Forecast of thunderstorm with increasing advance time

A Kh Kagermazov,†, L T Sozaeva

†Federal state budgetary institution «High-Mountain Geophysical Institute», Nalchik, Lenina av. 2, Russia

E-mail: ka5408@mail.ru

Annotation. A method for assessing the hail hazard of the atmospheric situation according to the output data of the global atmospheric model (GFS NCEP) with increasing lead time is proposed. The method is implemented on the basis of multivariate discriminant analysis. The evaluation of the success of the method is carried out and the preservation of its predictive potential is shown for a lead time of up to five days.

1. Introduction

Dangerous weather events associated with atmospheric convection make a significant contribution to the damage caused to the economies of countries by natural disasters, and have a pronounced upward trend, compounded by their lack of predictability. The main reasons for the lack of prevention of convective hazards are the lack of initial information, as well as the lack of traditional approaches for predicting local rapidly developing processes.

In recent years, the output of global mathematical models of the atmosphere has become available, including the fields of meteorological elements (stratifications of the fields of temperature, humidity, wind direction and speed) with different lead times. The output of the current model of the NOAA global Forecasting System (GFS NCEP), an analogue of the aerological probe, used in this work, has a sufficiently high reliability and allows obtaining information with a lead time of up to 10 days. There are a number of publications [1-4] using this data.

In this paper, we study the dependence of the validity of the hail forecast on the increasing up to five days in advance.

2. Materials and methods of research

The NOAA Global Forecasting System (GFS NCEP) became widely used and well-known in the early 90s [5-7]. With the growth of computing resources and changes in computer architecture, its resolution also increased. In June 2019, the global forecasting system under the NGS (Next Generation Global Prediction System) project was significantly updated. The transition to a new non-spectral block for solving dynamic equations was made, and a number of improvements were introduced in the description of physical processes at the grid scale. The dynamic core splits the atmosphere into small cubic cells located on the grid and calculates the changes in the parameters within each cell. This allows you to increase the scale of the calculated grid when modeling mesoscale storm systems to improve their forecast. The time discreteness of the model is three hours for a lead time of 0-180 hours and 12 hours for a lead time of 180-384 hours.
The part of the output of the global model required for use in methods for predicting convective phenomena (analogous to an aerological sound) includes the prognostic fields of the following meteorological elements:
- isobaric surfaces (hPa) and their corresponding heights (m);
- air temperature (°C);
- dew point temperature (°C);
- wind direction (deg);
- wind speed (m/s).
These weather elements are calculated at various levels from 1000 to 20 hPa.

The algorithm used in this paper for calculating the parameters of the atmosphere and clouds that cause the occurrence, development and intensity of hail processes is a generalization of the schemes for calculating the parameters used in existing methods for predicting convection and related weather hazards, and is, in fact, an electronic aerological diagram [8].

To predict hail, this algorithm uses the output of the global atmospheric model as input data with a lead time of 12 hours to five days: air temperature and dew points, wind direction and speed at standard levels (Earth, 800, 700, 650, 600, 500, 400, 300 and 200 hPa). About 45 atmospheric parameters are preliminarily calculated, which are known from the existing methods of forecasting convection and associated dangerous weather phenomena. Such a number of features imposes too stringent (often unrealizable) requirements for the amount of empirical data. Therefore, it became necessary to select the most informative of them signs. This problem was solved in two ways: using the biserial correlation coefficient and factor analysis taking into account the biserial correlation coefficient [8-12]. Then, using the selected parameters, discriminant functions are compiled to predict hazardous events (hail).

The discriminant function obtained on the basis of the parameters selected using the biserial correlation coefficient has the form:

\[
L = 0.0557D_j - 0.0696 \sum_{j=1}^{500} q - 0.0002 H_p - 0.036 TT - 0.0069\Delta S + 0.057 \tau_{dpd} - 20.3, \quad (1)
\]

where
- \(D_j\) – George’s instability index;
- \(H_p\) – height of the layer of potential instability;
- \(TT\) – index of Miller’s integral sum;
- \(\sum_{j=1}^{500} q\) – summary specific humidity in the Earth layer – 5 km;
- \(\Delta \tau_{dpd}\) – summary dew point deficit in a 5 km layer from the condensation level;
- \(\Delta S\) – energy characteristic of the subcloud layer.

According to the discriminant function (1), when \(L \geq 0\) «hail» is predicted and \(L < 0\) – «without hail».

The discriminant function obtained on the basis of the parameters selected using factor analysis and the biserial correlation coefficient has the form:

\[
L_2' = -0.05 V_{700} + 0.3205 \Delta t_m + 0.0071 \Delta t_1 + 0.00359 \sum_{j=1}^{500} q - 0.03606 \sum_{j=1}^{500} \tau + 0.0389 D_j - 12.57, \quad (2)
\]

where
- \(V_{700}\) – ordered vertical air movements at 700 hPa;
- \(\Delta t_m\) – maximum temperature difference between the cloud and the surrounding air;
- \(\Delta t_1\) – vertical temperature gradient in the layer above the condensation level by 2 km;
$\sum_{Earth}^{500} q$ – summary specific humidity in the layer from the Earth's surface up to 5 km height;

$D_j$ – George's index;

$0.03606 \sum_{Earth}^{500} \tau$ – summary dew point deficit in the layer from the Earth's surface up to 500 hPa.

According to the discriminant function (2), «hail» is also predicted at $L_2^* \geq 0$, and at $L_2^* < 0$ is predicted «without hail».

Note that the parameters included in the discriminant functions mainly reflect the margin of instability, moisture content and temperature regime in the hail growth zone. A distinctive feature of the discriminant function (2) is that it allows one to quantitatively take into account large-scale atmospheric movements (ordering of vertical currents), against the background of which convective phenomena develop. This is the reason for the choice of this function for further calculations.

To assess the success of the proposed hail forecasting method, a contingency table is drawn up with the following elements:

- $n_{11}$ – the number of correct forecasts with the formulation «hail»;
- $n_{12}$ – the number of unjustified forecasts with the formulation «hail»;
- $n_{21}$ – the number of unjustified forecasts with the formulation «without hail»;
- $n_{22}$ – the number of correct forecasts with the formulation «without hail»;
- $n_{10} = n_{11} + n_{12}$ – the number of all estimated forecasts with the formulation «hail»;
- $n_{20} = n_{21} + n_{22}$ – the number of all estimated forecasts with the formulation «without hail»;
- $n_{01} = n_{11} + n_{21}$ – the number of all assessed cases «hail»;
- $n_{02} = n_{12} + n_{22}$ – the number of all assessed cases «without hail»;
- $n_{00} = n_{10} + n_{20} = n_{01} + n_{02}$ – the number of all assessed cases.

Based on the contingency table, the main parameters are calculated for assessing hail hazard forecasts:

$$P_{gaf} = (n_{11} + n_{22})/100,$$  
$$P_{wpp} = (n_{11} + n_{01})/100,$$

where $P_{gaf}$ – where is the general accuracy of the forecast;

$P_{wpp}$ – warning of the presence of the phenomenon.

Additional indicators are calculated for assessing hail hazard forecasts:

$$P_{ppop} = \frac{n_{11}}{n_{10}} \cdot 100,$$  
$$P_{jap} = \frac{n_{22}}{n_{20}} \cdot 100,$$  
$$P_{wap} = \frac{n_{22}}{n_{02}} \cdot 100,$$

where $P_{ppop}$ – predictability of the prediction of the occurrence of the phenomenon;

$P_{jap}$ – justification of the absence of the phenomenon;

$P_{wap}$ – warning of the absence of a phenomenon.

The criterion for the quality of the Pearcy - Obukhov forecasts is calculated by the formula:

$$T = \frac{n_{11}}{n_{01}} \frac{n_{12}}{n_{02}} \frac{n_{22}}{n_{02}} - \frac{n_{21}}{n_{01}}.$$
Reliability criterion according to N.A. Bagrov looks like:

\[ H = \frac{U - U_0}{1 - U_0}, \]

where

\[ U_0 = \frac{(n_{01}n_{02} + n_{20}n_{02})}{n_{00}}; \quad U = \frac{n_{11} + n_{22}}{n_{00}}. \]

The risk error of the method (the phenomenon was not predicted, but was observed) is equal to:

\[ a = \frac{n_{21}}{n_{01}}. \]

The method insurance error (the phenomenon was predicted, but not observed) is equal to:

\[ P = \frac{n_{12}}{n_{02}}. \]

Accuracy criterion no A.M. Obukhov takes on the meaning:

\[ Q = 1 - a - P. \]

3. Results and discussion

Based on the high values of the correlation coefficients between the data of actual upper-air sounding and the corresponding predicted values according to the global model with increasing lead time [4], in this work, for the forecast of hail, the values of meteorological parameters were used according to the data of the global atmospheric model.

The hail forecast was compiled using discriminant functions for May-August 2020 according to the geographical coordinates of the «Mineralnye Vody» station. To assess the success of the forecast, a contingency table was compiled (table 1), the data of which made it possible to calculate the criteria for the quality of the hail forecast (table 2).

Table 1. Contingency table with increasing lead times.

| Forecast                | Observation of the phenomenon | Total |
|-------------------------|-------------------------------|-------|
|                         | Hail                         | Without hail |
| Lead time, 24 hours     | 26                           | 7     | Hail |
|                         | 6                            | 16    | Without hail |
| total                   | 32                           | 23    | total |
| Lead time, 72 hours     | 14                           | 7     | Hail |
|                         | 5                            | 13    | Without hail |
| total                   | 19                           | 20    | total |
| Lead time, 132 hours    | 10                           | 5     | Hail |
|                         | 6                            | 18    | Without hail |
| total                   | 16                           | 23    | total |
The division of phenomena into «hail» or «without hail» was carried out according to the observations of the paramilitary services on the active impact on meteorological and other geophysical processes located within the radius of representativeness of the actual upper-air sounding data at the «Mineralnye Vody» station (table 1).

The results of hail forecasts by discriminant functions for 24, 72 and 132 hours showed that the method under consideration meets all the criteria for the quality of forecasts. Table 2 clearly shows that the indicators of the success of the forecast with a lead time of 24 hours turned out to be high. As for the indicators for a 72-hour lead time, they turned out to be slightly lower than the indicators for hail forecasting with the same functions with a shorter lead time (24 hours). In the case of a 132-hour lead time, the forecast quality criteria turned out to be at the level or slightly higher than for the previous lead time. This is due to the fact that the calculations for this period were carried out according to the data of the global model for 2019, and as noted above, the global atmospheric model itself was improved in the same year.

Table 2. Quality criteria for hail forecast with increasing lead time.

| The name of the forecast quality criteria | Lead time, hour |
|------------------------------------------|-----------------|
|                                          | 24  | 72  | 132 |
| General accuracy of the forecast         | 77  | 69  | 72  |
| Warning of the presence of the phenomenon| 88  | 74  | 63  |
| Predictability of the prediction of the occurrence of the phenomenon | 80  | 67  | 67  |
| Justification of the absence of the phenomenon | 73  | 72  | 75  |
| Warning of the absence of a phenomenon   | 70  | 65  | 78  |
| The criterion for the quality of the Pearcy - Obukhov | 0,51 | 0,39 | 0,41 |
| The criterion of the N.A. Bagrov        | 0,52 | 0,39 | 0,41 |
| The risk error of the method            | 0,18 | 0,26 | 0,38 |
| The method insurance error              | 0,30 | 0,35 | 0,22 |
| Accuracy criterion no A.M. Obukhov      | 0,51 | 0,39 | 0,41 |

Note that the justification of the proposed method is equal to ≈70% for a lead time of 72 hours, in contrast to the similar forecast method for 24 hours (≈77%) and for 132 hours (≈72%). Thus, the criteria for the quality of forecasts, in particular, hail, with increasing lead time, meets all criteria for the quality of forecasts.

4. Conclusion
The criterion for the quality of forecasts of hail with increasing lead time meets all the criteria for the quality of forecasts.

Thus, the predicted values of atmospheric stratification can be used to make predictions of hazardous phenomena and weather elements (replacing the actual upper-air sounding), which cannot yet be predicted by the global models themselves. In this case, the lead time of forecasts can be brought to the periods corresponding to the medium-term periods of the weather forecast.

References
[1] Kagermazov A Kh and Sozaeva L T 2019 Validation of the output data of the global atmospheric model on days with the development of convection according to the data of aerological sounding with a two-day lead time Proc. of Voeikov Main Geophysical Observatory 593 150–159
[2] Kagermazov A Kh and Sozaeva L T 2020 Validation of the output data of the global atmospheric model, obtained with a lead time of up to three days, according to the data of upper-air sounding Materials of the VI All-Russian Scientific Conference «Problems of
military-applied geophysics and control of the state of the natural environment» (16-18 September St. Petersburg: A. F. Mozhaikii VKA) p 131

[3] Kagermazov A Kh and Sozaeva L T 2020 Validation of prognostic fields of meteorological elements of the global atmospheric model for average periods according to aerological sounding data for the Central part of the North Caucasus Proc. of Voeikov Main Geophysical Observatory 599 104–114

[4] Kagermazov A Kh and Sozaeva L T 2020 Validation of the output of the global atmospheric model from upper-air sounding data with increasing lead time Nauka. Innovacii. Tehnologii. 4 137–148

[5] Kalnay E, Kanamitsu M and Baker W E 1990 Global numerical weather prediction at the National Meteorological Center. Bull Amer. Meteor. Soc. 71 1410–28

[6] Kanamitsu M 1989 Description of the NMC global data assimilation and forecast system Wea. and Forecasting 4 335–342

[7] Kanamitsu M at al. 1991 Recent changes implemented into the global forecast system at NMC Wea. and Forecasting 6 425–435

[8] Kagermazov A Kh 2015 Digital atmosphere. Modern methods and methodology for the study of hazardous meteorological processes and phenomena (Nalchik: Pechatnyy dvor) p 215

[9] Kagermazov A Kh 2012 Hail Prediction from Global Atmosphere Model Output (T254 NCEP) Meteorology and Hydrology 3 28–34

[10] Kagermazov A Kh and Sirotenko L A 2010 Results of approbation of software modules for calculating meteorological parameters and computer forecasting of hail according to aerological sounding data in the North Caucasus Proceedings of Higher Educational Institutions. North Caucasian region. Physics of the atmosphere. Natural Sciences Special issue 49–53

[11] Kagermazov A Kh and Sozaeva L T 2020 Hail forecast with a lead time of up to three days according to the output of the global atmospheric model Proc. of Voeikov Main Geophysical Observatory 598 204–214

[12] Kagermazov A Kh, Fedchenko L M, Sozaeva L T and Zhaboeva M M 2021 Medium-term hail forecast based on the output of the global atmospheric model Nauka. Innovacii. Tehnologii. 2 91–106