Precipitation Maximum Spatial Dependence Structures Modelling

Eugeny Yu. Shchetinin, Veronica M. Martynova

Applied Mathematics Department
Moscow State Technology University "STANKIN"

Email: riviera-moltol@mail.ru

The theory of extremes provides a wide range of fundamental capabilities for its application in hydrology, statistical analysis of extreme precipitation, as well as planning measures to reduce the damage of their consequences.

The aim of this work is to develop mathematical models of the structures of spatial dependence between the maximum precipitation for various weather stations. This natural solution is the application of models for multivariate extremes. However, consider the flow measurements have the properties of random fields, as a rule, are non-stationary. In this regard, it needs to develop spatial and temporal models of statistical dependence structures, including characterizing the context of an extreme type. In this aspect, the spectral representation of de Haan-Pickands is no alternative for multivariate distributions of extreme. The complexity in modeling and assessment of such structures leads to the necessity of development of methods of its representation in the spatial structure of pairwise distributions of extreme. Effective method of solving this problem is to develop adequate copula models in combination with vine structures branching into pairwise copulas.

In addition to modeling structures of branching, which is a separate complicated task, it is necessary to evaluate and analyze the pairwise dependence. To solve this problem, we used the model of copulas of extreme type. Proposed pairwise dependence estimation method of madogram, taking into account the extreme nature of dependency structures. Also proposed a modified non-parametric method of madogram with high stability of statistical estimates.

As a practical application, this paper investigates the flow of data on precipitation in the European part of Russia over the last 20 years. The analysis of rainfall maxima showed that the type of dependence structure of extreme precipitation can be characterized by three main factors: the distance between the two stations, the season (summer or winter) and duration of precipitation (daily, monthly, etc.). Increase the duration of precipitation increases, the spatial dependence of extreme precipitation. Complete independence between them is approximately 50 km (100 km) for summer (winter) with a duration not exceeding one hour, and for a long time after only a few hundred kilometers. In addition, this dependence is always greater in winter than in summer, regardless of the length of precipitation.

Key words and phrases: extreme value, precipitation, spatial dependence, copula, vine structure, madogram.

In: K. E. Samouilov, L. A. Sevastianov, D. S. Kulyabov (eds.): Selected Papers of the VII Conference "Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems", Moscow, Russia, 24-Apr-2017, published at http://ceur-ws.org

О моделировании пространственных структур статистической зависимости экстремальных осадков

Е. Ю. Щетинин, В. М. Мартынова

Кафедра прикладной математики
ФБГОУ ВО МГТУ СТАНКИН

Email: riviera-moltol@mail.ru

Теория экстремальных величин предоставляет широкий спектр фундаментальных возможностей по ее применению в области гидрологии, анализа статистики экстремальных осадков, а также планированию мер по снижению ущерба их последствий.

Целью настоящей работы является разработка математических моделей структур пространственной зависимости между максимумами осадков для различных метеостанций. Для этого естественным решением является применение моделей многомерных экстремальных величин. Однако рассматриваемые потоки измерений обладают свойствами случайных полей, как правило, являющихся нестационарными. В связи с этим потребовалось разработать пространственно-временные модели структур статистической зависимости, в том числе характеризующих связи экстремального типа. В этом аспекте спектральное представление де Хаана–Пикендса является безальтернативным для многомерных распределений экстремумов. Сложность в моделировании и оценивании подобных структур приводит к необходимости разработки методов их представления в виде пространственных структур попарных распределений экстремумов. Эффективным методом решения этой задачи является разработка адекватных моделей копул в сочетании со структурами ветвления на попарные копулы.

Помимо моделирования структур ветвления, что является самостоятельной сложной задачей, необходимо оценивать и анализировать попарные зависимости. Для решения этой задачи в работе использовались модели копул экстремального типа. В работе предложено попарные зависимости оценивать с помощью метода мадограмм, учитывающий экстремальный характер структур зависимости. Предложен также модифицированный непараметрический метод λ-мадограмм, обладающий высокой устойчивостью статистических оценок.

В качестве практического применения в статье исследован поток данных по осадкам в европейской части России за последние 20 лет. Анализ максимумов осадков показал, что тип структуры зависимости экстремальных осадков может быть охарактеризован тремя основными факторами: расстояние между двумя станциями, сезон (лето или зима), а также продолжительностью осадков (ежечасно, ежедневно, ежемесячно и т. д.). Увеличение продолжительности осадков усиливает пространственную зависимость экстремальных осадков. Полная независимость между ними достигается примерно через 50 км (100 км) для летнего (зимнего) при длительности, не превышающей один час, а в течение длительного времени только после нескольких сотен километров. Кроме того, эта зависимость всегда более значима зимой, чем летом, вне зависимости от продолжительности осадков.

Ключевые слова: экстремальные величины, структуры зависимости экстремального типа, копулы, вариограмма, осадки.

1. Введение

Статистический анализ максимумов основан на теории экстремальных значений (EVT) [1,2,3,6], утверждающей, что обобщенное распределение экстремальных величин является предельным распределением независимых последовательностей максимумов. Метеостанции предоставляют, как правило, последовательности измерений, связанных в общем случае нестационарными временными и пространственными зависимостями, что может приводить при использовании классической EVT к недооценке последствий наступления различных терминальных событий, а также повлиять на расчеты надежности и устойчивости гидрологических сооружений. В данной работе мы ограничиваемся исследованиями структур пространственной зависимости как функции различных параметров, таких как размер блока максимумов, длительность осадков и сезон. Для гауссовских случайных векторов (X, Y) хорошо известно, что ковариационная матрица полностью описывает их структуру зависимости. Для негауссовских распределений существует несколько подходов, чтобы выразить совместное распределение $P(X \leqslant x, Y \leqslant y)$ через их частные распределения. Один из них основан на следующем равенстве

$$P(X \leqslant x, Y \leqslant y) = [P(X \leqslant x, Y \leqslant y)]^{\theta(x,y)},$$

где $\theta(x, y)$ — неотрицательная функция. Если $\theta = 1$, то (X, Y) независимы. В настоящей работе предложен устойчивый непараметрический метод оценивания функции $\theta(x, y)$, основанный на концепции λ-мадограммы, описанной в [5].

2. Математические основы моделей пространственных структур статистической зависимости

Определим вектор $M_i(t) = \max(Z_{i,1}(t), Z_{i,2}(t), ..., Z_{i,m}(t))$ как максимум осадков, где m означает количество периодов длительности t осадков, зарегистрированных за период времени, где n соответствует количеству метеостанций. Случайная величина $Z_{i,j}(t)$ описывает количество осадков с длительностью t, выпавших на i-й станции. Например, если $M_i(t)$ представляет месячный максимум осадков часовой длительности, то $T = 1$ месяц, $t = 1$ час и $m = 24 \times 30$. Для частных распределений максимумов $F_i(x) = P(M_i(t) < x)$ и их совместных попарных распределений между двумя станциями i и k $F_{ik}(x, y) = P(M_i(t) < x, M_k(t) < y)$ фундаментальный результат многомерной теории экстремальных величин состоит в том [7], что функция $F_{ik}(x, y)$лежит в области притяжения функции распределения экстремальных величин $G_{ik}(x, y)$ [2,7]

$$G_{ik}(x, y) = \exp\left[-V_{ik}\left(\frac{-1}{\ln G_i(x)}, \frac{-1}{\ln G_k(y)}\right)\right],$$

$$V_{ik}(x, y) = 2\int\limits_0^1 \max\left(\frac{\omega}{x}, \frac{1-\omega}{y}\right) dH_{ik}(\omega), \quad H_{ik}(\omega)$$

определена на $[0, 1]$, функции $G_i(x)$, $G_k(y)$ являются частными GEV-распределениями, и $\int\limits_0^1 \omega dH_{ik}(\omega) = 0.5$. Функция $V_{ik}(\cdot)$ называется функцией парной экстремальной зависимости.

Заметим, что $\theta(x, x) = V_{ik}(1, 1)/2$. Величина $V_{ik}(1, 1)$ называется экстремальным коэффициентом, равным 2 в случае независимости и 1 в случае полной зависимости. Для оценивания $V_{ik}(1, 1)$ мы используем метод мадограмм [4], показавший, что экстремальный коэффициент может быть оценен непосредственно из выражения

$$\nu_{ik} = \frac{1}{2}E(|F_i(M_i(t)) - F_k(M_k(t))|).$$

Для независимых максимумов $\nu_{ik} = 1/6$, если $\nu_{ik} < 1/6$, то существует зависимость. Можно показать, что если вектор $(M_i(t), M_k(t))$ удовлетворяет (2), то

$$V_{ik}(1,1) = \frac{0.5 + \nu_{ik}}{0.5 - \nu_{ik}}.$$

Ограниченность подхода (2) состоит в том, что он дает оценку только $V_{ik}(1,1)$, но не дает характеризации всей функции зависимости $V_{ik}(x,y)$. Для решения этой задачи в работе [4] введен дополнительный параметр λ

$$\nu_{ik} = \frac{1}{2} E\left(\left|F_i^{\lambda}(M_i(t)) - F_k^{1-\lambda}(M_k(t))\right|\right), \quad \lambda \in [0,1].$$

Из (2) следует, что

$$\nu_{ik}(\lambda) = \frac{V_{ik}(\lambda, 1-\lambda)}{1 + V_{ik}(\lambda, 1-\lambda)} - c(\lambda),$$

где

$$c(\lambda) = \frac{3}{2(1+\lambda)(2-\lambda)}.$$

Так как

$$\lambda = \frac{x}{x+y}, \quad V_{ik}(x,y) = \frac{1}{x+y} V_{ik}(\lambda, 1-\lambda),$$

то λ-мадограмма полностью характеризует функцию зависимости $V_{ik}(x,y)\, x, y \in R^2$.

Также λ-мадограмма удовлетворяет условию $\nu_{ik}(0) = \nu_{ik}(1) = 0.25$. В итоге, это позволило нам предложить следующую оценку (2) для ν_{ik}

$$\tilde{\nu}_{ik}(\lambda) = \frac{1}{L}\sum_{l=1}^{L}\left|\tilde{F}_i^{\lambda}\left(M_{i,l}(t)\right) - \tilde{F}_k^{1-\lambda}\left(M_{k,l}(t)\right)\right| - \frac{\lambda}{2L}\sum_{l=1}^{L}\left(1 - \tilde{F}_i^{\lambda}\left(M_{i,l}(t)\right)\right) - $$
$$- \frac{1-\lambda}{2L}\sum_{l=1}^{L}\left(1 - \tilde{F}_k^{1-\lambda}\left(M_{k,l}(t)\right)\right) + \frac{1-\lambda+\lambda^2}{2(2-\lambda)(1+\lambda)}.$$

3. Исследование пространственных статистических структур зависимости экстремальных осадков

Изложенный выше метод мадограмм использован нами для исследования пространственных статистических связей осадков, наблюдаемых в европейской части России на протяжении 1999–2015 г.г. Нами был проведен полный статистический анализ измерений, который показал нестационарную временную зависимость для всех источников, тяжелохвостый характер частных распределений наблюдений, а также наличие ненулевой пространственной корреляции между практически всеми источниками наблюдений.

Исследования показали, что сильные пространственные корреляции присутствуют между максимумами, начиная с однодневных и до месячных включительно. Они также свидетельствуют о существенных отличиях между летним и зимним сезонами, как для коротких так и длинных длительностей осадков. Для обоих сезонов оказалось, что попарная независимость между максимумами, начиная с дневных значений, достигается только после расстояний между источниками около 200 км. На более коротких дистанциях зависимость несколько сильнее зимой, чем летом, что вполне ожидаемо в силу характерных для зимы метеорологических явлений.

Эта особенность, очевидно, отражает основной характер динамических процессов осадков: летом экстремальные осадки связаны в основном с грозами, масштаб и длительность которых невелики; для зимы характерно то, что поле осадков обычно занимает очень большое пространство и длительно во времени. Для более длительных дюраций мадограмма убывает по всем источникам наблюдений, но, при этом зимой мадограмма образует форму плато, на которой она практически постоянна для дюраций от 10 до 20 дней, а затем снова начинает убывать. Предположительно, это связано с распространением особых временных структур бароклинических волн в зимний период. Летом этого плато не видно, но убывание зависимости на больших расстояниях гораздо менее значимо заметно, чем зимой. Это, вероятно, отражает конвективный характер динамики осадков, что приводит к слабой зависимости для расстояний более 200 км даже для длительных осадков.

На рис. 1 представлены графики оценки λ-мадограммы по методу (2) как функции различных расстояний между метеостанциями: интервалы (0,10) км, (50,70) км, (130–150) км и (210–230) км. Непрерывная линия соответствует независимости максимумов, пунктирная линия соответствует полной зависимости.

4. Выводы

В настоящей работе были исследованы структуры пространственных зависимостей максимумов временных рядов осадков с использованием метода λ-мадограмм и предложена непараметрическая оценка коэффициента экстремальной зависимости как функции расстояния между источниками и длительностью осадков. Были получены новые результаты о статистических свойствах осадков с различной длительностью и в разные климатические периоды.

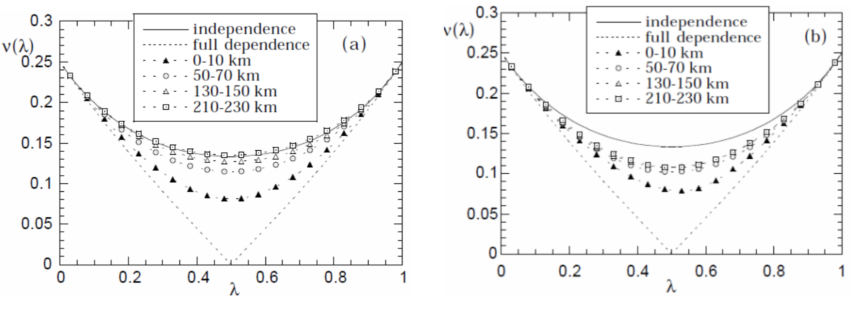

Рис. 1. λ-мадограмма для различных интервалов летом (a) и зимой (b) ежедневных максимумов

Литература

1. M. A. Ancona-Navarrete, J. A. Tawn, Diagnostics for pairwise extremal dependence in spatial processes, Extremes, 5, pp. 271–285, 2002.
2. J. Beirlant, Y. Goegebeur, J. Segers, J. Teugels, Statistics of Extremes: Theory and Applications, Wiley Series in Probability and Statistics, 2004.
3. D. Cooley, P. Naveau, P. Poncet, Variograms for spatial maxstable random fields, Chapter of the book Statistics for dependent data, Lecture Notes In Statistics, Springer, 2006.
4. D. Cooley, D. Nychka, P. Naveau, Bayesian Spatial Modelingof Extreme Precipitation Return Levels, J. Am. Statist. Assoc., 2007.
5. P. Embrechts, C. Kluppelberg, T. Mikosch, Modelling Extremal Events for Insurance and Finance, Vol. 33 of Applications of Mathematics, Springer-Verlag, Berlin, 1997.
6. T. Hsing, C. Kluppelberg, G. Kuhn, Dependence estimation and visualisation in multivariate extremes with applications to financial data, Extremes, 7, p.99–121, 2004.
7. В. А. Акимов, А. А. Быков, Е. Ю. Щетинин, Введение в статистику экстремальных значений и ее приложения, М.: ФГУ ВНИИ ГОЧС (ФЦ), 2009, с. 536.
8. M. Schlather, Models for stationary max-stable random fields, Extremes (2002), 5, pp. 33–44.
9. M. Schlather, J. Tawn, A dependence measure for multivariate and spatial extreme values: Properties and inference, Biometrika (2003), 90, pp. 139–156.