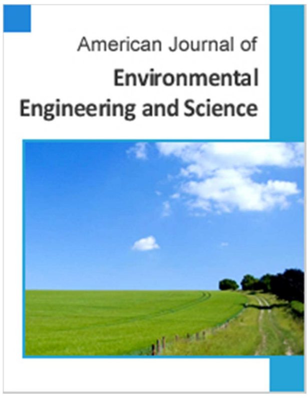
American Journal of Environmental Engineering and Science
2017; 4(6): 71-77
http://www.aascit.org/journal/ajees
ISSN: 2381-1153 (Print); ISSN: 2381-1161 (Online)

American Association for Science and Technology



On Statistical Forecasting of the Atmospheric Temperature for Month

Vladimir Stepanovich Mukha

Department of Automated Data Processing Systems, Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Email address
mukha@bsuir.by

Citation
Vladimir Stepanovich Mukha. On Statistical Forecasting of the Atmospheric Temperature for Month. *American Journal of Environmental Engineering and Science.*
Vol. 4, No. 6, 2017, pp. 71-77.

Keywords
Statistical Weather Forecast, Numerical Weather Forecast, Extrapolation of a Random Sequences, Forecast of the Atmospheric Temperature

Received: September 7, 2017
Accepted: November 16, 2017
Published: December 7, 2017

Abstract
Irregular nature of the meteorological data requires obviously the probability-statistical data processing methods, but these methods didn't find the significant application in meteorology. In this paper we investigate statistical algorithm based on Kolmogorov's theory random sequences extrapolation for forecasting of atmospheric temperature. On the many real forecasts we compare the accuracy of statistical forecast with deterministic numerical and simple climatological forecasts when forecasting is performed for a month ahead. This analysis shows advantages and disadvantages as statistical as numerical forecasts.

1. Introduction

At present enormous amount of data about actual weather condition is accumulated in the world. It seems that these data can and must be used for estimation of variability of the weather in the course of time and for forecasting of the weather. The acquaintance with available data shows that they have an irregular nature. We have the need to build the mathematical models of such data and decision on this base the constructive problems of weather analysis and forecasting. Now we can used the probabilistic-statistical theory of irregular data that is well designed and not have a convincing alternative. (The theory of the fuzzy sets can't serve as alternative to the probabilistic-statistical theory because of its insufficient development). However it seems that both this theory and accumulated data are used in insufficient degree on elementary level even. It is proposed to use the statistical methods for serving the numerical weather forecast models [1]. There are a few examples of "pure" application of statistical methods in weather forecasting [2, 3], but they are difficult to reproduce for comparison because of different formulations, methods, regions, insufficiently detailed description of the algorithms and their parameters, etc. In this article the meteorological data are used for calculation and analysis of the statistical forecast of the atmospheric temperature within the framework of existing probabilistic-statistical mathematical models [4].

In works [5, 6] the empirical analysis of accuracy of the statistical forecast of the atmospheric temperature at the meteorological station Minsk for 13 days ahead in comparison with accuracy of the climatological forecast and the numerical forecast Gismeteo [7] was executed. In this work, we execute the analysis of accuracy of the statistical forecast for month ahead.

2. Empirical Analysis of Accuracy of Statistical Forecasting of the Atmospheric Temperature for Month

For forecasting of the atmospheric temperature was used algorithm offered in [8] with improvements given in works [9, 10]. The parameters of the forecasting algorithm was calculated on the data for previous history not less than for 11 years in each 3 hours. For forecasting was used data for previous history in 224 measurements in each 3 hours, and forecast was executed also on depth in 224 measurements (the forecast 224/224 in numbers of measurements, or 672/672 in hours, or 28/28 in days). Most forecasts were calculated on condition for nine hours UTC. Yearly forecast and month forecasts were considered. In month forecast were included forecasts executed at days given month, in yearly – all executed forecasts. The retrospective forecasting was executed for 2016 and 2017 years: at January – 31 forecasts, at February – 31, at March – 61, at April – 54, at May – 52, at June – 57, at July – 54, at August – 28, at September – 29, at October – 30, at November – 29, at December – 31, the whole 487 forecasts. For fixed depth of the forecast the average module error (a.m.e.) of the forecast $s_{a,i}$ for the maximum temperature daytime and minimum in the night was calculated:

$$s_{a,i} = \frac{1}{k}\sum_{j=1}^{k} | t_{i,j} - t_j |, \; i = \overline{1,52}, \quad (1)$$

where $t_{i,j}$ – predicted value of the temperature at the time j, t_j – actual value of the temperature at the time j; k – a number of executed forecasts for considered period, i – a depth of the forecast with unit of the measurement in 12 hours. By expression (1) with $i = \overline{1,52}$ we calculated a.m.e. of statistical, climatological and Gismeteo forecasts. (The Gismeteo forecast is a numerical forecast, using, on statement of its authors, one of the best predictive models [7]. Gismeteo is Russian meteorological internet portal, http://www.gismeteo.ru. The climatological forecast is a forecast on the average temperature over many years).

The graphs of a.m.e. of yearly forecast calculated on formula (1) with $k = 487$ on the grounds of forecasts for 2016–2017 years are presented in the figure 1. The graphs of a.m.e. are very irregular. For the simplification of the comparison, a cubic approximations using the least squares error method was performed (LSE-approximations). The cubic LSE-approximations of the a.m.e. in figures 1, 2, 3 are smooth. The square marker in figures 1, 2, 3 points to midday. We see that statistical forecast more accurate than climatological if forecasting less than on 5 days, and they are approximately equivalent in accuracy if forecasting more than on 5 days. At the same time the numerical forecast Gismeteo is less accurate than statistical and climatological forecasts if forecasting more then on 8 days.

On the whole, climatological forecast should be considered as a baseline for evaluating of other forecasts. Climatological forecast is very easy, cheap and easily implemented, so any less accurate forecast should be considered as bad. As we can see in figure 1, the numerical Gismeteo forecast with depth of forecasting more than 8 days is less accurate than climatological forecast.

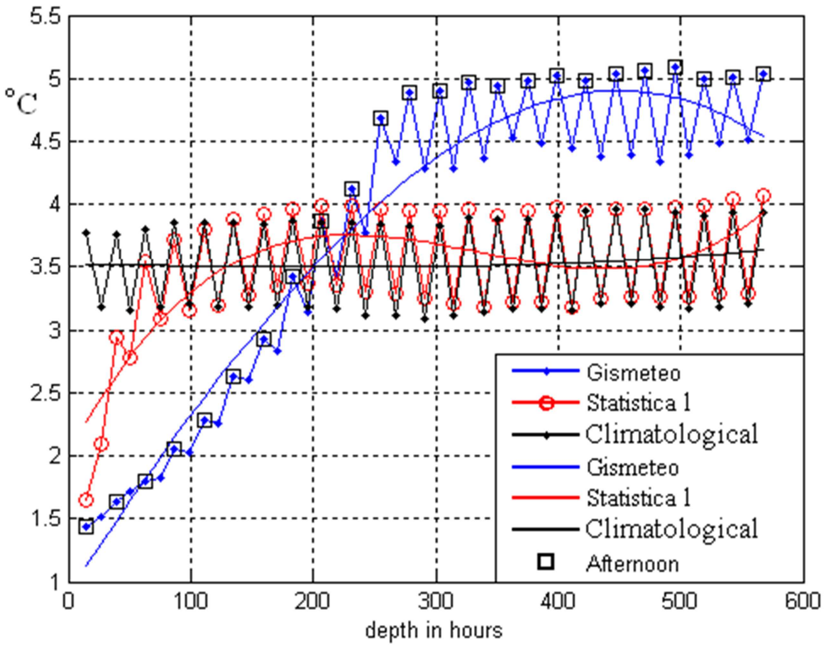


Figure 1. *Graphs of a.m.e. of the yearly forecasts of the temperature.*

Graphs of a.m.e. for separated months are presented in figures 2, 3. Statistical forecast exceeds climatological forecast in accuracy under short-time forecasting and is equivalent it under long-time forecasting. Gismeteo forecast obviously loses statistical and climatological forecasts under long-term forecasting. This reveals itself especially clearly at January, February, March, April, May, August, September, October, November and December. Gismeteo forecast may be worse than climatological forecast already at depth of the forecast 5 days (refer to January, November).

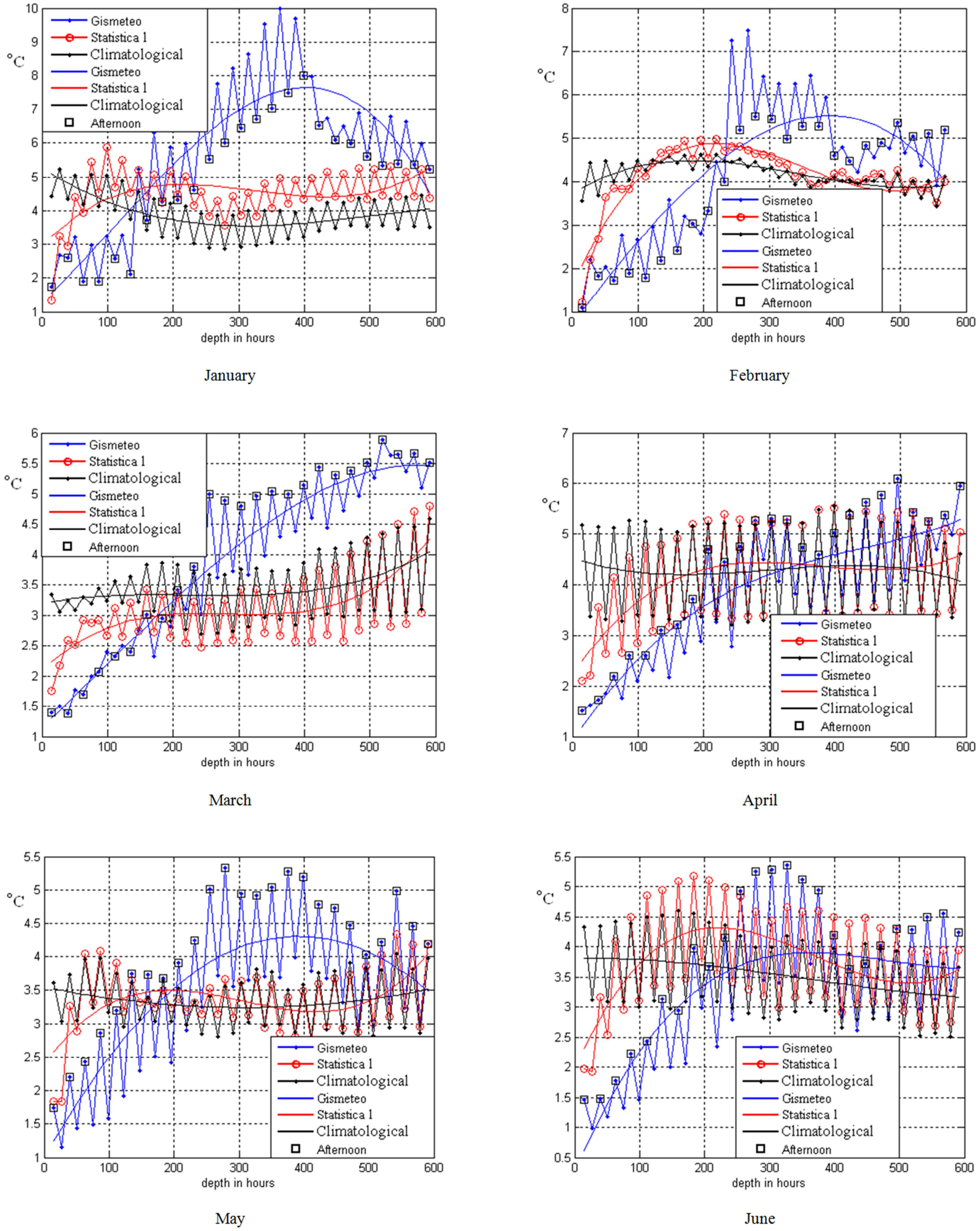


Figure 2. *Graphs of a.m.e. of the monthly forecasts of the temperature (January – June).*

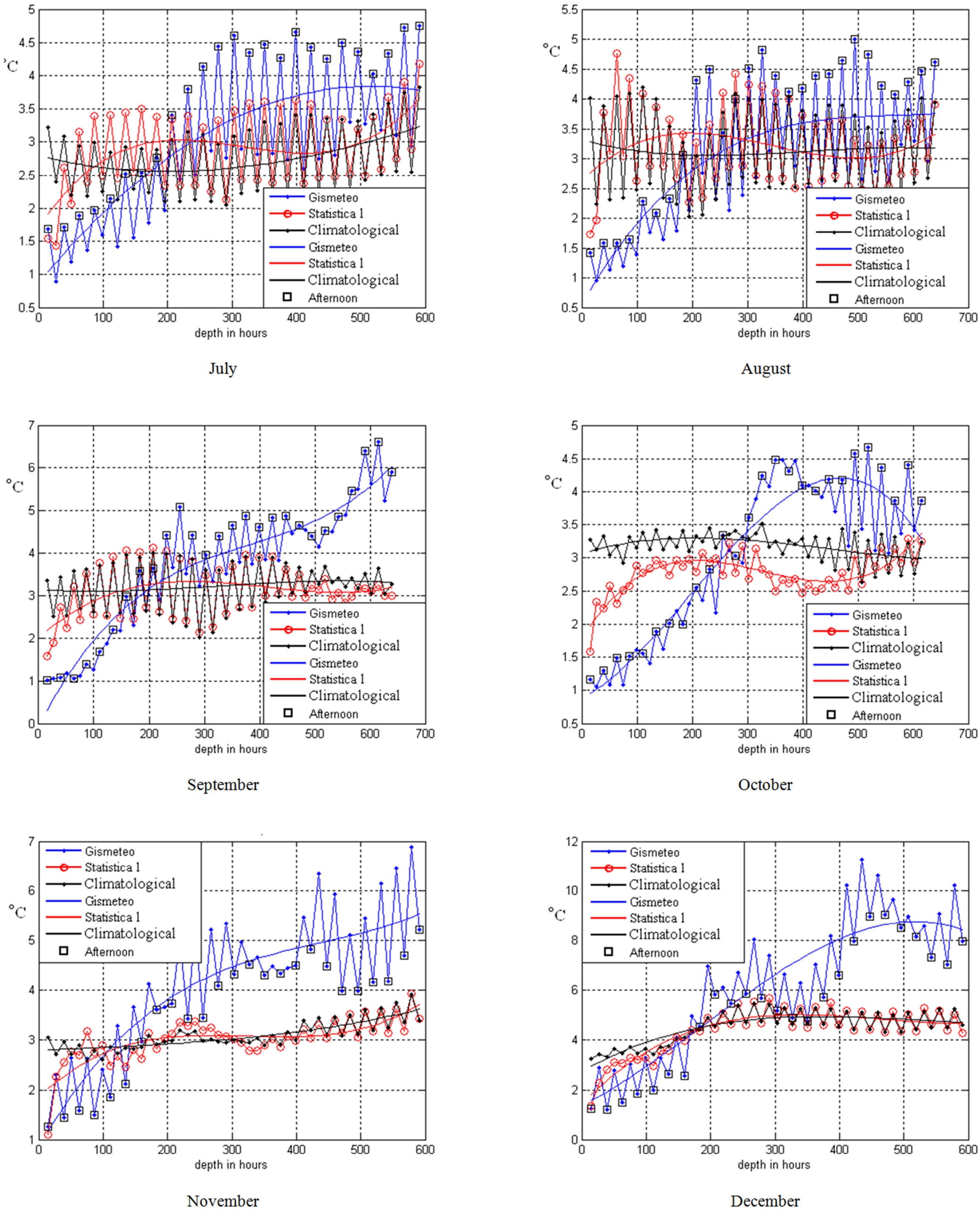


Figure 3. *Graphs of a.m.e. of the monthly forecasts of the temperature (July – December).*

3. Analysis of Separate Forecasts

Analysis of separate forecasts allows to do qualitative conclusions about the nature of different forecasts of the temperature. We will show in the figures 4–8 the most salient illustrations. The real motion of the temperature in this figures is showed in black color, Gismeteo forecast – in blue color and statistical forecast – in red color.

Figure 4 illustrates inaccuracy of the forecast Gismeteo at forecasting on not so far of depth of the time. 22.02.2017 Gismeteo predicts on the period 04–16.03.2017 negative values of temperature up to -20°C, but real temperature in this period is positive and reaches +10°C. Gismeteo insists on this forecast also 23.02.2017 and 24.02.2017 and changes its decision only 25.02.2017. Statistical forecast in this case is more appropriate.

Gismeteo is characterized by a radical changes of its own predictive decisions. For example, in the figure 5a is presented the forecasts 31.12.2015 where Gismeteo predicts significant decrease of temperature up to -27°C on 06–10.01.2016. For example, in the figure 5a is presented the forecasts 31.12.2015 where Gismeteo predicts significant decrease of the temperature up to -27°C on 06–10.01.2016. But in forecast 01.01.2016 (figure 5b) this decision was changed in favor of positive values of temperature on the same period. Statistical forecast in this case consistently points on increase of temperature.

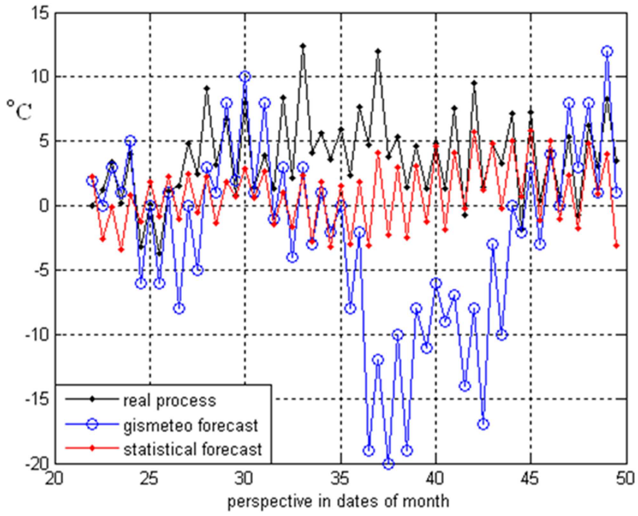
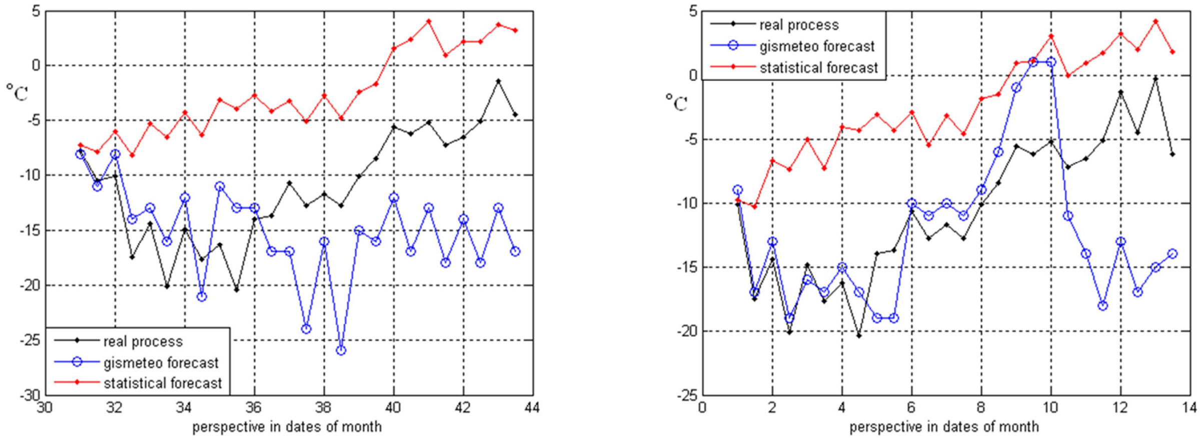


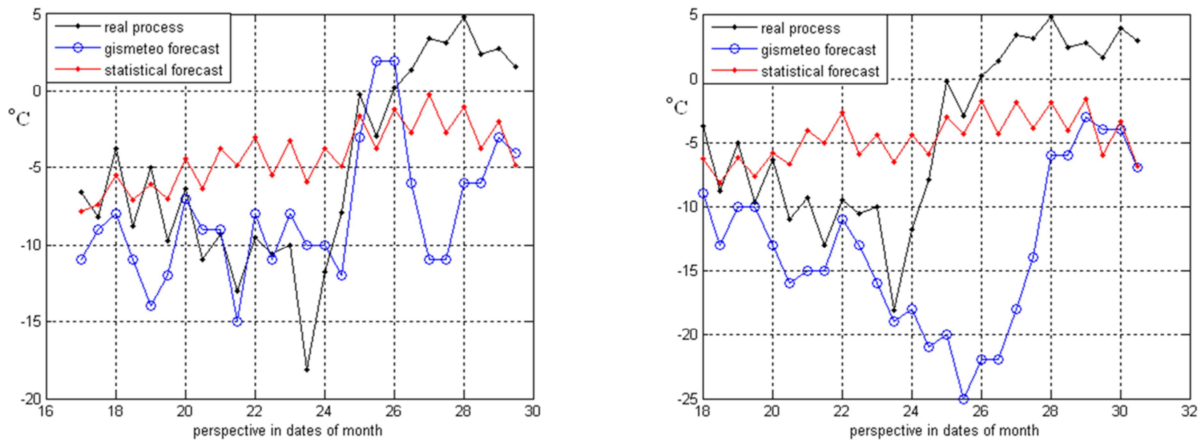
Figure 4. *Forecasts of the temperature 22.02.2017 at 09:00.*



a b

Figure 5. *Forecasts of the temperature 31.12.2015 (a) and 01.01.2016 (b) at 09:00.*

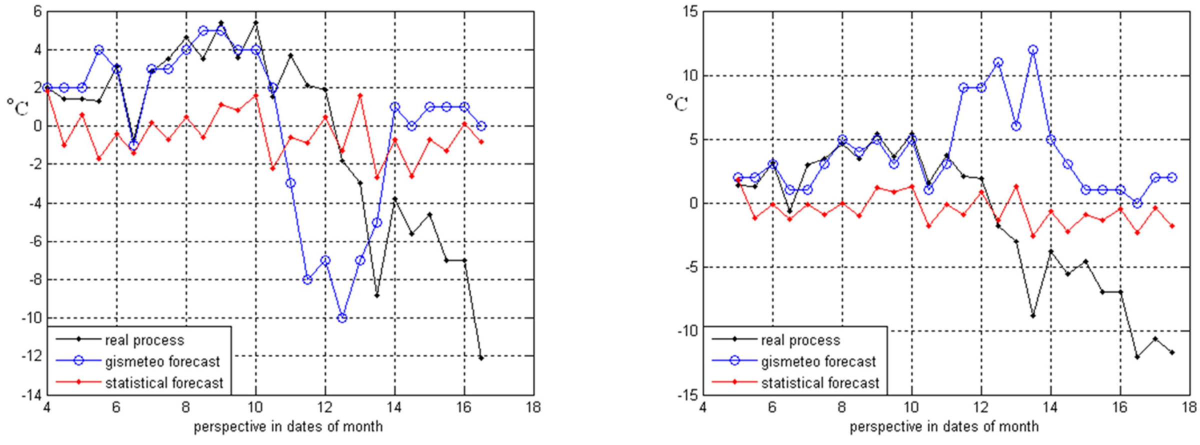
Radical change of decisions of Gismeteo is illustrated by forecasts 17.01.2016 and 18.01.2016 (figure 6) too. Forecast Gismeteo 18.01.2016 on the period after 24.01.2016 (figure 6b) differs significantly from the forecast 17.01.2016 on the same period, and in the worst way. Statistical forecast don't change its nature and indicates the actual temperature trend.



a b

Figure 6. *Forecasts of the temperature 17.01.2016 (a) and 18.01.2016 (b) at 09:00.*

Forecasts 04.01.2014 and 05.01.2014 in the figure 7 show sudden Gismeteo solution change. Forecast 04.01.2014 on 11–13.01.2014 (figure 7a) predict decrease of temperature up to -10°C, but forecast 05.01.2014 on the same period (figure 7b) predict its increase up to +12°C.

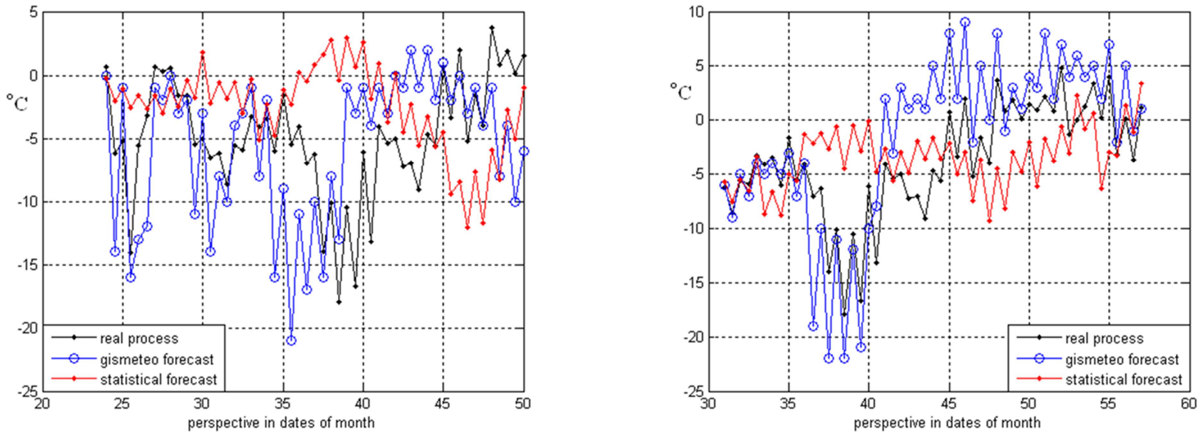


a b

Figure 7. *Forecasts of the temperature 04.01.2014 (a) and 05.01.2014 (b) at 09:00.*

We illustrate inaccuracy of short-term statistical forecast of temperature by figure 8. Forecast Gismeteo on 25.01.2017 performed 24.01.2017 (figure 8a) is much more accurate than statistical forecast. In addition, forecasting 24.01.2017 (figure 8a) and 31.01.2017 (figure 8b) show that statistical forecast haven't tracked short duration decrease of temperature on 05–10.02.2017. However, forecasts Gismeteo in the figure 8 can't be called good. So in the figure 8a the negative temperatures predicted by Gismeteo is bad justified both in time and values. Forecast Gismeteo in the figure 8b is bad justified in values.

Analysis of this and other forecasts evidence that statistical forecast doesn't track of sudden change of temperature and it lead to increase of errors of short-time forecasts. The advantage of statistical forecast consist in its balance (high inertia) what allows to correctly evaluate the trend of evolution of process. Numerical forecast Gismeteo allows to track of sudden change of weather, but this can led to big errors especially for long-term forecasts.



a b

Figure 8. *Forecasts of the temperature 24.01.2017 (a) and 31.01.2017 (b) at 09:00.*

4. Conclusion

The data presented in this paper allow to do certain conclusions about accuracy of linear statistical forecast of atmospheric temperature in comparison with accuracy of other forecasts. Short-time statistical and numerical forecasts (up to 6–8 days) are more accurate than climatological forecast for any month of year. Numerical forecast Gismeteo is more accurate than statistical forecast on $0-1{,}5\,^{\circ}C$ for depth of forecast up to 6–8 days. At the same time both statistical and climatological forecasts exceed the numerical forecast Gismeteo in accuracy when forecasting is performed on more than 6–8 days. Therefore the numerical forecast of the atmospheric temperature cannot be considered as consistent for depth of forecast more than 6–8 days. This conclusion is in a good agreement with estimates of work [11]. As regards statistical forecast, it can be implemented into local automatic meteorological stations due to its simplicity and enough high accuracy.

References

[1] Vilfand, R. M., Tischenko, V. A., Khan, V. M. The Statistical forecast of the move of the temperature of the air inwardly month with use output data of hydrodynamic models. Meteorology and hydrology. No 3. Pp. 5-13. 2007. In Russian.

[2] Malone, T. F. Application of statistical methods in weather prediction. *Proceedings of the National Academy of Sciences USA*. – Vol. 41.– Pp. 806-815. 1955.

[3] Campbell, S. D., Diebold, F. X. Weather Forecasting for Weather Derivatives. Journal of the American Statistical Association. Vol. 100. No 469. Pp. 6-16. 2005.

[4] Kolmogorov, A. N. Interpolation and extrapolation of stationary random sequences. Proceedings of the Academy of Sciences of USSR. Mathematical series. Vol. 5. Pp. 3-14. 1941. In Russian.

[5] Mukha, V. S., Trofimovich, A. F. Numerical empirical analysis of accuracy of linear statistical forecasting of atmospheric temperature. Doclady BGUIR. No. 8 (62). Pp. 14-21. 2011. In Russian.

[6] Mukha, V. S. Experience in Statistical Forecasting of the Atmospheric Temperature. Open Journal of Atmospheric and Climate Change. – In Press. Pub. Date (Web): 2015-10-06. http://www.scipublish.com/journals/ACC/ Papers in Press. -10 p.

[7] Gismeteo: http://www.gismeteo.ru/

[8] Mukha, V. S. Statistical vector forecasting of the quantitative features of the weather. Information systems and technologies (IST'2004). Materials of International conference (Minsk, November 8–10, 2004.). Part 2. Pp. 195-200. 2004. In Russian.

[9] Mukha, V. S., Trofimovich, A. F. Estimation of mathematical expectation and covariance function of stationary random sequence by averaging over time and many realizations. Doclady BGUIR. No. 1 (39). Pp. 93-99. 2009. In Russian.

[10] Mukha, V. S., Kozyachy, A. N. Multidimensional model of data for analytical processing. Doclady BGUIR. No 1 (47). Pp. 100-105. 2010. In Russian.

[11] N. Gustafsson. Statistical issues in weather forecasting. Scandinavian Journal of Statistics. V. 29. No. 2. Pp. 219-239. 2002.