The forecast of the thunderstorm on the basis of settlement fields of meteorological parameters of global prognostic model (GFS NCEP) for areas of Western Caucasus

A Kh Kagermazov
Russia, KBR, 360024, Nalchik, Kabardinskaya St. 208-76

E-mail: ka5408@mail.ru

Abstract. The main reasons for the lack of warning of dangerous convective phenomena are the lack of initial information, as well as the lack of traditional approaches to predict local rapidly developing processes. At present, the results of atmospheric parameters prediction based on global and mesoscale models have been successfully used in the study and forecast of thunderstorm processes occurring over a limited area. They are based on non-stationary three-dimensional equations of atmospheric hydrodynamics and parametrization of atmospheric processes (shortwave and longwave radiation fluxes, convective processes, boundary layer, moisture microphysics, atmospheric turbulence, heat and moisture exchange in the underlying surface).

In this article, according to the calculated output of the global atmospheric model (Global Forecast System GFS), a method is proposed for predicting lightning processes based on the statistical interpretation of the “model output” for the regions of southern European Russia.

For different climatic zones of the North Caucasus, statistical forecast models were constructed and the results of assessments of the development of thunderhair processes by the corresponding discriminant functions for characteristic regions of the southern European Russia — the Western and Central parts of the North Caucasus — are presented.

Introduction
In recent years, significant changes have taken place in the information support of methods for forecasting dangerous weather phenomena associated with convective processes in the atmosphere. The intensive development of computer facilities, information technologies (Internet) and, as a result, the rapid application of the results of calculations of global, regional, mesoscale models of the atmosphere, opens up previously inaccessible opportunities for the development of forecasting methods of hazardous weather phenomena and their consequences on the basis of a new information base [1-3]. In this work, as in [3], the method based on using the output data of the Global Atmosphere Model is used to predict the development of thunder-grave processes with subsequent discriminant analysis. The MOS (model output statistics) concept assumes that the predictive hydrodynamic model is unchanged, however, the output products of the current GFS model used by us - stratification of temperature, humidity, wind direction and speed fields - an analogue of upper-air atmospheric sounding, already has a sufficiently high accuracy. This was confirmed by the validation of these fields carried out by us with a lead time (earliness) of 12-24 hours. The results of validation based on actual data of upper-air sounding of the global model products showed a high degree of coincidence of
their values [7]. The correlation coefficients of the predicted values from the model and real data for the upper-air probe at the level of 0.9 and higher.

Therefore, the continuous improvement of the global model will have a slight effect on the part of the output that we use, and the discriminant functions will not lose their relevance.

**Materials and research methods**

As a result of the research [2, 4, 11], it was found that the formation of phenomena of a convective nature (hail, thunderstorm, heavy rain, etc.) has the greatest effect on the below listed atmospheric parameters. This is the speed of the orderly ascent of air at 700 mb (V700); total specific humidity in the layer of the earth's surface - 500 hPa (SQZ5); energy characteristics of the sub-cloud layer (DSS); the average moisture deficit in the layer Hk + 5 km (TDSR5); Miller thermodynamic index (TTMI). When forecasting lightning processes, to take into account the climatic features of the territory, the data were grouped in accordance with the following zones southern European Russia - the Western and Central parts of the North Caucasus:

- The central part (Stavropol Territory, Karachay-Cherkessia, KBR, Ossetia, Chechen Republic);
- The western part (the Black Sea coast of the North Caucasus, points Anapa, Gelendzhik, Tuapse, Sochi).

The result of discriminant analysis is the construction of a discriminant model (discriminant function), which has the form:

\[ d = a + b_1 x_1 + b_2 x_2 + \cdots + b_n x_n, \]

where \( d \) the grouping (dependent) variable; \( b_n \) coefficient of discriminant function; and \( a \) the constant term (constant); \( x_n \) discrimination (independent) variables.

Discriminant analysis was carried out using the statistical package of computer programs SPSS [9-11]. Discriminating (independent) variables are predictors, calculated from the global model GFS (the concept of output statistics of MOS models).

The investigated groups (phenomena) we have are:

- thunderstorm
- the absence of thunderstorm “without thunderstorm”

The grouping variable is L2 which has two values: 0 (absence of thunderstorm), 1(thunderstorm fallout).

43 cases with thunderstorm and thunderstorm were considered, all these data were involved in discriminant analysis.

The construction of the discriminant model consists in calculating and analyzing the coefficients of the discriminant function. The constructed discriminant model should clearly distinguish the groups under study. The quality of the constructed discriminant model is characterized by the value of the correlation coefficient between the calculated values of the discriminant function and the actual membership of the group equal to "0.838" in the case under consideration, which indicates the high accuracy of the constructed discriminant model.

Not standardized (initial) coefficients of discriminant function are presented in Table 1, they are used for creation of discriminant model.

**Table 1. Initial coefficients of discriminant function**

| Function | DTM | DSS | DTK | V700 | SQZ5 | (Constant) |
|----------|-----|-----|-----|------|------|------------|
| 1        | 0.644 | -0.004 | 0.428 | -0.053 | 0.119 | -10.932    |

In accordance with the data presented in Table 1, the discriminant model constructed as a result of the discriminant analysis has the following form:
\[ L2 = -10.932 + 0.119 \times SQZ5 + 0.644 \times DTM - 0.004 \times DSS + 0.428 \times DTK - 0.053 \times V700 \] (1)

The accuracy of the predictions based on the constructed discriminant model is determined from the data of the summary table of classification results (Table 2), i.e. assigning the objects of research to one of the groups under study.

### Table 2. The results of classification

| L2 value | Forecast of accessory to groups | general |
|---------|--------------------------------|---------|
| 0       | 27                             | 2       | 29 |
| 1       | 0                              | 14      | 14 |
| %       | 93.1                           | 6.9     | 100.0 |
| 1       | 0.0                            | 100.0   | 100.0 |

From data of the Table 2 "results of classification" it is visible that the studied group "without thunderstorm" consists of 29 cases. According to the constructed discriminant model, 27 cases from 29 were correctly ranked as this group, 2 are by mistake ranked as group "with the thunderstorm". Correct results of classification were 93.5%, and inaccurate 6.9%.

According to the table the studied group "with the thunderstorm" consists actually of 14 cases. According to the constructed discriminant model of 14 cases from 14, were correctly ranked as this group. Total correct results of classification were 100%.

In general, the correct results of classification were 95.3%. This means that in 95.3% of cases the actual belonging of the investigated case "with thunderstorm" or "without thunderstorm" will coincide with the predicted one, determined on the basis of the constructed discriminant model. This makes it possible to conclude that the accuracy of the predictions made on the basis of the constructed discriminant model is about 95%.

The results of an operational check of statistical models of the thunderstorm forecast

An operative check of the alternate method of the forecast of a hailstorm for the Western part of the North Caucasus was carried out also on earlier received discriminant functions for all region of the North Caucasus according to the concept of the ideal forecast of perfect prognoses (PP) [8-10].

The calculations were carried out using the formula for the prediction of hail, which has the form:

\[ L1 = -0.146 \times SQZ5 + 0.53 \times DTM + 0.023 \times DSS + (0.717 \times DTK1 + +0.717 \times DTK2) - 0.114 \times V700 - 7.61 \] (2)

To assess the success of the forecast method for materials for the period May-September 2015, a contingency table was compiled (Table 3).

### Table 3. Contingency table

| forecast       | It was observed | sum |
|----------------|-----------------|-----|
| thunderstorm   |                 |     |
| n11 number of the come true forecasts with a formulation «a thunderstorm» | 11   | 13  |
| n12 number of not come true forecasts with a formulation «a thunderstorm» | 2    | 12  |
| n21 number of not come true forecasts with a formulation «not a thunderstorm» | 3    | 30  |
| n22 number of the come true forecasts with a formulation «not a thunderstorm» | 27   | 43  |
| sum            |                 |     |
| n01            |                 |     |
| n02            |                 |     |
| n00            |                 |     |

n10 = n11 + n12 number of all estimated forecasts with a formulation "a thunderstorm",
n20 = n21 + n22 number of all estimated forecasts with a formulation "not a thunderstorm".
n01 = n11 + n21 number of all estimated cases with the thunderstorm,
n02 = n12 + n22 number of all estimated cases without thunderstorm,
n00 = n10 + n20 = n01 + n02 number of all estimated cases.

The general justification of the forecast P = (n11 + n22) / n00 x 100; P = (11 + 27) / 43 x 100 = 88%.
The warning of the presence of the phenomenon P1 = (n11 / n01) x 100; P = 11 / 14x100 = 78.6%

Additional indicators for the assessment of urban hazard forecasts are:
The justification for the prediction of the presence of the phenomenon P2 = (n11 / n10) x 100 = 
(11/13) x100 = 84.6%; The justification for the absence of the phenomenon P3 = (n22 / n20) x 100 = 
(27/30) x 100 = 90%; The warning of the absence of the phenomenon P3’ = (n22 / n02) x 100 = 
(27/29) x 100 = 93%; The quality criterion for the forecasts of Pirs-O, Obukhov T = (n11 / n01) - (n12 / n02) = 
(27/11) / (27/29) = 0.71 ~ 0.71. The error risk of the method (the phenomenon was not pre
dicted, but was observed): a = n21 / n01 = 0.1 The error of insurance of the method (the 
phenomenon was predicted, but not observed): P = n12 / n02 = 0.07. The accuracy of criterion on 
A.M. Obukhov: Q = 1 - a - p ~ 0.83.

Based on the results of calculations and analysis of the success of the forecasts, it can be concluded
that the tested method of hail forecast meets all the criteria for the quality of forecasts.

Based on the materials of 2015 (Anapa, Gelendzhik, Tuapse, Sochi), a new discriminant function 
(1) was obtained directly for the black sea coast of the North Caucasus, which was used in the 
operational test of the hail prediction method.

Assessment of the success of the separation of the thunderstorm phenomenon based on the 
materials of these points for the period may-september 2015 is presented in the contingency table 4, 
where the corresponding (n) are the same as in Table 3.

| Forecast | It was observed | sum |
|-----------|----------------|-----|
|           | thunderstorm   | not thunderstorm |
| thunderstorm | 14 (n11) | 2 (n12) | 16 (n10) |
| not thunderstorm | 0 (n21) | 27 (n22) | 27 (n20) |
| sum | 14 (n01) | 29 (n02) | 43 (n00) |

The general justification of the forecast P = (n11 + n22) / n00 x 100; P = (14+27) /43 x 100 = 95%.
The warning of the presence of the phenomenon P1 = (n11/n01) x 100; P = 14/14x100 = 100% And 
additional indicators for assessing the quality of thunder hazard forecasts have the following 
meanings: The justification for the prediction of the presence of the phenomenon P2 = (n11 / n10) x 
100 = (14/16)x100 = 87 %; The justification for the absence of the phenomenon P3 = (n22/ n20) x 100 
= (27 /27)x 100 = 100 %; The warning of the absence of the phenomenon P3’ = (n22/ n02) x 100 = (27 
/ 29) x 100 = 93 %; The quality criterion for the forecasts of Pirs-O, Obukhov T = (n11 / n01) - (n12 / n02) = 
(27/11) / (27/29) = 0.93~0.93. The error risk of the method (the phenomenon was not pre
duced, but was observed): a = n21 / n01 = 0.1 The error of insurance of the method (the phenom
enon was predicted, but not observed): P = n12 / n02 = 0.07. The accuracy of criterion on A.M. Obukhov: Q 
= 1 - a - p ~ 0.93.

The final results of the estimation of the hail forecast method for statistical models are presented in 
Table 5.

| Table 5. Forecasts and their accuracy |
|--------------------------------------|
| Discriminant functions | Total Forecasts | Predictions with phenomena | Forecast without phenomena |
|                        | Issued | Justification | Issued | Justification | Issued | Justification |
|                        | amt   | %            | amt   | %            | amt   | %            |
Based on the results of the calculations and analysis of the success of the forecasts, it can be concluded that the proposed statistical models of the thunderstorm forecast correspond to all the quality criteria for the forecasts.

Discussion
For the noted climatic zones of the North Caucasus, discriminant functions have been constructed that can be used to predict thunder-grave processes:

Central part:

\[ L1 = -0.146 \times SQZ5 + 0.53 \times DTM + 0.023 \times DSS + (0.717 \times DTK1 + 0.717 \times DTK2) - 0.114 \times V700 - 7.61 \]  (2)

West Side:

\[ L2 = 0.119 \times SQZ5 + 0.644 \times DTM - 0.004 \times DSS + (0.428 \times DTK1 + 0.428 \times DTK2) - 0.053 \times V700 - 10.9 \]  (1)

From equations (1) - (2) it can be seen that they are distinguished by free members and a set of independent variables and coefficients for them, which indicates that climatic features significantly influence the genesis of lightning-graded processes. The processes on the Black Sea coast are less dependent on the large-scale atmospheric circulation (the V700 parameter is missing) and are more associated with local conditions.

Summary
Multidimensional statistical schemes have been developed, the training of which was carried out both on the basis of the concept of an “ideal” PP forecast (actual data) and according to the concept of output data of the global atmosphere model (MOS concept).

The parameters included in the discriminant functions reflect the instability margin (DTM, DTK) of the moisture content (SQZ5) of the atmosphere and the energy saturation (DSS) of the sub-cloud layer. The corresponding nature of large-scale circulation (V700) is a necessary background in the formation of intense convection, leading to lightning-grave processes.

Based on the results of the calculation of the criteria, the quality of the statistical models obtained is quite good, differing with the technology and ease of application in operational practice.

The applied approach should be promising when forecasting other consequences of strong convection in the atmosphere, such as strong wind, heavy rain, and associated floods, landslides, and mudflows.

References
[1] Akimov V A, Durnev R A, Sokolov Yu I 2009 Dangerous hydrometeorological phenomena in Russia (Moscow, VNII GOCHS).
[2] Kagermazov A Kh 2012 Forecast of hail on the output data of the global atmosphere model (T254 NCEP) (J. Meteorology and hydrology) 3 28-34.
[3] Kagermazov A Kh, Fedchenko L M 2012 Using the concept of the “ideal” method of statistical analysis (perfect prognos methods-pp) for predicting hail (J. Izvestija of the KBNTS RAN) 4 (48) 75-80.
[4] Kagermazov A Kh, Sirotenko L A 2010 Results of approbation of software modules for calculating meteorological parameters and computer forecast of hail according to aerological sounding data in the North Caucasus (J. News of Higher Educational Establishments. North-Caucasian region. Physics of the atmosphere. Natural Sciences) Special Issue 49-53.
[5] Perekhodtseva E V 2011 On the hydrodynamic and statistical prediction of up to two days of strong winds and heavy precipitation events for the North Caucasus (Proceedings of the Hydrometeorological Center of Russia) 347 113-125.

[6] Alekseyeva A A, Losev V M 2014 Forecast of strong summer precipitation based on the output of the regional model of the Russian Hydrometeorological Center (Proceedings of the State Medical Center of Russia) 351 43-63.

[7] Kagermazov A Kh 2014 Validation of the Global Forecasts System output data with upper-air sounding results (J. Izvestija of the KBNTS RAN) 3 (59) 32-36.

[8] Tolstykh M A, Zhelen Zh F, Volodin E M, Bogoslovsky N N, Vilfand R M, Kiktev D B, Krasyuk D B, Kostrykin S V, Mizyak V G, Fadeev R Yu, Shashki V V, Shlyaeva A V, Ezau I N, Yurova A Yu 2015 Development of a multi-scale version of the global atmospheric model PLAV (J. Meteorology and Hydrology) 6 25–35.

[9] Golovlev K N, Lukin A A, Perekhodtseva E V 2015 About the hydrodynamic-statistical forecast of the storm wind in the Northern, Norwegian and Barents Seas (Proceedings of the Hydrometeorological Center of Russia) 354 63-78.

[10] Information on http://www.emc.ncep.noaa.gov/GFS/doc.php

[11] Fedchenko L M, Kagermazov A Kh 1988 The use of statistical methods for the prediction of hail processes and their characteristics (J. Meteorology and hydrology) 4 41-50.

[12] Moosmyuller G, Rebik N N 2009 Marketing research with SPSS (INFRA-M, Moscow).

[13] Information on http://r-analytics.blogspot.de/2012/07/blog-post.html

[14] Information on http://math.semestr.ru/corel/cheddok.php