

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ УРОВНЯ ВОДЫ В РЕКЕ ГОРНОГО ТИПА (НА ПРИМЕРЕ РЕКИ МЗЫМТА)

**Семенчин Е.А., Титов Н.Г., Кузякина М.В., Лебедев К.А.**

*ФГБОУ ВПО «Кубанский государственный университет», Краснодар, e-mail: rector@kubsu.ru*

Некоторые реки Краснодарского края являются реками горного типа. Реки данного типа обладают крутым падением водотока (угол наклона течения) и площадью водосбора, отличающейся от рек равнинного типа своей подстилающей поверхностью. Такие особенности реки делают невозможным использование ранее разработанных математических моделей для рек горно-равнинного типа. В данной статье предлагаются две математические модели прогнозирования уровня воды в реках горного типа. Первая модель позволяет производить прогноз с помощью рассчитанного уравнения линейной регрессии, описывающего изменения уровня воды в русле горной реки Мзымта; другая – производить прогноз уровня воды, основываясь на математической модели созданной на основе нейронной сети архитектуры многослойного персептрона. Также в статье проведено сравнение полученных математических моделей.

**Ключевые слова:** математическое моделирование, регрессионный анализ, метод наименьших квадратов, нейронные сети, прогнозирование, многослойный персептрон, метод обратного распространения, паводковые ситуации

## COMPARATIVE ANALYSIS OF METHODS OF MATHEMATICAL MODELING OF THE WATER LEVEL IN THE RIVER MOUNTAIN TYPE (FOR EXAMPLE, THE RIVER MZYMTA)

**Semenchin E.A., Titov N.G., Kuzyakina M.V., Lebedev K.A.**

*Kuban State University, Krasnodar, e-mail: rector@kubsu.ru*

Some rivers of Krasnodar region are mountain type rivers. Rivers of this type have a steep watercourse drop(the angle of flow) and a catchment area different from river plain type by its underlying surface. These features make it impossible to early developed mathematical models for rivers with mountain-flat type. In this paper, we propose two mathematical models for predicting the water level in the mountain type river. The first model allows to produce prediction by using linear regression equation describing the water level change in the bed of a mountain river called Mzymta; the other – to produce a prediction of the water level, based on a mathematical model developed on basics of neural network architecture of multilayer perceptron. Also, the article compared the obtained mathematical models.

**Keywords:** mathematical modeling, regression analysis, the method of least squares, neural networks, forecasting, multilayer perceptron, back propagation, the flood situation

На территории Краснодарского края находится немало рек, которые следует отнести к рекам горного типа. В основном эти реки протекают в горных районах Сочи – Туапсе. Рекам такого типа свойственны высокая скорость течения и резкое изменение уровня воды в русле за счет обильных осадков, ледникового и снегового питания, что не так характерно для рек равнинного и горно-равнинного типов. Как показывают вычислительные эксперименты, методики, описывающие изменения уровня воды в руслах рек равнинного и горно-равнинного типов, не применимы для описания изменения уровня воды в руслах рек горного типа. В данной статье предлагается анализ двух математических моделей прогнозирования уровня воды в реках горного типа. Первая модель позволяет производить прогноз с помощью рассчитанного уравнения линейной регрессии, описывающего изменения уровня воды в русле горной реки

Мзымта; другая – производить прогноз уровня воды, основываясь на математической модели, созданной на основе нейронной сети архитектуры многослойного персептрана.

**Цель работы** – апробировать ранее разработанную математическую модель прогноза уровня воды в реке горного типа на основе уравнения линейной регрессии на промежутке времени в один год; апробировать ранее разработанную математическую модель прогноза уровня воды в реке горного типа на основе построенной нейронной сети архитектуры многослойного персептрана на промежутке времени в один год; провести анализ полученных результатов методами математической статистики и на основе данного анализа выбрать математическую модель, вычисляющую наиболее приближенный к реальному значению будущий уровень воды в реке горного типа.

### Оценка коэффициентов в уравнении линейной регрессии, описывающем изменения уровня воды в русле горной реки

Предположим, что значения уровня воды в реке на момент времени  $t_i$  совпадают со значениями функции [4].

$$y(t_i) = f(x_1(t_{i-5}), x_1(t_{i-4}), x_1(t_{i-3}), x_1(t_{i-2}), x_1(t_{i-1}), x_2(t_{i-5}), x_2(t_{i-4}), x_2(t_{i-3}), x_2(t_{i-2}), x_2(t_{i-1})), \quad (1)$$

$$t_{i-5} < t_{i-4} < t_{i-3} < t_{i-2} < t_{i-1} < t_i, i = 6, 7, \dots, n,$$

где  $y(t_i)$  – прогнозируемый уровень воды в створе данной реки в районе гидрологического поста,  $x_1(t_{i-5})$ ,  $x_1(t_{i-4})$ ,  $x_1(t_{i-3})$ ,  $x_1(t_{i-2})$ ,  $x_1(t_{i-1})$  – уровень воды в моменты  $t_{i-5}$ ,  $t_{i-4}$ ,  $t_{i-3}$ ,  $t_{i-2}$ ,  $t_{i-1}$  соответственно,  $x_2(t_{i-5})$ ,  $x_2(t_{i-4})$ ,  $x_2(t_{i-3})$ ,  $x_2(t_{i-2})$ ,  $x_2(t_{i-1})$  – количество осадков, выпав-

ших в соответствующие моменты времени в окрестности русла горной реки. Пусть соотношение (1) определяет математическую модель процесса изменения уровня воды в реке горного типа.

Пусть функция (1) является линейной:

$$y(t_i) = b_1 + a_1 \cdot x_1(t_{i-5}) + a_2 \cdot x_1(t_{i-4}) + a_3 \cdot x_1(t_{i-3}) + a_4 \cdot x_1(t_{i-2}) + a_5 \cdot x_1(t_{i-1}) + \\ + a_6 \cdot x_2(t_{i-5}) + a_7 \cdot x_2(t_{i-4}) + a_8 \cdot x_2(t_{i-3}) + a_9 \cdot x_2(t_{i-2}) + a_{10} \cdot x_2(t_{i-1}), \quad (2)$$

где  $a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}, b_1$  – некоторые постоянные.

На основе статистических данных, взятых из гидрологических бюллетеней за 2010 год по реке Мzymта, методами ре-

грессионного анализа [1] были вычислены в статье [4] значения коэффициентов и свободного члена в (2) (табл. 1).

Согласно данным, приведенным в табл. 1, функция (2) (регрессионная модель) имеет вид

$$y(t_i) = 3,014 + 0,0465x_1(t_{i-5}) + 0,0312x_1(t_{i-4}) + (-0,155)x_1(t_{i-3}) + \\ + 0,236x_1(t_{i-2}) + 0,830x_1(t_{i-1}) + 0,00457x_2(t_{i-5}) + (-0,0377)x_2(t_{i-4}) + \\ + (-0,312)x_2(t_{i-3}) + 0,166x_2(t_{i-2}) + 0,214x_2(t_{i-1}), \quad (3)$$

**Таблица 1**  
Значения коэффициентов и свободного члена в регрессионной модели (2)

| Наименование коэффициента | Значение коэффициента |
|---------------------------|-----------------------|
| $b$                       | 3,014                 |
| $a_1$                     | 0,047                 |
| $a_2$                     | 0,031                 |
| $a_3$                     | -0,155                |
| $a_4$                     | 0,236                 |
| $a_5$                     | 0,830                 |
| $a_6$                     | 0,005                 |
| $a_7$                     | -0,038                |
| $a_8$                     | -0,313                |
| $a_9$                     | 0,166                 |
| $a_{10}$                  | 0,214                 |

Точность оценки прогноза значений  $y(t_i)$  определяется числовыми значениями характеристик регрессионной модели (3), приведенными в табл. 2, коэффициентом детерминации [3]

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

и значением критерия Фишера [3]

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - k}{k - 1}, \quad (5)$$

где  $n$  – размер выборки (количество проведенных замеров);  $k$  – число коэффициентов модели (3);  $y_i$  – наблюдаемое значение  $y(t)$ ;  $\tilde{y}_i$  – значение  $y(t_i)$  вычисленное с помощью уравнения регрессии (3) в момент  $t_i$ ,  $i = 6, 7, \dots, n$ ;  $\bar{y}$  – среднее значение  $y(t)$ , вычисленное на основе статистических данных.

**Таблица 2**  
Характеристики регрессионной модели (2)

|                              |        |
|------------------------------|--------|
| $R^2$                        | 0,971  |
| Нормированное значение $R^2$ | 0,970  |
| Стандартная ошибка $\sigma$  | 10,217 |
| Количество наблюдений $n$    | 545    |
| $F$ -критерий Фишера         | 1808   |

Если  $F$  превышает некоторое критическое значение  $F_{kp}$ , определяемое по таблице значений  $F$ -критерия Фишера при заданных  $n$ ,  $k$  и заданной доверительной вероятности, а величина  $R^2$  близка к 1, то значение  $y(t_i)$ , вычисленное по формуле (3), с вероятностью, близкой к 1, совпадает с экспериментальными данными. Согласно статистическим данным и допущениям, указанным выше,  $F_{kp} = 10,534$ ,  $\alpha = 0,95$ ,  $n = 545$ ,  $k = 10$ ,  $R^2 = 0,97131498$ ,  $F = 1808$ .

Так как  $F > F_{kp}$  ( $1808 > 10,534$ , т.е. значение  $F$  во много раз превышает  $F_{kp}$ ), то с вероятностью 0,95 при стандартной ошибке  $\sigma = 10,217$  значения  $y(t_i)$  совпадают с реальными значениями  $y(t_i)$ ,  $i = 6, 7, \dots, n$  [2].

На рис. 1 приведены графики значений  $y(t_i)$ , построенные на основе статистических данных и значений регрессионной модели (3). Из этих графиков видно, что прогнозируемые с помощью (3) значения  $y(t_i)$  несущественно отличаются от статистических данных.

График распределения реального и предсказанного с помощью регрессионного анализа значения уровня воды



Рис. 1. Графики значений  $y(t_i)$ , построенные на основе статистических данных регрессионной модели (3)

### Прогнозирование уровня воды в реке горного типа на основе нейронной сети, имеющей архитектуру многослойного персептрона

Для разработки нейронной сети, позволяющей построить прогноз уровня воды в русле горной реки, с архитектурой многослойного персептрона и последующего ее обучения были использованы статистические данные о метеоусловиях на гидрологическом посту, расположенным на реке Мзымта, предоставленные Краснодарским центром гидрометеорологии и мониторинга окружающей среды. Использовался пакет прикладных программ Statistica 6.1 [8].

Обучение производилось с помощью метода обратного распространения, основной идеей которого является распространение сигналов ошибки от выходов сети к ее входам в направлении, обратном прямому распространению сигналов в обычном режиме работы.

Архитектура построенной согласно (1) нейронной сети для прогноза паводков в русле указанной реки определяется как многослойный персептрон (трехслойный), имеющий на первом слое 8 нейронов, на втором – 5, на третьем – 1 [7]. Данная архитектура представлена в виде упрощенной схемы на рис. 2.

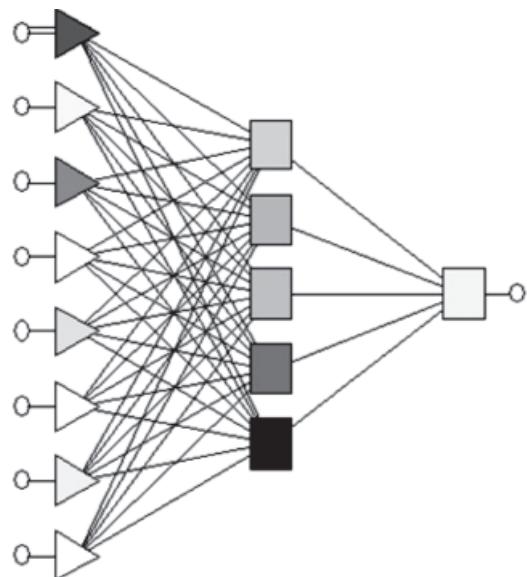


Рис. 2. Схема работы нейронной сети архитектуры трехслойного персептрона:  
8 – входных сигналов, изображенных кругами;  
8 – синапсов первого слоя, изображенных треугольниками; 5 – синапсов второго слоя, изображенных квадратами;  
1 – синапс третьего слоя, изображенный прямоугольником; 1 – выходной сигнал, изображенный кругом

**Таблица 3**  
Значения весов на слое 1

| Вес синапса | Значение веса синапса |
|-------------|-----------------------|
| $w_1$       | 0,181                 |
| $w_2$       | -1,136                |
| $w_3$       | -0,568                |
| $w_4$       | -0,657                |
| $w_5$       | -0,987                |
| $w_6$       | 0,045                 |
| $w_7$       | 0,033                 |
| $w_8$       | -0,006                |
| $w_9$       | 0,038                 |
| $w_{10}$    | -0,316                |
| $w_{11}$    | -0,237                |
| $w_{12}$    | 0,619                 |
| $w_{13}$    | -0,396                |
| $w_{14}$    | -0,078                |
| $w_{15}$    | -0,609                |
| $w_{16}$    | -0,434                |
| $w_{17}$    | 0,249                 |
| $w_{18}$    | 0,344                 |
| $w_{19}$    | -0,442                |
| $w_{20}$    | -0,942                |
| $w_{21}$    | 0,026                 |
| $w_{22}$    | 0,780                 |
| $w_{23}$    | 0,818                 |
| $w_{24}$    | 0,155                 |
| $w_{25}$    | 0,667                 |
| $w_{26}$    | 0,236                 |
| $w_{27}$    | -0,376                |
| $w_{28}$    | -0,372                |
| $w_{29}$    | -0,582                |
| $w_{30}$    | -0,069                |
| $w_{31}$    | -0,502                |
| $w_{32}$    | -0,775                |
| $w_{33}$    | 1,323                 |
| $w_{34}$    | -0,059                |
| $w_{35}$    | 0,321                 |
| $w_{36}$    | 0,655                 |
| $w_{37}$    | 0,512                 |
| $w_{38}$    | -0,300                |
| $w_{39}$    | -0,709                |
| $w_{40}$    | -1,127                |

В каждый нейрон подаются входные значения  $x_i$  с некоторой поправкой  $w_i(x_i)$ , затем сумма  $w_1(x_1) + w_2(x_2) + w_3(x_3)$  поступает на передаточную функцию

$$f(w_1(x_1) + w_2(x_2) + w_3(x_3)),$$

где  $w_i(x_i)$ ,  $i = 1, 2, 3$  принято называть соответственно входными дендритами и синапсами [5].

Описанная нейронная сеть использует алгоритм обучения обратного распространения, который использует несколько слоев нейронов, связанных между собой. Задача обучения нейронной сети сводится к нахождению функциональной зависимости  $y = f(x)$ . Для сужения области поиска необходимых весов необходимо уменьшить функцию ошибки нейронной сети. Это достигается с помощью метода обратных квадратов [6]:

$$E(w) = \frac{1}{2} \sum_{j=1}^p (y_j - d_j)^2,$$

где  $d_j$  – целевое значение  $j$ -го выхода;  $p$  – число нейронов в выходном слое.

Вторым этапом обучения нейронной сети являлось обучение методом градиентного спуска. Каждая итерация сопровождалась изменением веса по формуле [6]:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}},$$

где  $\eta$  – параметр, определяющий скорость обучения.

Полученные в результате обучения веса представлены в табл. 3, 4.

**Таблица 4**  
Значения весов на слое 2

| Вес синапса | Значение веса синапса |
|-------------|-----------------------|
| $w_1$       | 0,277                 |
| $w_2$       | -1,017                |
| $w_3$       | 0,846                 |
| $w_4$       | -0,316                |
| $w_5$       | 0,634                 |

**Таблица 5**  
Характеристики нейронной сети

|                                |       |
|--------------------------------|-------|
| Производительность обучения    | 0,388 |
| Контрольная производительность | 0,306 |
| Тестовая производительность    | 0,574 |
| Ошибка обучения                | 0,061 |
| Контрольная ошибка             | 0,050 |
| Тестовая ошибка                | 0,081 |
| Стандартная ошибка $\sigma$    | 8,694 |

Для обучения описанной нейронной сети были использованы данные об уровне воды и количестве осадков в реке Мзымта

за январь–декабрь 2010 года. Характеристики построенной нейронной сети приведены в табл. 5.

### График распределения реального и предсказанного с помощью нейронной сети значения уровня воды



*Рис. 3. Графики значений  $y(t)$ , построенные на основе статистических данных с помощью нейронной сети*

### График ошибок



*Рис. 4. График ошибок*

На рис. 3 приведены графики значений  $y(t)$ , построенные на основе статистических данных и значений, полученных с помощью нейронной сети. Из этих графиков видно, что прогнозируемые значения  $y(t)$  несущественно отличаются от статистических данных.

### Выводы

Сравнивая полученные результаты, можно сделать следующие выводы.

Оба предложенных в статье метода дают адекватный результат. Однако метод, основанный на регрессионном анализе, показывает большую стандартную ошибку ( $10,21772 > 8,693615$ ) по сравнению с методом, основанным на использовании нейросетевых технологий. Визуализация полученных результатов, представленная на рис. 1, 3, 4, подтверждает проведенные расчеты.

### Список литературы

1. Статистическое моделирование и прогнозирование [под ред С.А. Гранборга]. – М.: Финансы и статистика, 2000. – 383 с.
2. Ивченко Г.И., Медведев Ю.И. Математическая статистика: учеб. пособие для вузов. – М.: Высш. шк., 1984. – 248 с.
3. Крамер Г. Математические методы статистики. – М.: ир, 1975. – 375 с.
4. Титов Н.Г., Семенчин Е.А., Об оценке коэффициентов в уравнении линейной регрессии, описывающем изменения уровня воды в русле горной реки // Известия Кубанского государственного университета. Естественные науки. – 2013. – № 1(2). – С. 49–51.
5. Медведев В. С., Потемкин В.Г. Нейронные сети. – М.: Диалог МИФИ, 2002. – 496 с.
6. Осовский С. Нейронные сети для обработки информации. – М.: Финансы и статистика, 2002. – 344 с.
7. Титов Н.Г., Семенчин Е.А., Лебедев К.А., Прогноз уровня воды в реке горного типа с помощью нейронной сети имеющей архитектуру многослойного персептрона // Современные концепции научных исследований: матер. IV международной научно-практич. конф. (Москва, 25 июля 2014 г.). – М., 2014. – № 4 (2) – С. 149–152.
8. Халафян А.А. Статистический анализ данных. STATISTICA 6.0. – Краснодар: Изд-во КубГУ, 2003. – 191 с.
2. Ivchenko G.I., Medvedev Y.I, Mathematical statistics: a manual for technical colleges, 1984. 248 p.
3. Kramer G. Mathematical statistic methods, 1975. 375 p.
4. Titov N.G., Semenchin E.A., About the estimation of coefficients in linear regression equation describing the water level change in the bed of a mountain river // Kuban State University News. Natural sciences, 2013. no. 1(2). pp. 49–51.
5. Medvedev V.S., Potemkin V.G. Neural networks, Dialog MIFI, 2002, 496 p.
6. Osovskiy S. Neural network for information processing, Finance and Statistics, 2002. 344 p.
7. Titov N.G., Semenchin E.A., Lebedev K.A. A prediction of the water level in the mountain type river using neural network having an architecture of a multi-layer perceptron // Modern concepts of scientific research: materials of IV International scientific and practical conference (Moscow, July 25, 2014), 2014. no. 4(2). pp. 149–152.
8. Halafyan A.A. Statistical data analysis. STATISTICA 6.0.. Krasnodar: KubGU Publishing House. 2003. 191 p.

### Рецензенты:

Халафян А.А., д.т.н., профессор, ФГБОУ ВПО «Кубанский государственный университет», г. Краснодар;

Ургенов М.А.Х., д.ф.-м.н., зав. кафедрой прикладной математики, ФГБОУ ВПО «Кубанский государственный университет», г. Краснодар.

Работа поступила в редакцию 19.12.2014.

### References

1. Statistical modeling and forecasting (by S.A.Granborg editing), Finance and Statistics, 2000. 383 p.