

МОДЕЛИРОВАНИЕ КОМПЛЕКСА МЕТЕОРОЛОГИЧЕСКИХ ПАРАМЕТРОВ С УЧЕТОМ ИХ ГОДОВОЙ НЕСТАЦИОНАРНОСТИ

Н. А. Каргаполова^{1,2}, В. А. Огородников^{1,2}

¹ *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

² *Новосибирский государственный университет, 630090, Новосибирск*

УДК 519.6

В работе предложена стохастическая модель комплекса из трех метеорологических элементов: индикатора наличия/отсутствия осадков в течение суток, минимальной и максимальной суточной температуры. Модель построена в предположении годовой нестационарности рассматриваемых метеопроцессов. Для моделирования индикаторного ряда используется латентный гауссовский процесс, а моделирование рядов температуры осуществляется в предположении гауссовости их одномерных распределений. При этом параметры распределений являются функциями времени. В работе также приведены аналитические преобразования уравнений, связывающих корреляционные функции латентного гауссовского и индикаторного процессов, позволяющие существенно сократить время вычислений значений корреляционной функции латентного процесса при заданных значениях корреляционной функции индикаторного.

Ключевые слова: стохастическое моделирование, нестационарный случайный процесс, приземная температура воздуха, индикатор осадков.

Введение

При решении многих прикладных задач в таких областях науки, как океанология, гидрология, агрометеорология, популяционная биология, достаточно часто требуется знание статистических свойств различных метеорологических процессов. Так, например, бывает необходимо оценивать вероятности возникновения сочетаний метеорологических элементов, благоприятных для распространения лесных пожаров, вероятности появления заморозков в весенне-летний период, среднее число засушливых дней и др. Во многих случаях в силу статистической ненадёжности оценок по малым выборкам исследование таких явлений по реальным данным является затруднительным, и поэтому приходится использовать различные стохастические модели.

В связи с этим в последние десятилетия бурное развитие переживает прикладная область математики, связанная с разработкой так называемых “генераторов погоды”. По своей сути “генераторы” являются пакетами программ, позволяющими численно моделировать длинные ряды случайных чисел, обладающих статистическими свойствами, повторяющими основные свойства реальных метеорологических рядов или их комплексов. Чаще всего моделируются ряды приземной температуры воздуха, суточного минимума и максимума температуры, количество осадков и количество солнечной радиации [2, 3, 9, 10, 11]. С помощью “генераторов погоды” моделируют не только одномерные случайные ряды, но случайные поля, с опорными точками, соответствующими местоположениям метеостанций [4, 5]. Каждый “генератор погоды” имеет временной интервал (несколько суток, месяц, год), на котором он “работает” наилучшим образом, и вне которого он дает неправдоподобные результаты. Так, например, известные “генераторы” WGEN [9, 10] и LARS [11] дают хорошие результаты при моделировании на относительно коротких временных интервалах (порядка 1 месяца), но неприменимы для моделирования метеорологических процессов на годовом интервале, поскольку они созданы в предположении стационарности метеопроцессов, а на годовом интервале метеорологические процессы сильно нестационарны.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты 15-01-01458-а, 16-01-00145-а, 16-31-00038-мол-а, 16-31-00123-мол-а) и гранта Президента РФ (МК-659.2017.1).

В данной работе предложена стохастическая модель комплекса из 3 метеорологических элементов (минимальной и максимальной суточной приземной температуры воздуха и индикатора осадков), пригодная для моделирования временных рядов комплексов этих метеозадаментов на интервале годовой длины. Модель основана на специальном нелинейном преобразовании гауссовских временных рядов, при котором за счет выбора параметров модели учитывается годовая нестационарность реального процесса, связанная с вращением Земли вокруг Солнца.

1 Алгоритм моделирования

В этом параграфе будет дано формальное описание модели, а также описаны предположения, в которых модель была построена.

Моделирование будет опираться на гипотезу о том, что одномерные распределения минимальной и максимальной температуры в каждые сутки являются нормальными, при этом параметры этих распределения изменяются день ото дня. Эта гипотеза хорошо согласуется с реальными данными, собранными на различных метеостанциях России (проверка гипотезы о нормальном распределении проводилась на основе реальных данных наблюдений с помощью критерия χ^2).

Опишем формально правила и алгоритмы, по которым проводилось моделирование. Строится процесс $M = (\vec{I}^T, \vec{A}^T, \vec{E}^T)$, где $\vec{I}^T = (I_1, I_2, \dots, I_N)^T$ — вектор минимальной суточной температуры в течение года, $\vec{A}^T = (A_1, A_2, \dots, A_N)^T$ — вектор максимальной суточной температуры, $\vec{E}^T = (E_1, E_2, \dots, E_N)^T$ — вектор индикаторов наличия/отсутствия осадков. Если $E_j = 1$, то считаем что в сутки номер j осадки наблюдались, если $E_j = 0$, то в сутки номер j осадки отсутствовали. Считаем число N дней в году неизменным и равным 365. Основная моделирующая формула имеет вид

$$\begin{aligned} I_j &= \sigma_j^I \xi_j^I + \mu_j^I, \\ A_j &= \sigma_j^A \xi_j^A + \mu_j^A, \\ E_j &= \begin{cases} 1, & \xi_j^E \leq c_j, \\ 0, & \xi_j^E > c_j, \end{cases} \end{aligned} \quad (1)$$

где векторы

$$\begin{pmatrix} \mu_1^I \\ \mu_2^I \\ \vdots \\ \mu_N^I \end{pmatrix}, \begin{pmatrix} \mu_1^A \\ \mu_2^A \\ \vdots \\ \mu_N^A \end{pmatrix}, \begin{pmatrix} \sigma_1^I \\ \sigma_2^I \\ \vdots \\ \sigma_N^I \end{pmatrix}, \begin{pmatrix} \sigma_1^A \\ \sigma_2^A \\ \vdots \\ \sigma_N^A \end{pmatrix}$$

есть векторы средних значений и среднеквадратических отклонений минимальной и максимальной температуры, значения компонент которых оценены по реальным данным. Пороговые значения c_j для каждого $j = 1, 2, \dots, N$ определяются из уравнения

$$P(E_j = 1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c_j} \exp\left(-\frac{t^2}{2}\right) dt = p_j,$$

где вероятности p_j наличия осадков в j -е сутки оценены по реальным данным. Отметим, что равенство $c_i = 0$ выполняется тогда и только тогда, когда $p_i = \frac{1}{2}$, неравенство $c_i > 0$ выполняется только при $p_i > \frac{1}{2}$, а неравенство $c_i < 0$ — только при $p_i < \frac{1}{2}$. Все дальнейшие рассуждения приведены для случая $p_j \neq 0$ и $p_j \neq 1$ при $j = 1, 2, \dots, N$. Величины $(\xi_j^I, \xi_j^A, \xi_j^E)$ есть компоненты совместного гауссовского ряда $\Omega = \left((\xi^I)^T, (\xi^A)^T, (\xi^E)^T \right)$ с нулевым средним и такой специально выбранной корреляционной матрицей

$$G = \begin{pmatrix} G_{II} & G_{IA} & G_{IE} \\ G_{AI} & G_{AA} & G_{AE} \\ G_{EI} & G_{EA} & G_{EE} \end{pmatrix},$$

что процесс M обладает корреляционной матрицей

$$R = \begin{pmatrix} R_{II} & R_{IA} & R_{IE} \\ R_{AI} & R_{AA} & R_{AE} \\ R_{EI} & R_{EA} & R_{EE} \end{pmatrix},$$

определяемой на основе реальных данных. Индекс XY у блоков G_{XY} , R_{XY} означают, что эти блоки составлены из взаимных коэффициентов корреляции между значениями метеозаписей X и Y во все возможные моменты времени i и j , где $X, Y \in \{I, A, E\}$, $i, j = 1, 2, \dots, N$. Отметим, что матрицы R и G имеют размерность $3N \times 3N$, т.е. 1095×1095 .

Возникает вопрос о том, как определить матрицу G такую, чтобы построенный согласно преобразованию (1) процесс обладал необходимой корреляционной матрицей R . Отметим, что в работе [6] был рассмотрен аналог преобразования (1) в более простом случае.

Определим последовательно все блоки матрицы G через R и p_j (т.е. выпишем уравнения относительно всех элементов каждого из блоков матрицы G). Очевидно, что

$$\begin{aligned} G_{II} &= R_{II}, \\ G_{AA} &= R_{AA}, \\ G_{IA} &= R_{IA}, \\ G_{AI} &= R_{AI}. \end{aligned}$$

Обозначим элементы матриц R_{EE} и G_{EE} , стоящие на пересечении i -ой строки и j -го столбца через $r_{EE}(i, j)$ и $g_{EE}(i, j)$, $i, j = 1, 2, \dots, N$. Очевидно, что $r_{EE}(i, i) = g_{EE}(i, i) = 1$, $i = 1, 2, \dots, N$. Из определения коэффициента корреляции и определения процесса \vec{E}^T следует выполнение равенств

$$r_{EE}(i, j) = \frac{P(E_i = 1, E_j = 1) - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}}$$

для $i, j = 1, 2, \dots, N$. При этом

$$P(E_i = 1, E_j = 1) = P(\xi_i^E \leq c_i, \xi_j^E \leq c_j) = F(c_i, c_j, g_{EE}(i, j)), \quad (2)$$

где

$$F(h, k, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^h \int_{-\infty}^k \exp\left(-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right) dx dy.$$

Равенства (2) позволяют численно определить значение всех элементов матрицы G_{EE} . Для того чтобы упростить вычисления при естественном предположении $|g_{EE}(i, j)| \neq 1$ вне диагонали матрицы G_{EE} , можно воспользоваться формулами Оуэна [1, 8]

$$F(c_i, c_j, g_{EE}(i, j)) = \frac{1}{2}\Phi(c_i) + \frac{1}{2}\Phi(c_j) - T(c_i, a_1) - T(c_j, a_2) + \frac{1}{2},$$

если $c_i \geq 0$, $c_j \geq 0$ или $c_i < 0$, $c_j < 0$, и

$$F(c_i, c_j, g_{EE}(i, j)) = \frac{1}{2}\Phi(c_i) + \frac{1}{2}\Phi(c_j) - T(c_i, a_1) - T(c_j, a_2),$$

если $c_i < 0$, $c_j \geq 0$ или $c_i \geq 0$, $c_j < 0$, где

$$\begin{aligned} \Phi(c_i) &= \frac{1}{\sqrt{2\pi}} \int_0^{c_i} \exp(-t^2/2) dt, \\ T(c, a) &= \frac{1}{2\pi} \int_0^a \exp\left(-\frac{c^2(1+t^2)}{2}\right) \frac{dt}{1+t^2}, \\ a_1 &= \frac{c_j - c_i g_{EE}(i, j)}{c_i \sqrt{1 - g_{EE}^2(i, j)}}, \quad a_2 = \frac{c_i - c_j g_{EE}(i, j)}{c_j \sqrt{1 - g_{EE}^2(i, j)}}. \end{aligned}$$

С учетом этих формул и равенств

$$p_i = \frac{1}{2} + \Phi(c_i), \quad i = 1, 2, \dots, N$$

уравнения для нахождения $g_{EE}(i, j)$ при $c_i > 0$, $c_j > 0$ или $c_i < 0$, $c_j < 0$ (т.е. при выполнении условия $c_i c_j > 0$) принимают вид

$$r_{EE}(i, j) = \frac{\frac{1}{2}p_i + \frac{1}{2}p_j - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}} - \frac{T(c_i, a_1) + T(c_j, a_2)}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}},$$

а при $c_i < 0, c_j > 0$ или $c_i > 0, c_j < 0$ (т.е. при выполнении условия $c_i c_j < 0$)

$$r_{EE}(i, j) = \frac{\frac{1}{2}p_i + \frac{1}{2}p_j - \frac{1}{2} - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}} - \frac{T(c_i, a_1) + T(c_j, a_2)}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}}.$$

Если же одно или оба пороговых значения равны 0, то используя равенства

$$\begin{aligned} F(0, 0, g_{EE}(i, j)) &= \frac{1}{4} + \frac{\arcsin g_{EE}(i, j)}{2\pi}, \\ T(0, a) &= \frac{\arctg(a)}{2\pi}, \\ T(h, a) &= -T(h, -a) \end{aligned}$$

можно получить следующие соотношения:

1) при $c_i = 0, c_j = 0$

$$r_{EE}(i, j) = \frac{2 \arcsin g_{EE}(i, j)}{\pi}.$$

2) при $c_i = 0, c_j \neq 0$

$$r_{EE}(i, j) = -\frac{2T\left(c_j, -\frac{g_{EE}(i, j)}{\sqrt{1-g_{EE}^2(i, j)}}\right)}{\sqrt{p_j(1-p_j)}},$$

3) при $c_i \neq 0, c_j = 0$

$$r_{EE}(i, j) = -\frac{2T\left(c_i, -\frac{g_{EE}(i, j)}{\sqrt{1-g_{EE}^2(i, j)}}\right)}{\sqrt{p_i(1-p_i)}}.$$

Для того чтобы определить элементы матрицы G_{IE} , необходимо воспользоваться соотношениями вида

$$r_{IE}(i, j) = \frac{EI_i E_j - EI_i E E_j}{\sqrt{DI_i} \sqrt{DE_j}} = -\frac{g_{IE}(i, j)}{\sqrt{2\pi p_j(1-p_j)}} \exp(-c_j^2/2).$$

Матрицы G_{EI}, G_{AE}, G_{EA} определяются аналогичным образом.

После того как матрица G полностью определена, можно приступить к моделированию гауссовского процесса Ω и последующему построению процесса M согласно основной моделирующей формуле (1). Моделирование гауссовского процесса с заданной корреляционной матрицей осуществляется с помощью широко известных алгоритмов, например, алгоритмов на основе спектрального или LL^T -разложения корреляционной матрицы [7]. Заметим, что в следствии вычислительных ошибок при вычислении элементов матрицы G , эта матрица может оказаться не положительно определенной. В таком случае перед моделированием гауссовского процесса необходимо провести процедуру регуляризации матрицы G [7]. Численные эксперименты показали, что при вычислении матрицы G указанным выше способом, в ряде случаев эта матрица оказывается не только отрицательно определенной, но и существенно искажается при применении регуляризации на основе искажения спектра матрицы. Поиск других способов регуляризации в таких случаях является задачей, которую ещё предстоит решить.

По смыслу минимальной и максимальной температуры в каждой моделируемой реализации во все моменты времени $j = 1, 2, \dots, N$ должно выполняться неравенство $I_j \leq A_j$. Однако моделирующая формула (1) выполнение этого неравенства не гарантирует. На практике, поскольку $\mu_j^I < \mu_j^A$ и σ_j^I, σ_j^A относительно невелики, вероятность события $I_j > A_j$ мала, поэтому реализации, в которых требуемое неравенство не выполняется, можно “отбрасывать”, т.е. не включать в выборку, по которой производится проверка адекватности модели и дальнейшие исследования.

2 Численные эксперименты

Описанная в предыдущем параграфе модель была применена для моделирования нестационарных совместных временных рядов метеозаписей на ряде станций, расположенных на территории России. Верификация модели показала, что для этих станций модель дает удовлетворительные результаты. В качестве примера, рассмотрим такую характеристику как среднее число дней в месяце, когда минимальная температура ниже

0°C, а максимальная — выше. Эта характеристика не является входным параметром модели, поэтому может быть использована для верификации. В табл. 1 приведены значения этой характеристики для различных месяцев, полученные оценкой по реальным данным и по ансамблю модельных траекторий. Как видно из табл. 1, модель воспроизводит данную характеристику с точностью до статистической погрешности.

Таблица 1: Среднее число дней с $I_j < 0, A_j > 0$. Санкт-Петербург.

Месяц	Среднее число дней	
	Реальные данные	Модельные данные
Октябрь	4.7	4.9
Ноябрь	8.9	9.1
Декабрь	5.7	5.3
Январь	9.3	9.4
Февраль	7.7	7.4
Март	16.3	16.8

Поскольку верификация модели прошла успешно, в дальнейшем эта модель будет использована для оценки вероятностей возникновения неблагоприятных сочетаний метеоэлементов (например, длительных периодов с высокой минимальной температурой и отсутствием осадков в летнее время, больших суточных колебаний температуры в зимний период и др.)

Список литературы

- [1] Смирнов Н.В., Большев Л.Н. Таблицы для вычислений функции двумерного нормального распределения // Изд. АН СССР, 1962.
- [2] Furrer E.M., Katz R.W. Generalized linear modeling approach to stochastic weather generators // Clim. Res. 2007. V. 34, iss. 2. P. 129–144.
- [3] Kargapolova N.A., Ogorodnikov V.A. Inhomogeneous Markov chains with periodic matrices of transition probabilities and their application to simulation of meteorological processes // Russ. J. Num. Anal. Math. Modelling. 2012. V. 27, iss. 3. P. 213–228.
- [4] Kleiber W., Katz R. W., Rajagopalan B. Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes // Water Resources Research. 2012. V. 48.
- [5] Ogorodnikov V.A., Kargapolova N.A., Seresheva O.V. Numerical stochastic model of spatial fields of daily sums of liquid precipitation // Russ. J. Num. Anal. Math. Modelling. 2013. V. 28, iss. 2. P. 187–200.
- [6] Ogorodnikov V.A., Khlebnikova E.I., Kosyak S.S. Numerical stochastic simulation of joint non-Gaussian meteorological series // Russ. J. Num. Anal. Math. Modelling. 2009. V. 24, iss. 5. P. 467–480.
- [7] Ogorodnikov V.A., Prigarin S.M. Numerical Modelling of Random Processes and Fields: Algorithms and Applications // the Netherlands, Utrecht: VSP, 1996.
- [8] Owen D.B. Table for computing bivariate normal probabilities // Ann. Math. Statist. 1956. V. 27, iss. 4. P. 1075–2000.
- [9] Richardson C. W. Stochastic simulation of daily precipitation, temperature and solar radiation // Water Resour. Res. 1981. V. 17. P. 182–190.
- [10] Richardson C. W., Wright D. A. WGEN: A Model for Generating Daily Weather Variables // U. S. Department of Agriculture, Agricultural Research Service, ARS-8, 1984.
- [11] Semenov M.A., Barrow E.M. LARS-WG: A Stochastic Weather Generator for Use in Climate Impact Studies. User Manual // Harpenden, 2002.

*Нина Александровна Каргаполова — к.ф.-м.н., науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
Новосибирский государственный университет;
e-mail: nkaargapolova@gmail.com;*

*Василий Александрович Огородников — д.ф.-м.н., гл.науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
Новосибирский государственный университет;
e-mail: ova@osmf.sscs.ru.*

Дата поступления — 15 мая 2017 г.