Analysis of the effects of climatic factors on flood peak formation

M E Bukovskiy and A V Semenova

Derzhavin Tambov State University, 33 Internatsional’naya Street, 392000 Tambov, Russia
E-mail: asv273@mail.ru

Abstract. Spring flooding is an annual significant overflow of water in the central Russia rivers. As a rule, the maximum flood levels in the Tambov region rivers are reached in the late March - early April. However, this parameter is very variable. Moreover, the maximum flood water level varies from year to year. In this paper, we study the effects of various climatic parameters on the spring flood peak formation. We consider the Tsna River (the Oka basin) near Tambov as the model object of our study. Using methods of multiple correlation analysis, we determine the most significant climatic factors which affect the maximum flood level. In the research, 16 climatic and hydrological parameters from 1970 to 2018 are analyzed. We assess the contribution of the various factors to flood wave formation by using multiple regression analysis. Using the analysis, we have identified the most significant flooding factors.

1. Introduction

Human development is closely linked to the state of water resources. Water land plays a key role in people's lives. It is widely used for industrial and domestic water supply, shipping, timber rafting, energy, irrigation and other purposes [1].

Every year more and more specialists in climatology, meteorology, and hydrology note, among the problems which threaten humanity, the problem of flooding of the inhabited territories [2].

European researchers believe that floods cause great economic losses and people’s death in many parts of Europe. Among the 100 most expensive natural disasters of the 20th century, flood damage amounts to 40% of the total losses (including earthquakes, typhoons, storms, fires, etc.), which are evaluated as 680 billion euros. Operative flood management and flood protection planning require detailed spatial and temporal information about flooding when the flood is active or in other states [3].

Local protection of buildings and infrastructure, early evacuation can be very effective measures to protect the population against the flood consequences. Reliable hydrological forecasts are necessary for all these measures [4].

Domestic researchers also consider flood as one of the most dangerous natural phenomena that cause emergencies and material damage to industrial and agricultural facilities and to human health and lead to their death [5]. Flood can be caused by high water, freshet and ice jams on rivers.

The causes of floods in the Tambov region are mostly spring snow melting and the water level rising in the rivers connected with it.

In this paper, one of the flood types – spring flood (high water) – is considered. It is a multifactorial phenomenon that is affected by a whole complex of natural factors: water reserves in snow, precipitation, temperature values, etc. [6].
The spring flood is a characteristic feature of the lowland rivers of Russia, including the Volga River basin [7]. The city of Tambov is located on the banks of the Tsna River, which relates to Volga basin. High floods near Tambov is a source of hydrological emergencies. It causes flooding of summer cottages, houses, and the infrastructure.

Many researchers have been studying the problem of flood forecasting in different aspects.

Over a long period of time, physical modeling was used to forecast hydrological phenomena, as well as to create hydraulic models of atmospheric, oceanic, and other flows [8,9,10,11]. Undoubtedly, physical-type models have shown high efficiency in forecasting various flooding options. However, the need for hydro-geomorphological monitoring data obtained during intensive calculations makes them unsuitable for short-term forecasts [12]. According to Kim et al., the creation of physical models requires fundamental knowledge of various hydrological parameters and experience in the designing of such models that complicate the forecasting activity [13]. In addition, many researchers note drawbacks of the short-term forecasting capabilities of physical-type models [11].

Besides the numerical and physical models, data-driven models are gaining more and more popularity in flood modeling. These methods combine fixed climatic indexes and hydrometeorological parameters for increasing the modeling capability.

The Russian scientists D.A. Burakov, V.F. Kosmakova, and I.N. Gordeev use the regression dependencies for creating a physical and mathematical model of long-term forecast for large Siberian rivers [14,15,16]. Their methodology is based on multiple linear regression, which makes it possible to study the interconnected multifactorial natural processes [14].

The most popular methods of flood frequency analysis (FFA) for the modelling of flood forecasts are statistical models of autoregressive moving average (ARMA), multiple linear regression (MLR), and autoregressive integrated moving average (ARIMA) [17,18,19].

The FFA is one of the first methods of statistics in flood forecasting [20]. In comparison with physical forecasting models; a more developed and, at the same time, a more efficient one is regional flood frequency analysis (RFFA) including the computation cost and generalization [21].

B.M. Belyaev and N.A. Varentsova developed a methodology for long-term forecasts of maximum flood level based on hydrometeorological information. They also considered its features, effectiveness, and prospects of the statistical approach in hydrological forecasts, especially for small rivers [6].

At the same time, A.N. Arzhakov and K.I. Kusatov tested the methodology for forecasting the maximum levels of spring floods based on integrated comparative geographic and statistical analysis taking into account satellite information [22].

Because of the computing power development, flood forecasting models, such as machine learning (ML) models, are becoming increasingly common. Another driving force for the widespread use of such models is their ability to numerically formulate flood nonlinearity based only on historical data without information on the basic physical processes. These models are really prospective for forecasting because of their fast development with minimal cost of it [23]. Continuous development of the ML methods over the past two centuries showed that these methods are better suited for forecasts with optimal speed than conventional methods, although it requires high technical equipment [24].

In the sphere of effective preparedness and prompt preventive measures, one of the most effective systems for reducing damage is a well-functioning warning system [25]. That is why long-, medium-, and short-term forecasting is an important scientific objective whose results can be demanded by the regional services of the Ministry of Emergency Situations [26].

Some foreign researchers indicate a significant role of precipitation and the spatial distribution of the hydrological cycle in flood modeling [27]. However, as it turned out later, for more accurate flood forecasting one cannot rely on a precipitation forecast. Thus, in the long-term forecast of a river it is advisable to take into account not only the precipitation, but also the estimates of the soil moisture in the catchment [28].

Analysis of the formation conditions of spring runoff performed by A. M. Shustov showed that «high floods are formed with large water reserves in cold winter snow and late united snow melting» [29].
At the same time, A.M. Vladimirov notes that significant factors which form a flood peak are climatic [30]. First, it is precipitation which can be solid and liquid. The solid one is represented by snow reserves, which are measured in a catchment at the flood start, and snowfalls during the flood. It determines the volume and the peak of the flood. The liquid one is rainfalls and drizzle, which are able to form new peaks on the rise and on the decline of the flood.

Temperature is the climatic factor that is second in significance. Its course strongly affects the intensity of snow melting. If the sum of positive temperatures during the snowmelt increases, it intensifies and the period of the rising flood level is reduced, and the peak is reached faster.

The goal of this paper is to assess the significance of various climatic factors in flood formation using the statistical methods.

2. Data and methods
The basis of our research is the data of the Tambov Center for Hydrometeorology and Environmental Monitoring.

Our model object is the Tsna River (the Oka basin) near Tambov.

A review of papers of other researchers makes it possible to highlight the main factors of flood formation, which include the water reserves in snow cover, the soil freezing depth, the temperature regime, and the precipitation before and during the flood.

The factors which affect the maximum level in a river during the flood were formalized in the following parameters:

- snow melting starting dates;
- dates of stable transition of the average daily air temperature above 0°C upwards (also below – “dates of transition above 0°C”);
- dates of complete snow melting;
- flood starting dates;
- maximum flood level dates;
- water reserves in the snow cover (also below – just “snow”) on the snow melting starting date;
- water reserves in the snow cover on the date of stable transition of the average daily air temperature above 0°C upwards;
- water reserves in the snow cover on the flood starting date;
- maximum flood levels;
- sums of average daily air temperatures above 0°C (also below – “sums of positive temperatures”) from the snow melting starting date to the maximum flood level date;
- sums of average daily air temperatures above 0°C from the dates of stable transition of the average daily air temperature above 0°C upwards to maximum flood level date;
- sums of average daily air temperatures above 0°C from the flood starting date to the maximum flood level date;
- sums of precipitation from the snow melting starting date to the maximum flood level date;
- sums of precipitation on the dates of stable transition of the average daily air temperature above 0°C upwards to the maximum flood level date;
- sums of precipitation from the flood starting date to the maximum flood level date;
- soil freezing depth on the snow melting starting date;
- soil freezing depth on the date of stable transition of the average daily air temperature above 0°C upwards;
- soil freezing depth on the flood starting date.

The analysis was performed for the period from 1970 to 2018.

At the first stage, we have created a data base containing values of the parameter mentioned above for each year.

Then we analyzed the degree of linking of the maximum flood level with each factor being considered. The partial Pearson’s correlation coefficients were verified by Student’s t-test.
The calculations were carried out by MS Excel. To characterize the correlation we used the Cheddock’s scale. After that we sorted the values of all parameters by the most significant criterion, which is the snow melting starting date. Thus, we got the periods of 15 days (at the end of February it is 13 (14) days, and at the end of March, 16 days).

To assess the contribution of climatic factors to the formation of flood wave and determine significant ones, we used multiple regression analysis.

Further analysis of significance of the factors which affect the flood peak formation was already carried out within these periods.

3. Results and discussion
During the correlation analysis of parameters which affect the maximum flood level, we determined the ones that have weak and middle direct links with the flood peak. The t-statistic value modulus in all cases exceeded the critical tabular values. Figure 1 shows the parameters whose correlation coefficients are higher than 0.4 at a significance level of 0.01.

In Figure 1 we can see that the maximum flood level of the above-listed climatic parameters is most correlated with the snow melting starting date, the date of stable transition of the average daily air temperature above °C upwards, the water reserves in the snow cover on the flood starting date, and the date of complete snow melting.

The snow melting starting date ($r=0.65$) and the date of complete of snow melting ($r=0.53$) have an average correlation. For other 3 parameters the correlation coefficient ranges from 0.42 to 0.49.
Then we sorted all parameters by the most significant one, the snow melting starting date. The values of all parameters relating to years with the same snow melting starting dates were combined by the following intervals: February 15-28 (29); March 1-15; and March 16-31.

Also, according to the State Hydrological Institute guidelines, we ranked the maximum flood levels as low (P<33.3%), average (33.3%≤P≤66.7%), and high (P>66.7%) [31].

As a result of the analysis, we determined that when snow starts melting before the 28th (29th) of February, floods are low and sometimes middle. If the snow melting starting date is in the range from the 1st to the 15th of March, we can observe low, middle, and high floods equally. But if the snow melting starts from the 16th to 31st, the floods are average and high. Such a distribution of the maximum flood level depending on the start of snow melting allows us to make preliminary qualitative assessments of the upcoming flood based on this date.

The next stage was analysis of the links between the maximum flood level and the different components of its formation: the sum of average daily air temperatures above 0 ºС, the sum of precipitation, the soil freezing depth, and the water reserves in the snow cover depending on the snow melting starting date. We used regression analysis with a reliability level of 95%.

| Table 1. Assessment of the contribution of climatic factors to flood formation |
|-----------------------------------------|-----------------|-----------------|
| Snow melting starting date              | 13.017          | 4.563           | 0.001          |
| Date of stable transition of the average daily air temperature above 0ºC upwards | 2.837           | 0.788           | 0.436          |
| Date of complete snow melting           | -0.472          | -0.131          | 0.897          |
| Flood starting date                     | -7.808          | -1.852          | 0.074          |
| Water reserves in snow on the snow melting starting date | -0.596          | -0.694          | 0.493          |
| Water reserves in snow on the date of transition above 0ºC | -1.598          | -2.086          | 0.045          |
| Water reserves in snow on the flood starting date | 2.209           | 2.463           | 0.020          |
| Sums of positive temperatures from the snow melting starting date to flood peak | 2.809           | 1.101           | 0.279          |
| Sums of positive temperatures from the flood starting date to flood peak | -7.857          | -2.196          | 0.036          |
| Sums of positive temperatures from the date of transition above 0ºC to flood peak | 4.443           | 1.659           | 0.107          |
| Sums of precipitation from the date of transition above 0ºC to flood peak | 3.919           | 1.525           | 0.137          |
| Sums of precipitation from the snow melting starting date to flood peak | 2.018           | 0.847           | 0.404          |
| Sums of precipitation from the flood starting date to peak | -0.110          | -0.044          | 0.965          |
| Soil freezing depth on the snow melting starting date | 2.193           | 1.170           | 0.251          |
| Soil freezing depth on the date of transition above 0ºC | 0.917           | 0.447           | 0.658          |
| Soil freezing depth on the flood starting date | -3.104          | -1.218          | 0.232          |

The multiple correlation coefficient is 0.85, and the standard error is 94.

According to Table 1, we determined several most statistically significant parameters which affect the flood. These are the snow melting starting date, the water reserves in snow cover on the flood starting date, and the sum of average daily air temperatures above 0 ºС from the flood starting date to the maximum flood level date.
4. Conclusions
According to the results of our research, the flood height has a strong dependence on the snow melting starting date. This parameter can be used for qualitative assessment of the height of an upcoming flood.

The probability of a high flood is one third when the snow melting starting date is from the 15th of February to the 28th (29th) of March. And if it the melting starts after the 16th of March, the probability of a high flood is 50%.

In our research we have analyzed the relationships between 16 climatic parameters and maximum flood water levels. The most significant factors which determine a flood peak are the snow melting starting date, the water reserve in the snow cover at the flood starting time, and the sum of average daily air temperatures above 0ºС from the flood starting time to the maximum flood level time.

References
[1] Dudnik S N, Bukovskiy M E, Kolkova K S, Surovikina I V and Chernova M A 2014 Dynamics of ice phenomenons on the rivers of Don basin within the territory of Tambov region Geopolitics and ecogeodynamics of regions 10 (1) 393-98
[2] Arzhanykh Yu P, Dolzhenkov V V and Zvyagintseva A V 2014 Forecasting of the hydrological situation (in the period of the high water) on water objects of the Voronezh region with use of geographical information systems Heliogeophysical researches 9 (9) 89-98
[3] Malnes E, Solbo S, Lauknes I, Evertsen G J, Tøllefsen T A, Solheim I and Indregard M 2005 FLOODMAN-Global near-real time flood monitoring for hydrological users ACTIF Proc. Int. Conf. on Innovation Advances and Implementation of Flood Forecasting Technology (Bergen-Tromso) (Electronic Materials) p 1-9
[4] Nester T, Komma J and Blöschl G 2016 Real time flood forecasting in the Upper Danube basin J. Hydrol. Hydromech. 64 (4) 404-14 doi: 10.1515/johh-2016-0033
[5] Golubev A B and Zemtsov V A 2013 Danger and flood risk assessment in Barnaul (Zaton village) Tomsk University Bulletin 373 183-8
[6] Belyaev B M and Varentsova N A 2018 Development of long-term hydrological spring high-water period forecasts based on modern hydrometeorological information and methods of its processing 3rd Vinogradov Readings «The Edge of Hydrology» (St. Petersburg: High Technology p 698–703
[7] Lavrov S A and Kalyuzhnyi I L 2016 Climatic Changes Impact upon the Spring High-water Runoff and its Formation Factors in the Volga Basin Russia's water industry: problems, technologies, management 6 42-60
[8] Zhao M and Hendon H H 2009 Representation and prediction of the Indian Ocean dipole in the POAMA seasonal forecast model Q. J. R. Meteorol. Soc. 135 337-52 doi: 10.1002/qj.370
[9] Costabile P, Costanzo C and Macchione F 2012 Comparative analysis of overland flow models using finite volume schemes J. Hydroinform. 14 122-35 doi: 10.2166/hydro.2011.077
[10] Xia X, Liang Q, Ming X and Hou, J 2017 An efficient and stable hydrodynamic model with novel source term discretization schemes for overland flow and flood simulations Water Resour. Res. 53 3730-59 doi: 10.1002/2016wr020055
[11] Costabile P and Macchione F 2015 Enhancing river model set-up for 2-D dynamic flood modelling Environ. Model. Softw. 67 89-107 doi: 10.1016/j.envsoft.2015.01.009
[12] Nayak P, Sudheer K, Rangan D and Ramasastri K 2005 Short-term flood forecasting with a neurofuzzy model Water Resour. Res. 41 doi: 10.1029/2004wr003562
[13] Kim B, Sanders B F, Famiglietti J S, Guinot V 2015 Urban flood modeling with porous shallow-water equations: A case study of model errors in the presence of anisotropic porosity J. Hydrol. 523 680-92 doi: 10.1016/j.jhydrol.2015.01.059
[14] Burakov D A, Gordeev I N and Kosmakova V F 2009 Applying of regression dependencies in long-term forecasts of maximum water levels on the example of the Yenisei River near Kyzyl Questions of geography of Siberia (Tomsk: Tomsk State University) pp 5-6
[15] Burakov D A, Kosmakova V F and Gordeev I N 2012 The results of testing a physical and statistical method for a long-term forecast of the maximum water levels of the spring flood of the Yenisei river near Kyzyl Results of testing new and improved technologies, models and methods of hydrological forecasts 39 121-6

[16] Burakov D A, Kosmakova V F and Gordeev I N 2015 About the results of operational tests of hydrological forecasting methods on the rivers of Eastern Siberia during the spring flood Results of testing new and improved technologies, models and methods of hydrological forecasts 42 50-9

[17] Valipour M, Banihabib M E and Behbahani S M R 2012 Parameters estimate of Autoregressive Moving Average and Autoregressive Integrated Moving Average Models and compare their ability for inflow forecasting J. Math. Stat. 8 330-8 doi: 10.3844/jmssp.2012.330.338

[18] Adamowski J, Fung Chan H, Prasher S O, Ozga-Zielinski B and Slisuarieva A 2012 Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada Water Resour. Res. 48 doi: 10.1029/2010wr009945

[19] Valipour M, Banihabib M E and Behbahani S M R 2013 Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir J. Hydrol. 476 433-41 doi: 10.1016/j.jhydrol.2012.11.017

[20] Chow V T, Maidment D R and Mays L W 1988 Applied hydrology (New York: MacGraw-Hill) p 149

[21] Aziz K, Rahman A, Fang G and Shrestha S 2014 Application of artificial neural networks in regional flood frequency analysis: A case study for Australia Stoch. Environ. Res. Risk Assess. 28 541-54 doi: 10.1007/s00477-013-0771-5

[22] Arzhakov A N and Kusatov K I 2009 The results of testing the method of forecasting the maximum water levels of the spring flood on the rivers of Yakutia for 2004-2006 Results of testing new and improved technologies, models and methods of hydrological forecasts 36 86-97

[23] Mekanik F, Imteaz M, S Gato-Trinidad and Elmahdi A 2013 Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes J. Hydrol. 503 11-21 doi: 10.1016/j.jhydrol.2013.08.035

[24] Mosavi A, Rabczuk T and Varkonyi-Kocz C R 2017 Reviewing the novel machine learning tools for materials design Advances in Intelligent Systems and Computing 660 50-8 doi: 10.1007/978-3-319-67459-9_7

[25] Kumani M V, S ysemko V I and Apuhtin A V 2012 Multi-level system for monitoring and forecasting the maximum flow rates and water levels of the spring flood Scientific notes: Stochastic and environmental research Kursk State University 4-2 (24) 213-7

[26] Kosheleva E D and Kushchin A V 2018 Short-term forecast of Ob river water level near Barnaul city during flood 2018 News of the Altai Branch of the Russian Geological Society 3 (50) 27-37

[27] Lafdani E K, Nia A M, Pahlavanravi A, Ahmadi A and Jajarmizadeh M 2013 Daily Runoff and Simulation Using ANN, ANFIS and Conceptual Hydrological MIKE11/NAM Models Int. J. Eng. Technol. 1 32-50

[28] Seo D J and Breidenbach J 2002 Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements J. Hydrometeorol. 3 93–111 doi: 10.1175/1525-7541(2002)003<0093:RTCOSN>2.0.CO;2

[29] Shustov A M 2004 Issues of engineering hydrology of the rivers of the Republic of Mordovia (Saransk: Mordovian University) p 80

[30] Vladimirov A M 2009 The factors determining occurrence of the high flow and highest water level during a flood Scientific notes of the Russian State Hydrometeorological University 9 22-39
[31] Ishkhanyan M V and Karpenko N V 2016 *Econometrics* vol 1 (Moscow: MIIT) p 117
[32] *Guidelines for determining the calculated hydrological characteristics in the absence of data of hydrometric observations* 2009 (Saint Petersburg: Nestor-History) p 193