

Марковские модели индикаторных рядов метеорологических процессов[†]

Каргаполова Нина Александровна[‡]

Введение

Рассматривается задача численного стохастического моделирования индикаторных рядов метеорологических процессов. Предлагаются 3 модели, основанные на марковских цепях различных типов, позволяющие оценивать вероятности возникновения длительных периодов времени, когда значение моделируемого метеоэлемента превышает некоторый заданный уровень. Рассматриваемые модели позволяют учитывать суточный ход реальных процессов.

Первая модель основана на неоднородных марковских последовательностях, у которых матрица переходных вероятностей является периодической функцией времени. Теоретические результаты изучения свойств таких цепей представлены в [1 - 4].

Моделирование скалярного случайного процесса во второй модели производится с помощью специального преобразования однородных векторных марковских цепей [1, 5].

В третьей модели используются дискретные скалярные однородные марковские процессы различного порядка. Такие процессы не обладают периодическими свойствами, однако, при моделировании метеопроцессов со слабо выраженной периодичностью, применение этих процессов, за счёт их высокого порядка связности, даёт достаточно хорошие результаты [6].

На основе реальных рядов метеоэлементов, полученных на метеостанциях Астрахани, Перми и Новосибирской области, проведены оценки вероятностей возникновения таких экстремальных явлений как длительные заморозки в летние месяцы, длительные засушливые и жаркие периоды времени и др.

[†] Работа выполнена при финансовой поддержке РФФИ, гранты № 11-01-00641-а, 12-01-00727-а.

[‡] Институт вычислительной математики и математической геофизики СО РАН, Новосибирск.

Марковские модели

Модель 1.

Рассматривается дискретная скалярная марковская последовательность с конечным числом состояний и матрицей переходных вероятностей являющейся периодической функцией времени. В работе [4] приведены теоретические результаты, показывающие, что такая последовательность, в случае двух возможных состояний, обладает периодическими распределениями, осциллирующей корреляционной функцией, периодическими распределениями длин серий постоянных значений. В случае большего числа состояний теоретические исследования свойств процесса оказываются достаточно трудоёмкими и громоздкими, однако численные эксперименты показывают, что свойства периодичности процесса сохраняются.

Модель 2.

Моделируется марковская последовательность m – мерных векторов $\vec{\omega}_1, \vec{\omega}_2, \dots, \vec{\omega}_t, \dots$, где компоненты ω_{it} , $i = 1, 2, \dots, m$ вектора $\vec{\omega}_t$ принимают одно из l возможных значений. Далее рассматривается скалярный процесс

$$\omega_{11}, \omega_{21}, \dots, \omega_{m1}, \omega_{12}, \omega_{22}, \dots, \omega_{m2}, \dots, \omega_{1t}, \omega_{2t}, \dots, \omega_{mt}, \dots$$

В [1, 5] показано, что такой скалярный процесс является периодическим по распределениям с периодом m . Однако, поскольку при моделировании векторного процесса учитываются вероятности перехода из вектора в вектор, а связи между компонентами внутри одного вектора не учитываются, то полученный скалярный процесс по своей корреляционной структуре отличается от процесса в Модели 1.

Модель 3.

Рассматриваются скалярные однородные марковские последовательности различного порядка. Напомним, что марковская цепь $\{X_i\}_{i \geq 0}$ является марковской цепью порядка k , если

$$\begin{aligned} P(X_{i+1} = j_{i+1} | X_i = j_i, X_{i-1} = j_{i-1}, \dots, X_0 = j_0) = \\ = P(X_{i+1} = j_{i+1} | X_i = j_i, X_{i-1} = j_{i-1}, \dots, X_{i-k+1} = j_{i-k+1}). \end{aligned}$$

В [6] приведены примеры использования марковских цепей третьего порядка для моделирования индикаторных последовательностей для

суточных сумм жидких осадков и оценки продолжительности дождливых и сухих периодов.

Результаты численных экспериментов

По ряду реальных данных ξ_t , $t = 0, 1, 2, \dots$ строится индикаторный процесс

$$\eta_t = \begin{cases} 1, & \xi_t > c, \\ 0, & \xi_t \leq c, \end{cases}$$

где c – некоторый заданный уровень. В качестве реальных данных были выбраны данные о температуре воздуха, модуле скорости ветра и других метеоэлементах в Перми и Астрахани за период с 1936 по 1984 гг. и с 1966 по 2000 гг. соответственно. Однако, оценки сложных характеристик процесса η_t , особенно связанных с редкими событиями, непосредственно по реальным данным обладают большой погрешностью в силу малости выборки. Поэтому на основе Моделей 1-3 строились последовательности из 0 и 1, обладающие статистическими свойствами как у реальных рядов. Эти последовательности использовались для оценки возникновения редких событий.

Для всех моделей входными характеристиками являются начальное распределение и матрица переходных вероятностей. Эти вероятности оценивались по реальному процессу. Оценки свойств процесса в Моделях 2, 3 производились по 1 000 000 реализаций, оценки для Модели 1 были проведены на основе формул, представленных в [4]. Оценивание производилось на временных интервалах месячной длины. Основное внимание уделялось оценке вероятностей вида

$$P(L_1 = k) = P(\chi_t = 0, \chi_{t+1} = 1, \dots, \chi_{t+k} = 1, \chi_{t+k+1} = 0) / P(\chi_t = 0, \chi_{t+1} = 1).$$

Такие вероятности соответствуют распределению длительностей серий постоянных значений процесса, что, например, при моделировании рядов температуры воздуха в зимние месяцы при уровне $c = 0^\circ C$ соответствует вероятностям возникновения оттепелей различной продолжительности.

В первом эксперименте в качестве процесса ξ_t , $t = 0, 1, 2, \dots$ выбраны данные четырёхсрочных наблюдений (в 0, 6, 12 и 18 часов) приземной температуры воздуха в Астрахани. В связи с этим, в Модели 1 был выбран

период $n = 4$, а в Модели 2 были взяты размерность m вектора $\vec{\omega}_i$ равная 4. На Рис.1 приведены графики вероятностей $P(L_1 = k)$ возникновения серий длины k постоянных значений для марта месяца при $c = 0^\circ C$ (эта характеристика соответствует вероятности появления оттепелей различной продолжительности) в Астрахани.

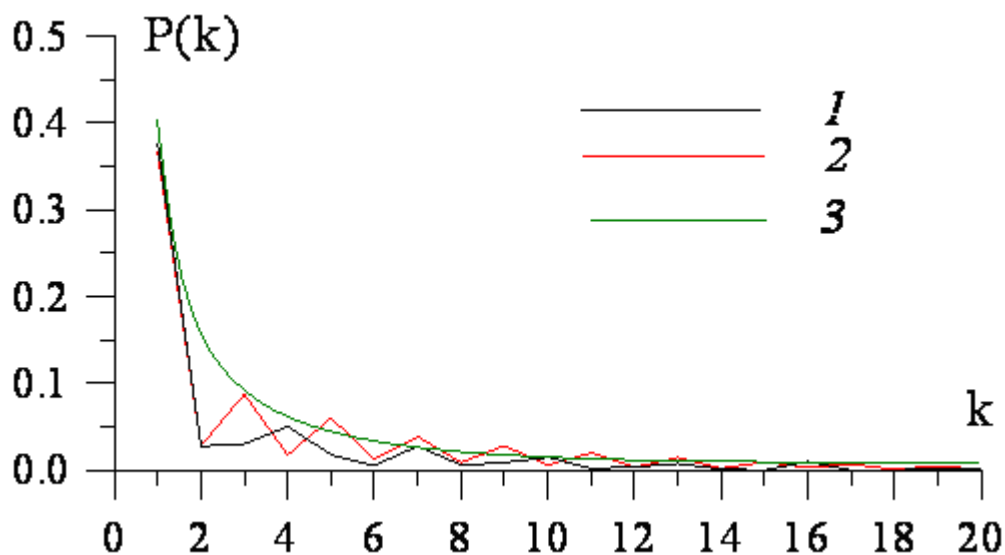


Рис.1 Вероятность $P(L_1 = k)$. 1 – реальные данные, 2 – по Модели 1, 3 – по Модели 3 с марковской цепью 3-го порядка.

Как видно из Рис.1, наилучшим образом эта характеристика при малых k воспроизводится Моделью 1. Поскольку длительные оттепели в Астрахани в марте месяце – явление достаточно редкое, то оценка вероятностей их возникновения по реальным данным менее чем за 40 лет является очень грубой. Т.к. для небольших длительностей оценки вероятности возникновения оттепелей по Модели 1 и по реальным данным достаточно близки, то оценки вероятностей для длительных временных промежутков по Модели 1 можно считать хорошо воспроизводящими свойства реального процесса.

В качестве исходных данных для следующего эксперимента были взяты ряды двенадцатичасовых сумм осадков в июле в Перми. Уровень c для построения индикаторного ряда был взят равным 0.1мм, т.е. количество осадков менее 0.1мм («след осадков») приравнивалось к отсутствию осадков. На Рис.2 приведены графики распределений длительностей серий с осадками. Рассматриваемый процесс не обладает выраженным суточным ходом, поэтому Модели 1 и 2 не дают

приемлемого результата. Для этого случая наилучший результат достигается при использовании Модели 3 с марковской цепью второго порядка.

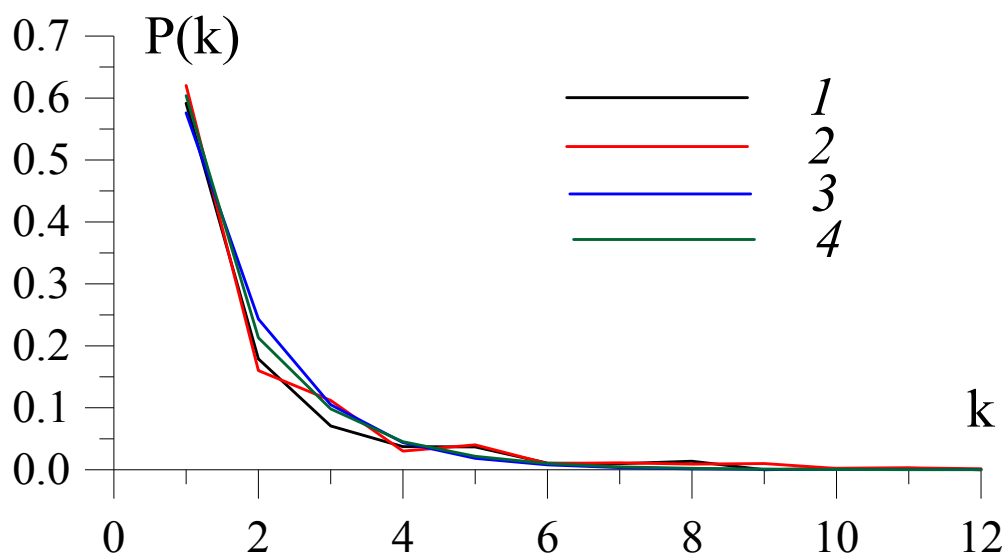


Рис.2. Вероятность $P(L_1 = k)$. 1 – реальные данные, 2 – по Модели 1, 3 – по Модели 3 с марковской цепью 1-го порядка, 4 – по Модели 3 с марковской цепью 2-го порядка. Осадки. Июль. Пермь.

Численные эксперименты по оценке аналогичных распределений для относительной влажности воздуха показывают, что наилучшие результаты даёт Модель 2.

В некоторых случаях можно улучшить результаты моделирования, используя комбинация Моделей 1 и 3. Т.е. можно строить марковские процессы высокого порядка с периодическими свойствами.

Заключение

Все 3 представленные модели могут применяться для моделирования индикаторных рядов метеоэлементов для различных уровней, месяцев и географических местоположений метеостанций.

Модели для оценки вероятностей появления экстремальных событий целесообразно выбирать на основе того, какая из моделей наилучшим образом воспроизводит те характеристики выбросов рассматриваемого реального процесса, которые достаточно точно оцениваются по имеющемуся объему выборки.

Литература

1. Каргаполова Н.А. «Марковские модели нестационарных временных рядов с периодическими свойствами», Труды конференции молодых учёных ИВМиМГ СО РАН, Новосибирск, 2011, с. 21-31.
2. Огородников В.А., Савельев Л.Я., Каргаполова Н.А. «Некоторые свойства неоднородных марковских последовательностей с периодическими матрицами переходных вероятностей», Труды международной конференции «Кубатурные формулы, методы Монте-Карло и их приложения», Красноярск, 2011, с. 86-90.
3. Kargapolova N.A., Saveliev L.Ya., Ogorodnikov V.A. «Modeling of nonstationary processes with periodic properties on basis of Markov chains», Proceedings of International Workshop «Applied Methods of Statistical Analysis. Simulations and Statistical Inference», Novosibirsk, 2011, pp. 323-330.
4. Kargapolova N.A., Ogorodnikov V.A. «Inhomogeneous Markov chains with periodic matrices of transition probabilities and their application to simulation of meteorological processes», Russian Journal of Numerical Analysis and Mathematical Modelling, Volume 27, Issue 3, pp. 213-228.
5. Огородников В.А., Деренок К.В., Толстых У.И. «Специальные численные модели дискретных случайных рядов», Новосибирск: Препринт/РАН. Сиб. отд-ние. ИВМиМГ (1166), 2009.
6. Марченко А.С., Огородников В.А. Вероятностные модели последовательности сухих и дождливых суток, Новосибирск: Препринт/Вычислительный центр СО АН СССР (933), 1991.