

## МОДЕЛЬ АНАЛИЗА ДИНАМИКИ ВЕКТОРНОГО МЕТЕОРОЛОГИЧЕСКОГО ПРОЦЕССА

М. Г. Матвеев\*, В. В. Михайлов\*\*, М. Е. Семенов\*\*, Е. А. Сирота\*

\* Воронежский государственный университет

\*\* ВУНЦ ВВС «Военно-воздушная академия  
имени профессора Н. Е. Жуковского и Ю. А. Гагарина»

Поступила в редакцию 14.02.2013 г.

**Аннотация.** Предлагается модель для описания динамики многомерного, нестационарного временного ряда атмосферных температур с изменяющимися состояниями атмосферы. Приводятся результаты сравнения с известными подходами к моделированию нестационарных рядов.

**Ключевые слова:** температура, многомерный временной ряд, векторная авторегрессия, искусственная нейронная сеть.

**Annotation.** A model for description of dynamics of the multidimensional, non-stationary time series of atmospheric temperature with the changing state of the atmosphere. Presents the results of a comparison with the known approaches to the modelling of non-stationary series.

**Keywords:** temperature, multivariate time series, vector autoregressive, an artificial neural network.

### ВВЕДЕНИЕ

Важной задачей метеорологии является анализ динамических процессов изменения атмосферной температуры. При этом рассматриваются скалярные поля температуры в трехмерном пространстве земной атмосферы. Дискретное представление этого пространства – трехмерная сетка с постоянными шагами по меридианам, параллелям и высоте от поверхности земли. Будем считать, что в узлах сетки производится ежедневное измерение температуры  $x$  и в каждом узле посредством случайного процесса с дискретным временем формируется временной ряд  $x(t)$ . Динамика температуры в узле с конкретными пространственными координатами зачастую определяется не только местными факторами погоды, но и влиянием конвективной составляющей, которая может оцениваться изменением аналогичных переменных в смежных узлах сетки. Действительно, логично предположить существование корреляционной зависимости процессов изменения температуры в смежных узлах. Такое предположение находит экспериментальное подтверж-

дение [1]. Будем рассматривать задачу построения математического описания совместной динамики дискретных случайных процессов в узлах сетки.

Пусть вид модели определяется как векторная авторегрессия [2]:

$$x(t) = A_1 x(t-1) + \dots + A_p x(t-p) + \varepsilon(t), \quad (1)$$

где  $x = (x_1; \dots; x_n)^T$  – вектор значений среднесуточной температуры в  $n$  смежных узлах сетки;  $A_k$  – матрица параметров модели,  $k = 1, \dots, p$  – временной лаг авторегрессии;  $\varepsilon(t)$  – вектор серийно некоррелированных ошибок, о которых предполагается, что они имеют нулевое среднее и интерпретируются как белый шум.

Если векторный процесс  $x(t)$  – стационарный и эргодический, то параметры  $A_k = \|a_{mk}\|$  векторной авторегрессии (1) находятся как МНК-оценки по однородным (полученным при одинаковых условиях) статистическим данным. В этом случае параметры представляют собой константы, не изменяющиеся во времени и отвечающие условиям стационарности.

В большинстве практических приложений, в метеорологии в частности, однородность данных обеспечивается лишь на сравнительно непродолжительных временных отрезках формирования временного ряда, более того границы этих отрезков, как правило, невозможно четко определить, что исключает возможность приме-

© Матвеев М. Г., Михайлов В. В., Семенов М. Е., Сирота Е. А., 2013

Работа выполнена при поддержке грантов РФФИ №№ 11-08-00032, 12-07-00252 и №13-08-00532

нения фиктивных переменных. Формально такое положение можно рассматривать как непрерывную зависимость параметров модели (1) от времени –  $A_k(t)$  и, следовательно, необходимость их корректировки по мере изменения условий. Изменение параметров  $A_k$  обуславливает нестационарность векторного процесса  $x(t)$ .

Наиболее известным подходом к правилам корректировки параметров являются адаптивные методы, например, экспоненциальное сглаживание статистических данных или корректировка параметров модели по рассогласованию измеренных и предсказанных значений членов ряда, например с помощью рекуррентного метода наименьших квадратов. Подробный обзор адаптивных методов приводится в [3]. Применение адаптации ограничено требованием медленного изменения условий функционирования, что часто не выполняется в метеорологии. Другим подходом к учету изменения параметров модели (1) является нечеткая классификация статистических данных ряда и применение модели Сугено [4]. Этот подход применительно к метеорологическим временным рядам, как показано в [4], может давать некоторое улучшение результатов, но не меняет их кардинально.

Следует отметить успешное применение (в отдельных случаях) искусственных нейронных сетей (ИНС) для моделирования поведения временного ряда, например, [4,5]. Можно предположить, что ИНС, как универсальный аппроксиматор, учитывает и нелинейные эффекты, возникающие в связи с изменением условий функционирования.

В данной работе предлагается рассматривать векторную авторегрессию (1) как модель типа «вход-выход» с измеряемыми входными переменными  $x(t-1), \dots, x(t-p)$  и выходными переменными  $x(t)$ , а ее параметры  $A_k(t) = \|a_{mk}(t)\|$  как совокупность внутренних характеристик процесса, определяющих его состояние (состояние атмосферы) в данный момент времени. Будем считать, что статистические данные однородны, если они получены при стабильности состояния атмосферы. Пространство состояний случайного процесса может быть задано системой уравнений состояния  $a_{mk} = f_{mk}(t)$ , где  $m, l = 1, \dots, n$ ;  $k = 1, \dots, p$ .

Построению системы векторной авторегрессии и уравнений состояния временного ряда для

анализа динамики нестационарного векторного метеорологического процесса посвящена данная работа.

### 1. Особенности моделирования динамики температуры в узлах плоской сетки

Подходы к моделированию динамики векторного случайного процесса рассмотрим на примере скалярного поля температур, представленного длинными нестационарными временными рядами в узлах плоской сетки, расположенной на некоторой геопотенциальной высоте от поверхности земли (рис. 1). Выбор плоской сетки обусловлен наличием достаточного количества статистических данных по узлам, образованным меридианами и параллелями, в отличие от геопотенциальных высот. Обозначим среднесуточное значение температуры, измеренное в узле с координатами  $i, j$  как  $x_{ij}$ . Взаимная корреляция случайных последовательностей в смежных узлах сетки (см. рис. 1) может быть достаточно высокой; для рядов отстоящих на два и более шага сетки оценки взаимной корреляции могут оказаться статистически незначимыми. В такой ситуации целесообразно рассматривать динамику температуры в узле  $ij$  только в зависимости от смежных узлов, обозначенных на рисунке 1. Аналогично строится схема зависимости для любого другого узла. Векторная авторегрессия в виде векторного марковского процесса для узлов, показанных на рисунке 1, примет вид системы уравнений

$$\begin{cases} x_{i-1,j}(t+1) = a_1 x_1(t) + \varepsilon_1(t+1); \\ x_{i,j}(t+1) = a_2 x_2(t) + \varepsilon_2(t+1); \\ x_{i+1,j}(t+1) = a_3 x_3(t) + \varepsilon_3(t+1); \\ x_{i,j-1}(t+1) = a_4 x_4(t) + \varepsilon_4(t+1); \\ x_{i,j+1}(t+1) = a_5 x_5(t) + \varepsilon_5(t+1). \end{cases} \quad (2)$$

Здесь  $x_l(t)$ ,  $l = 1, \dots, 5$  – вектор, составленный по схеме рисунка 1 для рассматриваемого узла сетки и включающий значения в смежных узлах;  $a_l$  – вектор параметров соответствующего уравнения.

Рассмотрим, например, авторегрессионное уравнение в узле с координатами  $ij$ , учитывая изменения состояний атмосферы:

$$\begin{aligned} x_{i,j}(t+1) = & a_{i,j}(t)x_{i,j}(t) + \\ & + a_{i-1,j}(t)x_{i-1,j}(t) + \\ & + a_{i+1,j}(t)x_{i+1,j}(t) + a_{i,j-1}(t)x_{i,j-1}(t) + \\ & + a_{i,j+1}(t)x_{i,j+1}(t) + \varepsilon(t+1). \end{aligned} \quad (3)$$

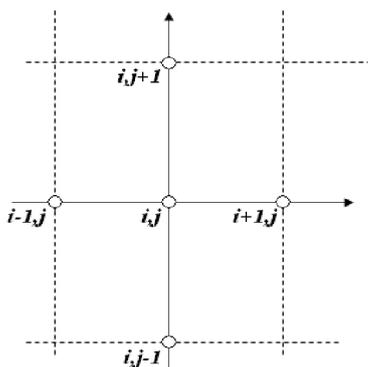


Рис. 1. Фрагмент схемы измерения метеорологических показателей, шаг сетки по параллелям и меридианам составляет  $2,5^\circ$

Для построения уравнений состояния  $a_{ml} = f_{ml}(t)$  необходимо определить признаки, по которым можно было бы оценивать однородность статистических данных. Допустим, что изменение среднесуточной температуры во многом определяется процессом адвекции [7], тогда в качестве такого признака можно принять конечно-разностную оценку градиента температуры в каждом узле пространственной сетки –  $\nabla = (\delta_i; \delta_j)$ , где компоненты вектора градиента оцениваются центральными разностями  $\delta_i = x_{i+1,j} - x_{i-1,j}$ ,  $\delta_j = x_{i,j+1} - x_{i,j-1}$ . Если градиент, вычисленный в окрестности узла с координатами  $i, j$  не изменяется (или почти не изменяется) во времени, то с определенными допущениями можно считать, что состояние атмосферы в данной окрестности стабильно, параметры  $a_{ml}$  также не меняются во времени.

Классы однородной статистики будем строить по двум признакам: направлению и длине вектора-градиента.

Теперь необходимо научиться находить классы однородной статистики и внутри этих классов получать МНК-оценки параметров модели (3). Для построения классов можно предложить следующую процедуру.

Пусть длина и направление вектора  $\nabla$  ограничены соотношениями:

$$|\nabla| \in [0, |\nabla_{\max}|], \quad (4)$$

$$\alpha \in [0^\circ, 360^\circ), \quad (5)$$

где  $\alpha$  – угол между вектором  $\nabla$  и вектором  $(0, 1)$ .

Разобьем промежутки в правой части соотношений (4) и (5) на несколько меньших, одинаковых промежутков  $\Delta x_i$ ,  $i = 1, \dots, k_1$ ; и  $\Delta \alpha_j$ ,

$j = 1, \dots, k_2$ ; так и х, что  $\bigcup_i \Delta x_i = [0, |\nabla_{\max}|]$ ;  $\bigcap_i \Delta x_i = \emptyset$ ;  $\bigcup_j \Delta \alpha_j = [0^\circ, 360^\circ)$ ;  $\bigcap_j \Delta \alpha_j = \emptyset$ . Получим два множества классов длин и направлений вектора градиента:  $\{\Delta x_i\}$ ,  $\{\Delta \alpha_j\}$ . Прямое произведение этих множеств даст множество классов  $(\Delta x_i, \Delta \alpha_j)$ . Каждая пара значений длины вектора  $|\nabla|$  и угла  $\alpha$  внутри класса  $(\Delta x_i, \Delta \alpha_j)$  определяется измеренными значениями температуры в смежных узлах сетки

$$|\nabla| = \sqrt{(x_{i+1,j} - x_{i-1,j})^2 + (x_{i,j+1} - x_{i,j-1})^2}; \quad (6)$$

$$\cos \alpha = \frac{\delta_j}{|\nabla|}. \quad (7)$$

Соотношения (6) и (7) означают, что попадание в тот или иной класс полностью определяется измеренными значениями температур  $x = (x_{i-1,j}; x_{i+1,j}; x_{i,j-1}; x_{i,j+1})$  и определяет возможность каждому классу  $(\Delta x_i, \Delta \alpha_j)$  ставить в однозначное соответствие некоторое подмножество однородных статистических данных и вектор оценок  $\hat{a}_m = (\hat{a}_{i,j}; \hat{a}_{i-1,j}; \hat{a}_{i+1,j}; \hat{a}_{i,j-1}; \hat{a}_{i,j+1})$ .

Продолжая эти рассуждения, при неограниченном увеличении количества классов ( $k_1 \rightarrow \infty, k_2 \rightarrow \infty$ ) приходим к возможности следующей замены аргументов в уравнении состояния

$$a_{ml}(t) = a_{ml}(x(t)) = a_{ml}(x). \quad (8)$$

Это означает, что состояние процесса полностью определяется входным вектором измеренных температур в момент времени  $t$ . Этот вывод представляет собой главную особенность моделирования динамики температуры в узлах пространственной сетки.

Следует отметить, что по статистическим данным внутри класса однородности можно получать МНК-оценки параметров –  $\hat{a}_m = (\hat{a}_{i,j}; \hat{a}_{i-1,j}; \hat{a}_{i+1,j}; \hat{a}_{i,j-1}; \hat{a}_{i,j+1})$ .

## 2. МЕТОДИКА ПОСТРОЕНИЯ МОДЕЛЕЙ НЕСТАЦИОНАРНОЙ ДИНАМИКИ ВЕКТОРНОГО МЕТЕОРОЛОГИЧЕСКОГО ПРОЦЕССА

Построение моделей будем начинать с определения классов однородной статистики в каждом узле сетки, используя в качестве признака классификации длину и направление вектора градиента температур. Очевидно, что чем больше классов удастся выделить, тем больше осно-

ваний для допущения об однородности статистики. Однако число классов будет ограничиваться требованием репрезентативности выборки и получения статистически значимых оценок параметров моделей авторегрессии.

Методику построения моделей будем описывать для узла сетки с координатами  $ij$ . Для смежных узлов сетки методика аналогична.

Для каждого класса однородной статистики найдем МНК-оценки одного уравнения векторной авторегрессии (2). То есть, найдем оценки вектора  $a_m = (a_{i,j}; a_{i-1,j}; a_{i+1,j}; a_{i,j-1}; a_{i,j+1})$ , для  $k$ -го класса однородной статистики. В результате получим совокупность уравнений вида

$$x_{i,j}(t+1) = \hat{a}_{i,j}^k x_{i,j}(t) + \hat{a}_{i-1,j}^k x_{i-1,j}(t) + \hat{a}_{i+1,j}^k x_{i+1,j}(t) + \hat{a}_{i,j-1}^k x_{i,j-1}(t) + \hat{a}_{i,j+1}^k x_{i,j+1}(t), \quad (9)$$

$$k = 1, \dots, K;$$

где  $K$  – число классов;  $x \in D^k$ ;  $D^k$  –  $k$ -й класс однородной статистики.

Вычислим для каждого класса выборочное среднее  $\bar{x}^k = (\bar{x}_{i-1,j}^k; \bar{x}_{i+1,j}^k; \bar{x}_{i,j-1}^k; \bar{x}_{i,j+1}^k)$  и поставим ему в соответствие вектор оценок параметров  $\hat{a}_i^k = (\hat{a}_{i,j}^k; \hat{a}_{i-1,j}^k; \hat{a}_{i+1,j}^k; \hat{a}_{i,j-1}^k; \hat{a}_{i,j+1}^k)$ . Таким образом, получили таблично заданную оценку зависимости вектора параметров  $\hat{a}_i^k$  от вектора входов  $\bar{x}^k$ .

Для аппроксимации этой зависимости целесообразно воспользоваться универсальным аппроксиматором в виде искусственной нейронной сети (ИНС) – многослойного персептрона.

Таким образом, искомая система уравнений состояний атмосферы будет задана в операторном виде

$$\hat{a}_m(x) = F_m(x_{i,j}; x_{i-1,j}; x_{i+1,j}; x_{i,j-1}; x_{i,j+1}), \quad (10)$$

где  $F_m$  – ИНС, играющая роль нелинейного оператора.

Полная модель векторной авторегрессии с изменяющимися состояниями для анализа динамики процессов в  $n$  смежных узлах сетки будет иметь вид

$$x_i(t+1) = \hat{a}_m x_i(t); \quad (11)$$

$$\hat{a}_m(x_l) = F_m(x_l); \quad (12)$$

$$i, m, l = 1, \dots, n.$$

### 3. ЭКСПЕРИМЕНТАЛЬНАЯ АПРОБАЦИЯ МОДЕЛИ ДИНАМИКИ ВЕКТОРНОГО ПРОЦЕССА

Для экспериментальной апробации использовались статистические данные реанализа параметров атмосферы [6] за 2008 год, приведенные к среднесуточным значениям. Рассматривались процессы изменения температуры при геопотенциале 300 ГПа (что позволяет не учитывать приземные факторы) в узле сетки с координатами  $-70^\circ$  северной широты;  $2,5^\circ$  восточной долготы.

Графическое отображение временного ряда в рассматриваемом узле показано на рисунке 2. Очевидно, что показанный процесс нестационарен, содержит возрастающий тренд, гармо-

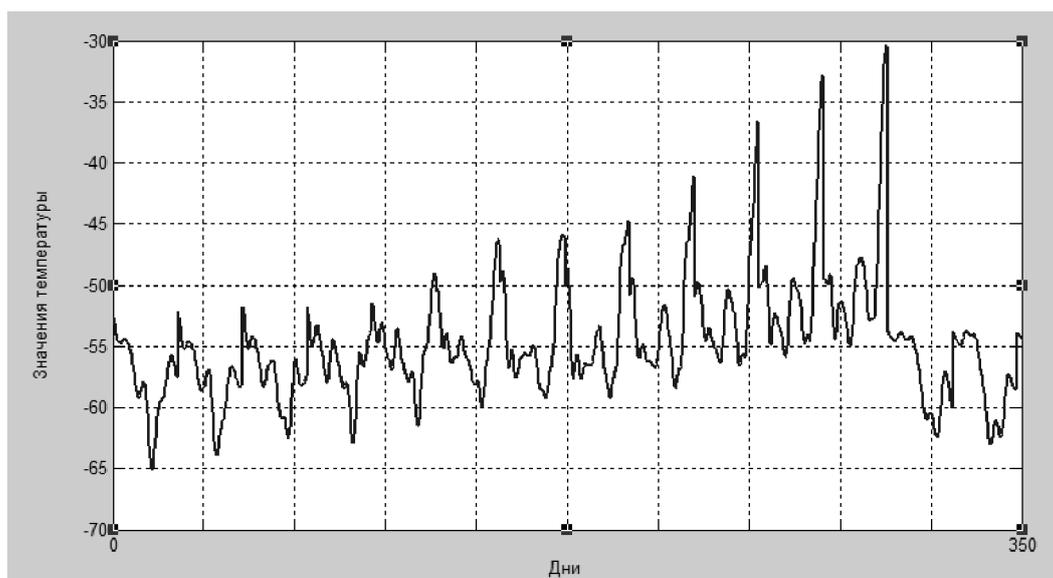


Рис. 2. Процесс изменения температуры в узле сетки с координатами  $-70^\circ$  северной широты;  $2,5^\circ$  восточной долготы

нические составляющие и резкое изменение поведения в окрестностях трехсотого дня.

Согласно предложенной методике моделирования в рассматриваемом узле было выделено 16 классов однородной статистики. В результате в узле оказалось семь классов с непредставительным объемом выборки и далее эти классы не рассматривались.

С использованием МНК было построено 9 авторегрессионных уравнений, номера которых соответствуют номерам классов, получены оценки соответствующих векторов параметров  $\hat{a}^k$ , которым в соответствие поставлены выборочные средние  $\bar{x}^k$ . Результаты представлены в таблице 1.

В качестве оператора  $F$  был принят трехслойный перцептрон с четырьмя нейронами входного слоя, четырьмя нейронами скрытого и пятью нейронами выходного слоя. Для обучения сети использовалась инструментальная среда MATLAB 10. В качестве функций активации входного и скрытого слоя использовалась сигмоидальная функция «logsig», для выходного слоя линейная функция «purelin». Обучение ИНС производилось с помощью алгоритма Levenberg-Marquardt («trainlm»).

Для оценки эффективности полученной модели динамики векторного процесса с переменными состояниями производилось сравнение результатов этой модели с моделями в виде многослойного перцептрона, моделью Сугено с векторной авторегрессией, моделью Сугено с многослойным перцептроном. Модель Сугено представляет собой систему нечетких продукционных правил, где каждое правило составлено для своего класса однородной статистики. Заключение каждого продукционного правила

содержит либо модель векторной авторегрессии, либо многослойный перцептрон. Моделирование во всех случаях осуществлялось на одной и той же статистике, что обеспечило возможность сравнения результатов. Для последних трех моделей использовались методики моделирования, описанные в [1, 5]. В качестве показателя эффективности каждой модели использовался коэффициент детерминации

$$R^2 = 1 - \frac{\sum_{t=1}^T (x(t) - \hat{x}(t))^2}{\sum_{t=1}^T (x(t) - \bar{x})^2}, \quad (13)$$

$x(t)$ ,  $\hat{x}(t)$  – фактические и расчетные значения уровней ряда соответственно,  $\bar{x}$  – среднее значение уровня ряда.

Также для сравнения использовался показатель процентной погрешности прогноза MAPE

$$MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{x(t) - \hat{x}(t)}{x(t)} \right| \cdot 100\%, \quad (14)$$

Считается, что MAPE не должен превышать 10–12 %.

Результаты моделирования показаны в таблице 2.

Численные эксперименты, проведенные в других узлах сетки, показали аналогичные результаты.

#### 4. ВЫВОДЫ

Результаты сравнения моделей, представленные в таблице 2, показывают работоспособность и очевидное преимущество модели динамики векторного процесса с переменными со-

Таблица 1.

Соответствие векторов средних и векторов оценок параметров

№ класса	Вектор средних, $\bar{x}^k$	Вектор оценок параметров, $\hat{a}^k$
1	(-56,39; -52,68; -53,85; -54,60)	(1,35; -0,03; -0,52; -0,003; 0,21)
2	(-56,51; -53,85; -54,80; 52,00)	(-6,15; 0,015; 6,80; -0,00004; 0,34)
3	(-58,05; -58,94; -58,00; -56,98)	(-5,30; 0,005; 6,28; 0,000015; 0,007)
4	(-55,22; -54,42; -54,27; -56,72)	(1,17; 0,0037; -0,23; 0,0028; 0,049)
6	(-55,30; -54,33; -56,66; -54,75)	(0,78; -0,076; -0,166; 0,057; 0,40)
7	(-56,37; -55,46; -55,96; -55,10)	(1,31; -0,00005; -0,34; 0,0016; 0,014)
11	(-57,58; -56,33; -54,22; -54,98)	(1,36; -0,0005; -0,40; 0,00007; 0,029)
15	(-54,66; -49,92; -43,32; -42,49)	(-4,25; -0,0016; 5,22; 0,000025; 0,017)
16	(-56,04; -56,94; -53,72; -58,56)	(1,61; -0,00013; -0,70; 0,00001; 0,09)

Результаты сравнения моделей описания динамики нестационарного векторного процесса

Вид модели	$R^2$	MAPE
Модель динамики векторного процесса с переменными состояниями	0,86	3,50 %
Трехслойный перцептрон	0,61	7,86%
Модель Сугено с векторной авторегрессией	0,65	6,91%
Модель Сугено с трехслойным перцептроном	0,65	6,83%

стояниями перед другими подходами к моделированию нестационарной динамики.

Однако главным результатом проведенного исследования можно считать экспериментальное обоснование возможности использования моделей векторной авторегрессии с переменными параметрами для анализа динамики нестационарных процессов в метеорологии.

СПИСОК ЛИТЕРАТУРЫ

1. Использование модели Сугено для прогнозирования метеорологических показателей. / М. Г. Матвеев, В. В. Михайлов, М. Е. Семенов // Вестник ВГУ, серия «Системный анализ и информационные технологии», научный журнал, № 2, 2011. – С. 164–169.

2. Айвазян С.А. Основы эконометрики. – М.: ЮНИТИ, 2001. – 495 с.

**Матвеев Михаил Григорьевич** – д. т. н., профессор кафедры программирования и информационных технологий, Воронежский государственный университет. E-mail: mgmatveev@yandex.ru

**Михайлов Владимир Владимирович** – д. т. н., начальник факультета метеорологии, ВУНЦ ВВС «Военно-воздушная академия имени профессора Н. Е. Жуковского и Ю. А. Гагарина». Тел. (473) 226-65-06

**Семенов Михаил Евгеньевич** – д.ф.-м.н., профессор кафедры теоретической гидрометеорологии, ВУНЦ ВВС «Военно-воздушная академия имени профессора Н. Е. Жуковского и Ю. А. Гагарина». Тел. (473) 226-13-66. E-mail: mk150@mail.ru

**Сирота Екатерина Александровна** – к.ф.-м.н, доцент кафедры цифровых технологий Воронежского государственного университета

3. Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов. – М.: Финансы и статистика, 2003. – 416 с.

4. Матвеев М.Г., Михайлов В.В. Управление организационно-технической системой в условиях метеорологической неопределенности – Воронеж: ВВАИИ, 2006. – 128 с.

5. Методические аспекты применения аппарата нечеткой логики при обработке и анализе метеорологической информации, необходимой для обеспечения боевой деятельности Вооруженных сил РФ / Матвеев М. Г., Михайлов В. В., Семенов М. Е., Первезенцев Р. Е. // Вестник Военного авиационного инженерного университета, № 3 (14), 2011. – С. 35–42.

6. NCEP/DOE AMIP II Reanalysis [Электронный ресурс]. URL: <http://www.cdc.noaa.gov/cdc/data/ncep.reanalysis2.html>.

7. Матвеев Л.Т. Курс общей физики атмосферы. 2-е изд. Л.: Гидрометеиздат, 1984. – 687 с.

**Matveev Mikhail G.** – doctor of technical sciences, professor department of programming and information technologies, Voronezh State University. E-mail: mgmatveev@yandex.ru

**Mikhailov Vladimir V.** – doctor of technical sciences, the chief of faculty meteorology, Military educational and scientific center of the AIR FORCE's «Air Force Academy im. professor N. E. Zhukovsky and Yuri Gagarin». Tel. (473) 226-65-06

**Semenov Mikhail E.** – doctor of physics-math. sciences, professor department of theoretical hydrometeorology, Military educational and scientific center of the AIR FORCE's «Air Force Academy im. professor N. E. Zhukovsky and Yuri Gagarin». Tel. (473) 226-13-66. E-mail: mk150@mail.ru

**Sirota Ekaterina** – PhD of sciences, associate professor of digital technology, Voronezh State University