

АНАЛИЗ ТОЧНОСТИ СТАТИСТИЧЕСКОГО ПРОГНОЗИРОВАНИЯ ТЕМПЕРАТУРЫ АТМОСФЕРНОГО ВОЗДУХА

В.С. МУХА

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь

Поступила в редакцию 26 мая 2017

Аннотация. Приводятся и анализируются эмпирические данные о точности линейного статистического прогноза температуры атмосферного воздуха в сравнении с точностью климатического и численного прогнозов.

Ключевые слова: статистический прогноз погоды, численный прогноз погоды, экстраполирование случайных последовательностей, прогноз температуры атмосферного воздуха.

Abstract. Empirical data about precision of linear statistical forecast of the atmospheric temperature in comparison with precision of the climatological and numerical forecasts is given and analyzed.

Keywords: statistical weather forecast, numerical weather forecast, extrapolation of a random sequences, forecast of the atmospheric temperature.

Doklady BGUIR. 2017, Vol. 108, No. 6, pp. 22-28

Analysis of accuracy of statistical forecasting of the atmospheric temperature

V.S. Mukha

Введение

Анализ литературных источников показывает, что в проблеме прогнозирования погоды доминирует детерминированный подход. Основу такого подхода составляют гидродинамические модели атмосферы в виде дифференциальных уравнений, описывающих атмосферные процессы. Ввиду того, что реализация прогностических моделей производится в основном численными методами, этот подход получил название численного прогноза погоды (Numerical Weather Prediction, NWP) [1, 2]. Численный прогноз является достаточно успешным, в связи с чем современные службы погоды полностью доверяются его моделям.

Параллельно существует определенный уровень интереса и внимания к статистическим методам прогнозирования погоды [3, 4], в том числе и со стороны отдельных профессиональных прогнозистов [5]. Однако, как отмечается в [5], «требуется определенное мужество, чтобы выступать за применение статистических методов, поскольку многие синоптические метеорологи рассматривают сегодня статистику с пренебрежением». Эта ситуация во многом сохраняется и в настоящее время, хотя ответственность специалистов должна состоять в том, чтобы видеть, что «ничто из инструментальных средств и методов современной науки не проигнорировано в попытках решения проблемы» [5].

В течение последних лет профессиональные специалисты в области прогнозирования погоды приходят к заключению, что, возможно, детерминированный подход достиг своих пределов предсказуемости, и мы видим тенденцию к более широкому использованию статистических методов [2]. Одно из направлений использования статистических методов связано с обслуживанием моделей численного прогноза погоды. Сюда относятся статистическая постобработка выхода модели, усвоение метеорологических данных, ансамблевое прогнозирование и др. [2, 6].

Имеются также примеры «независимого», или «чистого», применения статистических методов в прогнозировании погоды [5, 7]. Однако они трудно воспроизводимы для сравнения в силу того, что отличаются постановками задач, прогнозируемыми переменными, методами решения, рассматриваемыми регионами, недостаточно детальным описанием используемых алгоритмов и их

параметров и т. д. Имеющиеся в статистической литературе чисто теоретические разработки чаще всего не доводятся до практического применения и сравнения (до алгоритмов прогнозирования и оценок их параметров, расчетов применительно к метеорологии, сравнительного анализа результатов расчетов).

В данной работе рассматривается «чистое» применение статистического подхода как задачи экстраполирования случайной последовательности [8]. В работе [9] был предложен алгоритм прогнозирования количественных характеристик погоды как задачи экстраполирования векторной гауссовской случайной последовательности. В [10] на основе данного алгоритма выполнен эмпирический анализ точности статистического прогнозирования температуры атмосферного воздуха. С тех пор были получены новые теоретические и практические результаты. В [11] выполнено обоснование применимости предложенного в [9] алгоритма для произвольной (не гауссовской) случайной последовательности. Предложен и реализован в программном обеспечении более совершенный алгоритм оценивания математического ожидания и ковариационной функции стационарной случайной последовательности усреднением по времени и множеству реализаций [12]. Предложена и реализована многомерная модель метеорологических данных [13], позволяющая продвигаться по оси времени с минимальной дискретностью 3 часа и тем самым получать более адекватные оценки параметров алгоритма прогнозирования. Получены новые метеорологические данные, позволившие улучшить оценки параметров алгоритма прогнозирования за счет увеличения размера обучающей выборки. Эти модификации способствовали повышению точности статистического прогнозирования и обоснованности выводов.

Эмпирический анализ точности прогнозирования температуры атмосферного воздуха

Количественные характеристики погоды (температура атмосферного воздуха, атмосферное давление, направление и скорость ветра, относительная влажность воздуха и некоторые другие) измеряются на метеостанциях с периодичностью в три часа начиная с нуля часов по Гринвичу. Эти измерения можно рассматривать как реализации векторной случайной последовательности. Предложенный в [9, 10] алгоритм был использован для прогнозирования температуры атмосферного воздуха на метеостанции 26850 Минск. Температура атмосферного воздуха в течение определенного периода предполагалась стационарной по ковариационной функции и нестационарной по математическому ожиданию случайной последовательностью. Параметры алгоритма прогнозирования рассчитывались по данным предыстории не менее чем за 9 лет. Для прогнозирования использовалась предыстория в 112 отсчетов, и прогнозирование выполнялось также на глубину в 112 отсчетов. Такой прогноз будем обозначать как 112/112, что соответствует прогнозу 336/336 в часах или 14/14 в сутках. Большинство прогнозов рассчитывалось по состоянию на девять часов по Гринвичу. Результаты прогнозирования группировались по месяцам. В отдельный месяц включались результаты прогнозов, выполненных в дни данного месяца. Было выполнено ретроспективное прогнозирование за 2011, 2012, 2013 и 2014 годы: в январе – 107 прогнозов, в феврале – 103, в марте – 107, в апреле – 106, в мае – 97, в июне – 103, в июле – 102, в августе – 72, в сентябре – 81, в октябре – 83, в ноябре – 82, в декабре – 86. Всего выполнено 1278 прогнозов. Это позволило получить фактическую ошибку прогнозирования. Для фиксированной глубины прогноза рассчитывалась средняя арифметическая по модулю ошибка (с.м.о.) статистического прогноза:

$$s_{\text{ст},i} = \frac{1}{k} \sum_{j=1}^k |t_{\text{пр},j} - t_{\text{ф},j}|, \quad i = \overline{1,26}, \quad (1)$$

где $t_{\text{ф},j}$ – прогнозное значение температуры атмосферного воздуха на момент времени j ; $t_{\text{п},j}$ – фактическое значение температуры атмосферного воздуха на момент времени j ; k – число выполненных прогнозов за рассматриваемый период. С целью сравнения с доступными прогнозами Гисметео [14] с.м.о. рассчитывались не по всем 112 отсчетам, а на день и ночь (максимальная температура днем и минимальная ночью соответственно). При

прогнозе 112/112 можно рассчитать 26 дневных и ночных значений с.м.о, как это указано в формуле (1). По такой же формуле были рассчитаны соответствующие с.м.о. прогнозов Гисметео и климатических прогнозов. (Прогноз Гисметео – это численный прогноз, использующий, по утверждениям его авторов, одну из лучших прогностических моделей. Климатический прогноз – это прогноз по средней за многолетний период температуре). Рассчитанные с.м.о. представлены в виде графиков на рис. 1–3.

Графики с.м.о. представляют собой ломаные кривые с достаточно нерегулярным характером. Поэтому для удобства сравнения выполнялась их кубическая аппроксимация методом наименьших квадратов (МНК-аппроксимация). МНК-аппроксимации с.м.о. представлены плавными кривыми. Квадратным маркером отмечен полдень.

На рис. 1 представлены графики с.м.о., рассчитанных по всем 1278 прогнозам за 2011–2014 гг. (годовые с.м.о. прогнозов). С увеличением глубины прогноза с.м.о. статистического и Гисметео прогнозов растут, в то время как с.м.о. климатического прогноза остается примерно постоянной. В прогнозировании до 6 дней статистический прогноз в среднем более точен по сравнению с климатическим прогнозом, при большей глубине прогноза их точность уравнивается. В прогнозировании до 10 дней прогноз Гисметео в среднем более точен по сравнению со статистическим и климатическим прогнозами. В то же время с.м.о. статистического прогноза больше с.м.о. прогноза Гисметео не более чем на 1,5 °С. Если учесть, что бытовой термометр может иметь погрешность до 2 °С, то такую ошибку можно считать не столь существенной.

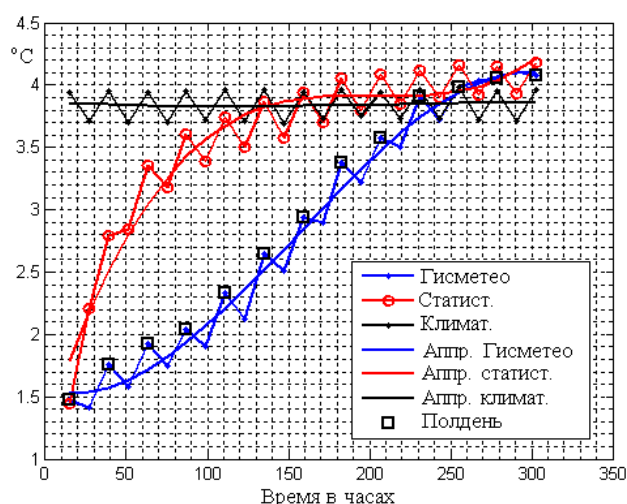
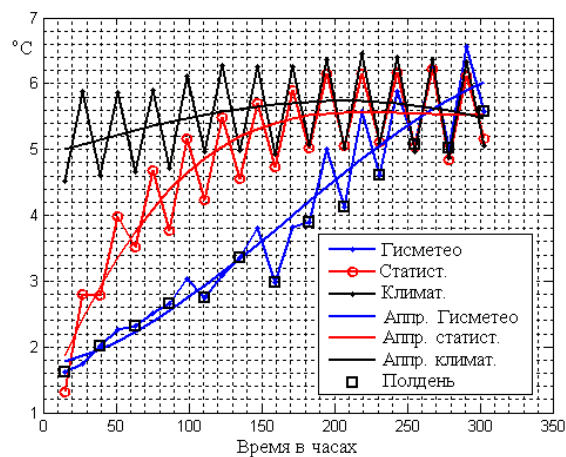
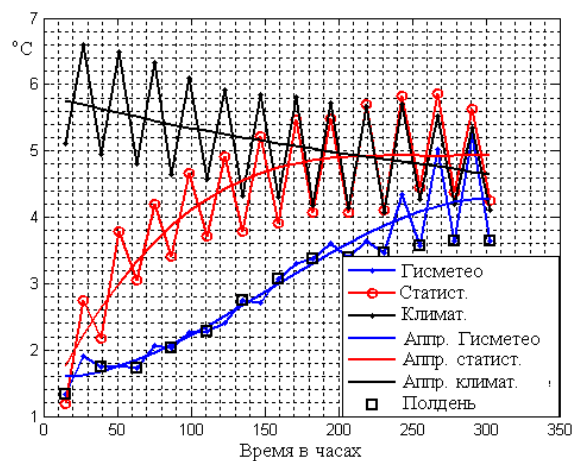


Рис. 1. Графики годовых выборочных с.м.о. прогнозов температуры

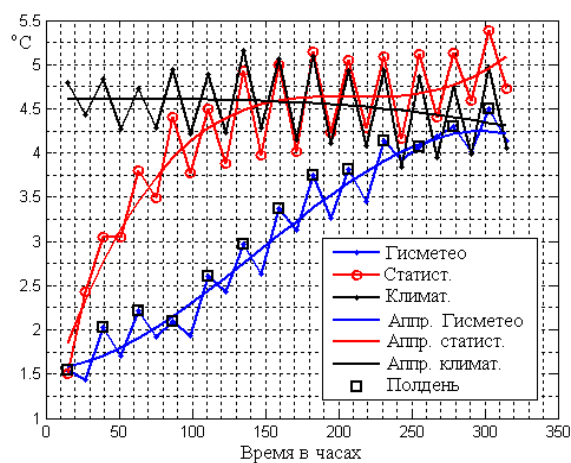
На рис. 2, 3 представлены графики с.м.о. для отдельных месяцев года. Здесь результаты не столь однозначны. Статистический краткосрочный прогноз остается лучше климатического прогноза и хуже прогноза Гисметео для всех месяцев года. В то же время для ряда месяцев достаточно явно прослеживается более низкая точность долгосрочного прогноза Гисметео, например, в январе, апреле, июне, августе, сентябре и ноябре.



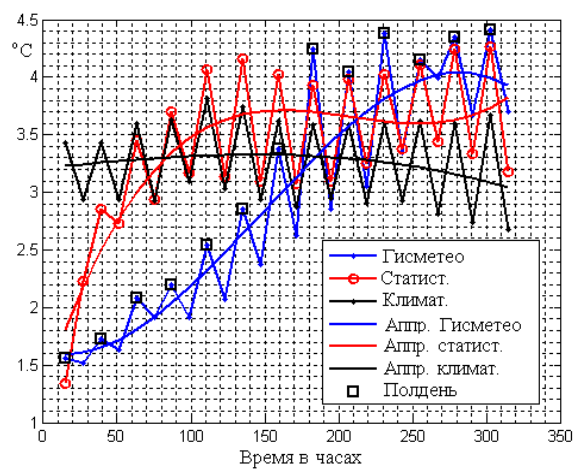
Январь



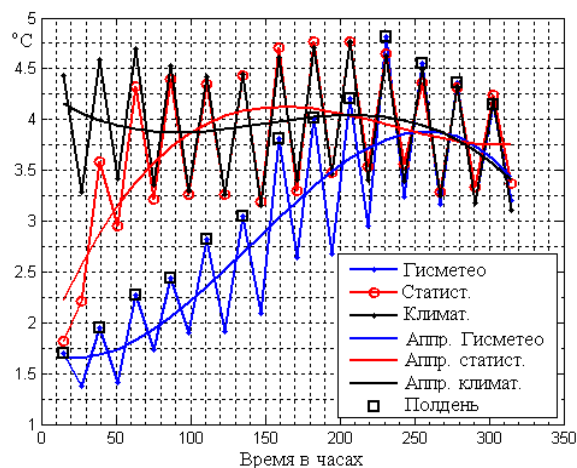
Февраль



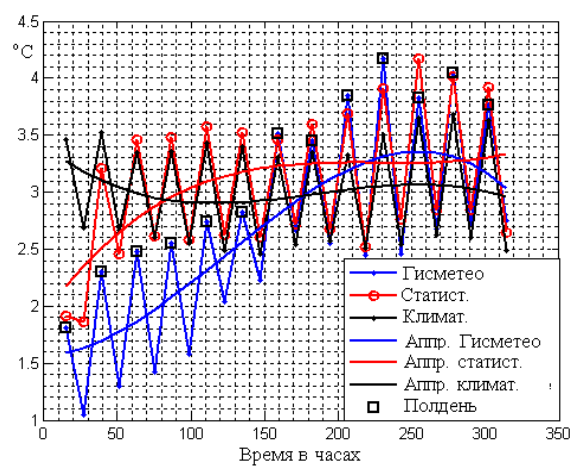
Март



Апрель



Май



Июнь

Рис. 2. Графики месячных с.м.о. прогнозов температуры (январь – июнь)

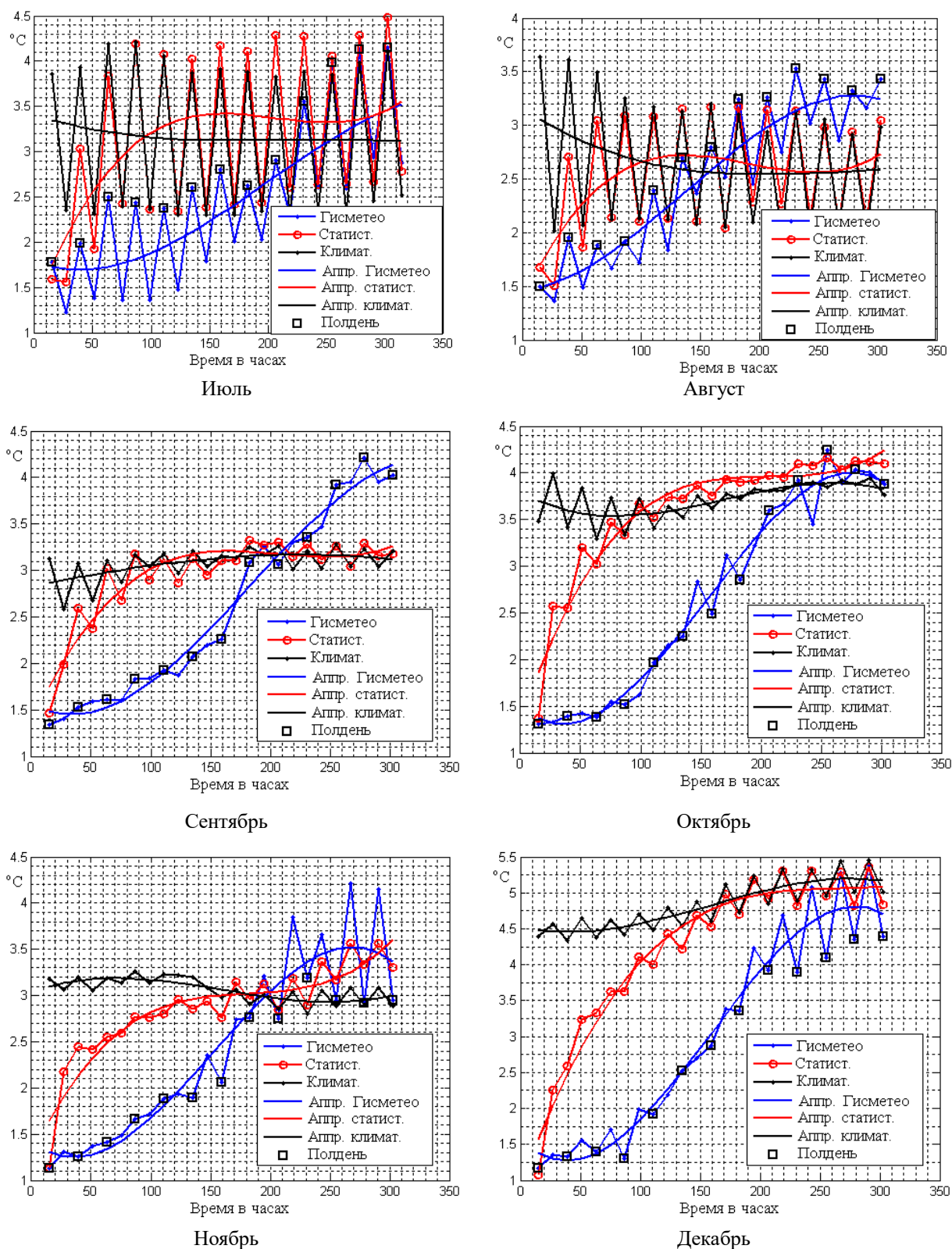


Рис. 3. Графики месячных с.м.о. прогнозов температуры (июль – декабрь)

Заключение

Представленные в статье данные позволяют сделать определенные выводы о точности линейного статистического прогноза температуры атмосферного воздуха по сравнению с точностью других прогнозов. Для любого месяца года краткосрочные (до 6–8 дней) статистический прогноз и численный прогноз Гисметео оказываются более точными по сравнению с климатическим прогнозом. Численный прогноз Гисметео более точен по сравнению со

статистическим прогнозом на 0–1,5 °С в зависимости от глубины прогноза. Вместе с тем для ряда месяцев года статистический прогноз точнее численного прогноза Гисметео при прогнозировании более чем на 6–8 дней.

С учетом положительных свойств линейного статистического прогнозирования (более высокой точности по сравнению с климатическим прогнозом, несущественного снижения точности по сравнению с численным прогнозом, невысоких требований к ресурсам и, как следствие, низкой стоимости прогноза) его можно рекомендовать к практическому использованию наряду с другими методами прогнозирования. В силу своих свойств он может быть реализован в выпускаемых в настоящее время локальных автоматических метеостанциях.

Список литературы

1. Белов П.Н., Борисенков Е.П., Панин Б.Д. Численные методы прогноза погоды. Л.: Гидрометеиздат, 1989. 376 с.
2. Gustafsson N. Statistical issues in weather forecasting // *Scandinavian Journal of Statistics*. 2002. Vol. 29, iss. 2. P. 219–239.
3. Колмогоров А.Н. К вопросу о пригодности найденных статистическим путем формул прогноза // *Журнал геофизики*. 1933. Т. 3. С. 78–82.
4. Вероятность и математическая статистика. Энциклопедия / Под ред. Ю.В. Прохорова. М.: Большая Российская энциклопедия, 1999. 910 с.
5. Malone T.F. Application of statistical methods in weather prediction // *Proceedings of the National Academy of Sciences USA*. 1955. Vol. 41. P. 806–815.
6. Вильфанд Р.М., Тищенко В.А., Хан В.М. Статистический прогноз хода температуры воздуха внутри месяца с использованием выходных данных гидродинамических моделей // *Метеорология и гидрология*. 2007. № 3. С. 5–13.
7. Campbell S.D., Diebold F.X. Weather Forecasting for Weather Derivatives // *J. of the American Statistical Association*. 2005. Vol. 100, № 469. P. 6–16.
8. Колмогоров А.Н. Интерполирование и экстраполирование стационарных случайных последовательностей // *Изв. АН СССР. Сер. математическая*. 1941. Т. 5. С. 3–14.
9. Муха В.С. Статистическое векторное прогнозирование количественных характеристик погоды // *Матер. Междунар. конф. «Информационные системы и технологии (IST'2004)»*. Минск, 8–10 ноября 2004 г. Ч. 2. С. 195–200.
10. Муха В.С., Трофимович А.Ф. Численный эмпирический анализ точности линейного статистического прогнозирования температуры атмосферного воздуха // *Докл. БГУИР*. 2011. № 8 (62). С. 14–21.
11. Муха В.С. Минимальный средний риск и эффективность оптимального полиномиального многомерно-матричного предиктора // *Кибернетика и системный анализ*. 2011. № 2. С. 121–130.
12. Муха В.С., Трофимович А.Ф. Оценивание математического ожидания и ковариационной функции стационарной случайной последовательности усреднением по времени и множеству реализаций // *Докл. БГУИР*. 2009. № 1 (39). С. 93–99.
13. Муха В.С., Козячий А.Н. Многомерная модель метеорологических данных для аналитической обработки // *Докл. БГУИР*. 2010. № 1 (47). С. 100–105.
14. Сайт Gismeteo. [Электронный ресурс]. – Режим доступа: <http://www.gismeteo.ru/>. – Дата доступа: 26.05.2017.

References

1. Belov P.N., Borisenkov E.P., Panin B.D. Chislennye metody prognoza pogody. L.: Gidrometeoizdat, 1989. 376 s. (in Russ.)
2. Gustafsson N. Statistical issues in weather forecasting // *Scandinavian Journal of Statistics*. 2002. Vol. 29, iss. 2. P. 219–239.
3. Kolmogorov A.N. K voprosu o prigodnosti najdennyh statisticheskimi putem formul prognoza // *Zhurnal geofiziki*. 1933. T. 3. S. 78–82. (in Russ.)
4. Veroyatnost' i matematicheskaya statistika. Jenciklopedija / Pod red. Ju.V. Prohorova. M.: Bol'shaja Rossijskaja jenciklopedija, 1999. 910 s. (in Russ.)
5. Malone T.F. Application of statistical methods in weather prediction // *Proceedings of the National Academy of Sciences USA*. 1955. Vol. 41. P. 806–815.
6. Vil'fand R.M., Tishhenko V.A., Han V.M. Statisticheskij prognoz hoda temperatury vozduha vnuti mesjaca s ispol'zovaniem vyhodnyh dannyh gidrodinamicheskikh modelej // *Meteorologija i gidrologija*. 2007. № 3. S. 5–13. (in Russ.)
7. Campbell S.D., Diebold F.X. Weather Forecasting for Weather Derivatives // *J. of the American Statistical Association*. 2005. Vol. 100, № 469. P. 6–16.

8. Kolmogorov A.N. Interpolirovanie i jekstrapolirovanie stacionarnyh sluchajnyh posledovatel'nostej // Izv. AN SSSR. Ser. matematicheskaja. 1941. T. 5. S. 3–14. (in Russ.)
9. Muha V.S. Statisticheskoe vektornoe prognozirovanie kolichestvennyh harakteristik pogody // Mater. Mezhdunar. konf. «Informacionnye sistemy i tehnologii (IST'2004)». Minsk, 8–10 nojabrja 2004 g. Ch. 2. S. 195–200. (in Russ.)
10. Muha V.S., Trofimovich A.F. Chislennyj jempiricheskij analiz tochnosti linejnogo statisticheskogo prognozirovanija temperatury atmosfernogo vozduha // Dokl. BGUIR. 2011. № 8 (62). S. 14–21. (in Russ.)
11. Muha V.S. Minimal'nyj srednij risk i jeffektivnost' optimal'nogo polinomial'nogo mnogomerno-matrichnogo prediktora // Kibernetika i sistemnyj analiz. 2011. № 2. S. 121–130. (in Russ.)
12. Muha V.S., Trofimovich A.F. Ocenivanie matematicheskogo ozhidaniya i kovariacionnoj funkcii stacionarnoj sluchajnoj posledovatel'nosti usredneniem po vremeni i mnozhestvu realizacij // Dokl. BGUIR. 2009. № 1 (39). S. 93–99. (in Russ.)
13. Muha V.S., Kozjachij A.N. Mnogomernaja model' meteorologicheskikh dannyh dlja analiticheskoi obrabotki // Dokl. BGUIR. 2010. № 1 (47). S. 100–105. (in Russ.)
14. Sajt Gismeteo. [Electronic data]. – Access mode: <http://www.gismeteo.ru/>. – Date of access: 26.05.2017. (in Russ)

Сведения об авторах

Муха В.С., д.т.н., профессор, профессор кафедры информационных технологий автоматизированных систем Белорусского государственного университета информатики и радиоэлектроники.

Information about the authors

V.S. Mukha, D. Sci., professor, professor of department of automated data processing systems of Belarusian state university of informatics and radioelectronics.

Адрес для корреспонденции

220013, Республика Беларусь,
г. Минск, ул. П. Бровки, д. 6,
Белорусский государственный университет
информатики и радиоэлектроники
тел. +375-17-293-88-23;
e-mail: mukha@bsuir.by
Муха Владимир Степанович

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovka Str., 6,
Belarusian state university
of informatics and radioelectronics
tel. +375-17-293-88-23;
e-mail: mukha@bsuir.by
Mukha Vladimir Stepanovich