	4,34	0,036	7,07
4	0,180	12,85	0,011	4,86
5	0,099	10,32	0,003	2,59
6	0,074	8,43	0,002	1,67
7	0,071	8,50	0,006	3,42
8	0,239	14,03	0,000	0,43
9	0,109	9,69	0,001	1,36
10	0,048	6,43	0,000	-0,83
11	0,060	7,39	0,004	-3,02

Окончание табл. 1

12	0,052	7,00	0,011	-5,07
13	0,044	6,57	0,039	-9,96
14	0,001	0,91	0,017	-5,70
15	0,006	1,83	0,018	-5,67
16	0,021	3,92	0,067	-11,65
17	0,000	0,40	0,108	-16,23

Отметим, что в табл. 1 курсивом выделены значимые на уровне $\alpha = 0,05$ тренды. Нетрудно видеть, что в холодный период на всей территории ЦФО отмечаются положительные тренды осадков, значимые на всем протяжении от границ с Белоруссией до Тамбовской области. Максимальный тренд, который отмечается вблизи границы с Белоруссией (узел 8, рис. 2), составляет более 14 мм/10 лет. В теплый период года, исключая крайний северо-восток ЦФО (узел 17), значимые тренды в осадках отсутствуют.

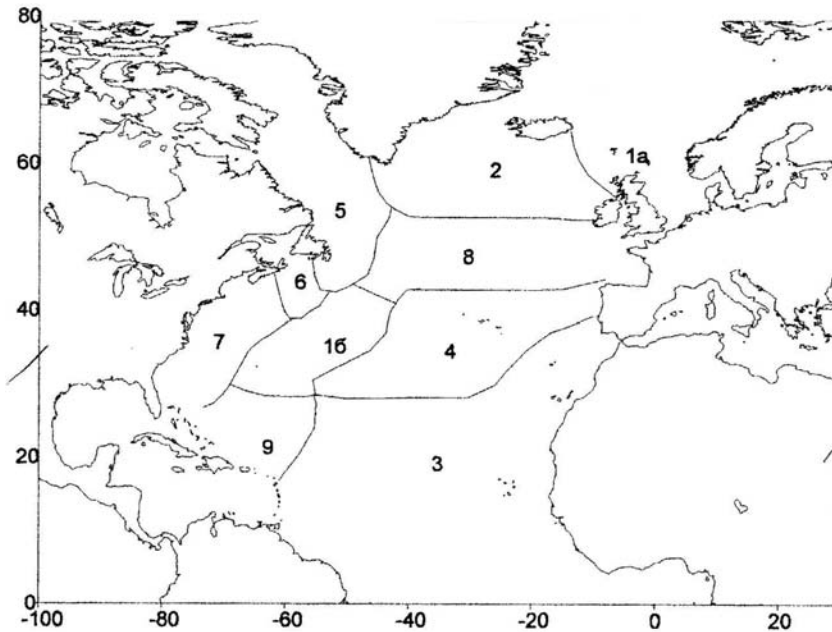


Рис. 2. Районирование акватории Северной Атлантики по межгодовым колебаниям ТПО в холодный период года на основе факторного анализа [Гордеева, 1999].

Перед построением прогностической модели местного стока в ЦФО предварительно весь архив данных об осадках был разбит на две выборки: зависимую (с 1950 по 1994 г.), по которой непосредственно рассчитывалась и тестировалась модель, и независимую (с 1995 по 2002 г.), используемую для оценки оправдываемости опытных прогнозов. Отметим, что стандартное отклонение местного стока с территории ЦФО составляет $\sigma_y = 22,5 \text{ км}^3/\text{год}$.

Первая ступень модели имела вид:

$$Q_i = f \left\{ \sum^{k1} P_{i-1}, \sum^{k2} P_{i-1} \right\}. \quad (3)$$

Заметим, что с целью увеличения заблаговременности осадки за холодный период брались за октябрь–декабрь. Это связано с тем, что осадки за октябрь–март и октябрь–декабрь очень тесно коррелированы между собой.

В результате использования метода включения переменных была получена следующая оптимальная модель, в которой предикторы расположены в порядке значимости:

$$Q^{(1)} = -a_0 + a_1 P_{12}^x + a_2 P_{15}^T + a_3 P_6^T - a_4 P_{12}^T + a_5 P_{17}^T. \quad (4)$$

Например, P_{12}^x и P_{12}^T означают соответственно осадки за холодный и теплый период в 12-м узле сеточной области. Основные параметры этой модели даны в табл. 2. Все коэффициенты регрессии данной модели значимы по критерию Стьюдента, а сама модель адекватна по критерию Фишера. Однако коэффициент детерминации все же не столь высок, как бы хотелось. Поэтому осуществлялось построение модели второй степени, в которой предикторами являются значения местного стока, предвычисленные по модели (4), и осадки за второй предшествующий стоку год. Рассчитанная модель второй степени имеет вид:

$$Q^{(2)} = -a_0 + a_1 Q^{(1)} + a_2 P_{12(i-2)}^T - a_3 P_{6(i-2)}^T. \quad (5)$$

Параметры уравнения (5) также приводятся в табл. 2.

Как видно из табл. 2, прогностическая модель (5) заметно улучшает качество модели первой степени. Безусловно, главным показателем надежности прогностических моделей служит составление опытных прогнозов, т.е. проверка ее точности на независимых данных. С этой целью обычно используется отношение $\delta = \sigma_{y(x)} / \sigma_y$, причем для долгосрочных прогнозов принимается, что независимые прогнозы должны удовлетворять условию $\delta < 1$. Действительно, обе модели удовлетворяют этому условию. На наш взгляд, модель (5) даже при использовании только внутренних стокоформирующих факторов уже вполне приемлема для ее использования на практике. Тем не менее была предпринята попытка включить в прогностическую модель внешние факторы.

Таблица 2

Оценки параметров регрессионных моделей для местного стока на территории ЦФО

Мо- дель	П а р а м е т р					
	Число предик- торов	Коэффици- ент детер- минации, R^2	Крите- рий Фи- шера, F	Стандартная ошибка модели, $\sigma_{y(x)}$, км ³ /год	Эффективность модели, $\delta = \sigma_{y(x)} / \sigma_y$	
					Зависимая выборка	Независимая выборка
4	5	0,60	12,1	15,1	0,67	0,88
5	3	0,66	26,2	13,4	0,60	0,82
6	3	0,78	18,4	11,2	0,50	0,48

Как было показано в работе [Гордеева, 1999], одним из важных «внешних» предикторов для предвычисления стока крупных рек ЕТР (Волга, Нева, Северная Двина) служит поле поверхностной температуры океана (ТПО) в Северной Атлантике в холодный период года. Действительно, колебания ТПО в холодный период по существу являются индикатором изменения теплосодержания верхнего квазиоднородного слоя и процессов крупномасштабного взаимодействия океана и атмосферы, в конечном счете, определяющих интенсивность переноса атмосферной влаги на территорию ЕТР. Естественно, что колебания ТПО должны оказывать воздействие и на межгодовую изменчивость водных ресурсов ЦФО.

В связи с этим нами сделана попытка построения третьей ступени модели местного стока в ЦФО. Предикторами для нее послужили среднемесячные значения ТПО в узлах географической сетки 5° широты и 10° долготы в зоне $10\text{--}65^\circ$ с.ш. (всего 64 точки) за 1950–2002 гг. и предвычисленный по модели (5) местный сток.

В связи с большим объемом данных о ТПО мы воспользовались результатами районирования акватории Северной Атлантики за холодный период, которое было выполнено с помощью классического факторного анализа. В результате было выделено 9 классов или 10 квазиоднородных районов [Гордеева, 1999], распределение которых дается на рис. 2. Для каждого класса путем простого арифметического осреднения находились центры районов, а корреляция центров с отдельными узлами района позволила определить наиболее репрезентативные узлы. Именно ряды ТПО в этих узлах и стали предикторами при построении третьей ступени модели. В результате удалось существенно снизить размерность исходных предикторов с 64 до 10 без больших потерь в описании дисперсии исходного поля ТПО.

Оптимальная модель третьей ступени включает, помимо предвычисленного по модели (5) стока, температуру поверхности океана во втором и четвертом районах Северной Атлантики, т.е.

$$Q^{(3)} = -a_0 + a_1 Q^{(2)} - a_2 T_2 + a_3 T_4. \quad (6)$$

Параметры этой модели также приводятся в табл. 2. Естественно, она улучшает предыдущую модель и, на наш взгляд, дальнейшее уточнение модели (6) хотя вполне возможно, однако с практической точки зрения вряд ли уже целесообразно. Отметим, что второй район расположен в зоне действия Исландской депрессии, а четвертый находится под влиянием Азорского максимума. Обратный характер связи ТПО в указанных районах способствует усилению зонального переноса влаги в атмосфере и соответственно формированию положительных аномалий осадков на территории ЦФО.

Обратимся теперь к рис. 3, на котором приведено сопоставление фактических и прогностических значений местного стока в ЦФО. Подчеркнем, что период 1995–2002 гг. является независимым.

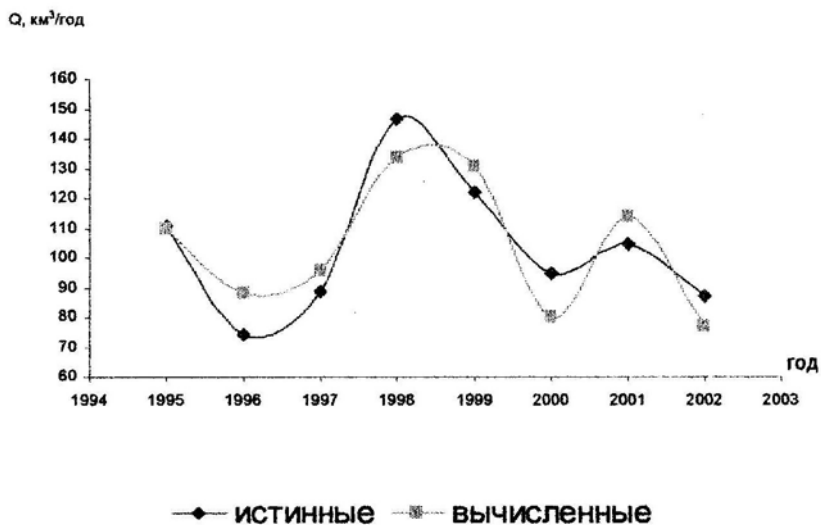


Рис. 3. Сопоставление фактических и прогностических значений местного стока в ЦФО.

Как следует из рисунка, систематическая ошибка практически отсутствует, а стандартная ошибка равна $10,8 \text{ км}^3/\text{год}$. Тогда $\delta = 0,48$, что даже чуть ниже, чем по зависимой выборке. Итак, введение в модель внешних предикторов позволило получить прогностическую модель стока в ЦФО, которая уже вполне отвечает практическим запросам.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант 05-05-65041).

Литература

1. Бабкин В.И., Малинин В.Н. О методике долгосрочного прогноза стока Волги. – Сб.: Условия форм. и методы прогноза стока Волги. – СПб.: Гидрометеоздат, 1995, с. 65–69 (<http://ingrid.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/.MONTHLY/?help+ datasetdatasetselection>).
2. Гордеева С.М., Малинин В.Н. К оценке влияния теплового состояния Северной Атлантики на колебания стока крупных рек Европейской территории России. – Труды РГГМУ, 1999, вып. 123, с. 92–100.
3. Малинин В.Н. Проблема прогноза уровня Каспийского моря. – СПб.: изд. РГГМУ, 1994. – 160 с.
4. Малинин В.Н. Физико-статистический метод долгосрочного прогноза годового стока Волги. – Сб.: Условия форм. и методы прогноза стока Волги. – СПб.: Гидрометеоздат, 1995, с. 52–54.
5. Малинин В.Н., Гордеева С.М. Физико-статистический метод прогноза океанологических характеристик (на примере Северо-Европейского бассейна). – Мурманск: изд. ПИНРО, 2003. – 164 с.
6. Малинин В.Н., Радикевич В.М., Гордеева С.М., Куликова Л.А. Изменчивость вихревой активности атмосферы над Северной Атлантикой. – СПб.: Гидрометеоздат, 2003а. – 171 с.
7. Смирнов Н.П., Скляренко В.Л. Методы многомерного статистического анализа в гидрологических исследованиях. – Л.: Изд. ЛГУ, 1986. – 190 с.