Journal of Hydrology 541 (2016) 1185–1198

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

# A regional model for extreme rainfall based on weather patterns subsampling

G. Evin [a,b,*], J. Blanchet [a,b], E. Paquet [c], F. Garavaglia [c], D. Penot [c]

[a] Univ. Grenoble Alpes, LTHE, 38000 Grenoble, France
[b] CNRS, LTHE, 38000 Grenoble, France
[c] EDF - DTG, 21 Avenue de l'Europe, BP 41, 38040 Grenoble Cedex 9, France

ARTICLE INFO

*Article history:*
Received 7 April 2016
Received in revised form 6 July 2016
Accepted 13 August 2016
Available online 20 August 2016
This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Uwe Haberlandt, Associate Editor

*Keywords:*
Regional frequency analysis
Extreme rainfall
Weather pattern

SUMMARY

Many rainfall generators rely on the assumption that statistical properties of rainfall observations can be related to physical processes via weather patterns. The MEWP (Multi-Exponential Weather Pattern) model belongs to this class. In this daily rainfall model, extremes above a threshold are distributed exponentially, for each season and atmospheric circulation pattern. A wide range of applications of this rainfall compound distribution has demonstrated its robustness and reliability. However, recent investigations showed that MEWP tends to underestimate the most extreme rainfall events in specific regions (e.g. the South-East of France).

In this paper, we apply different versions of a generalized MEWP model: the MDWP (Multi-Distribution Weather Pattern) model. In the MDWP model, the exponential distribution is replaced by distributions with a heavier tail, such as the Generalized Pareto Distribution (GPD). Unfortunately, local applications of the MDWP model reveal a lack of robustness and overfitting issues. To solve this issue, a regional version of the MDWP model is proposed.

Different options of a regionalization approach for excesses are scrutinized (e.g. choice of the scale factor, testing of homogeneous regions based on neighborhoods around each site, choice of the distribution modelling extreme rainfall). We compare the performances of local and regional models on long daily rainfall series covering the southern half of France. These applications show that the local models with heavy-tailed distributions exhibit a lack of robustness. In comparison, an impressive improvement of model robustness is obtained with the regional version, without a loss of reliability.



## 1. Introduction

Design of water-related infrastructures often applies weather generators as a probabilistic tool to assess risks related to extreme events. Among the great variety of weather generators, an important class of models makes the link between the generating physical processes of meteorological events and statistical models via a limited number of weather "types" or "patterns" (Ailliot et al., 2015). Following the assumption that such classification generally improves the statistical homogeneity of fitted observations, MEWP (Garavaglia et al., 2010, 2011) categorizes rainfall events according to sub-populations corresponding to eight atmospheric circulation patterns identified at a large scale (e.g. over Western Europe). For each sub-population and season, an exponential distribution is fitted to a Peak-Over-Threshold (POT) sample composed of the largest rainfall events, i.e. exceeding a pre-determined quantile (e.g. the 0.7-quantile at a daily scale). The global MEWP distribution is obtained by combining the exponential distributions.

Several studies in France (Garavaglia et al., 2010, 2011), Canada, Austria (Brigode et al., 2013) and Norway (Blanchet et al., 2015) show that MEWP often outperforms other distributions commonly used to model extreme rainfall (e.g. Gumbel, GPD and GEV distributions) in terms of reliability and robustness. However, the assumption of severe rainfall events being exponentially distributed can be put into question, especially in regions where flash-floods are likely to occur (e.g. the South-East of France, see Veysseire et al., 2012; Neppel et al., 2014). For this reason, we propose to apply the MDWP model, for which the exponential distribution can be replaced by distributions with a heavier tail, such as the GPD. Unfortunately, local applications of the MDWP model reveal a lack of robustness and overfitting issues (Garavaglia et al., 2011). Indeed, the application of heavy-tailed distributions to short record lengths of rainfall extremes leads to potential

* Corresponding author at: Univ. Grenoble Alpes, LTHE, 38000 Grenoble, France.
*E-mail address:* guillaume.evin@ujf-grenoble.fr (G. Evin).

http://dx.doi.org/10.1016/j.jhydrol.2016.08.024
0022-1694/

estimation biases (Papalexiou and Koutsoyiannis, 2013; Serinaldi and Kilsby, 2014).

To solve these estimation issues, we propose to apply a regionalized version of MDWP. A regional frequency analysis (RFA) is developed in order to increase the sample size for each season and weather pattern (WP). The RFA approach (Hosking and Wallis, 2005) combines observations at a given site with the information gathered in a region around this site. While the RFA approach has been extensively applied to annual rainfall maxima (for recent applications see, e.g. Carreau et al., 2013; Du et al., 2014), fewer applications to POT rainfall series can be found, with the notable exceptions of Madsen et al. (1997), Roth et al. (2012), and Mailhot et al. (2013). In this paper, we develop a RFA methodology specific to the MDWP model and study the impact of several choices (scale factor, homogeneity tests) on its performances. The regionalized MDWP is expected to show a significant gain of robustness when heavy-tailed distributions are applied, while preserving the reliability qualities of the MEWP model in regions where the local MEWP shows adequate performances. In addition, the possibility of extending the application of MDWP to ungauged sites is of strong interest.

The development of the RFA methodology is based on an extensive dataset of 773 daily rainfall series. Neighborhoods around each site are considered, following the concept of regions-of-influence (RoI) proposed by Burn (1990). The paper is organized as follows: The rainfall dataset and the localization of the meteorological stations are presented in Section 2. Section 3 discusses the nonstationarity of rainfall extremes in France. Section 4 details the local MDWP model and Section 5 the regional MDWP model. A bottom-up algorithm for the selection of RoIs is proposed, different options of scale factors are tested and several statistical homogeneity tests are reviewed. Section 6 describes the criteria used to evaluate the reliability and robustness of the different probabilistic models. Section 7 assesses the performance of local and regional models, when different alternatives to the exponential distribution are considered. Section 8 concludes and draws some perspectives.

## 2. Data

Daily rainfall observations from $N = 773$ rain gauges belonging to EdF (Électricité de France), the French meteorological office Météo-France and the Italian meteorological Service are used in this study. Only rain gauges with at least 20 years of records over the period 1948–2013 have been selected, most of the stations having more than 50 years of data (see Table 1). These stations are mainly located in the South-Eastern part of France and at borders (see Fig. 1). Rain gauges networks are historically denser over mountainous areas, where water management stakes are important. Regions covering the Alps, Pyrenees and Massif Central are thus particularly well monitored, although high-altitude zones are usually poorly instrumented (Gottardi et al., 2012).

Table 1 gives some properties of daily rainfall measurements and shows a great variability of rainfall regimes, the average annual total $\overline{TOT}$ ranging from 523 [mm] to 2177 [mm] and the average annual maxima $\overline{MAX}$ ranging from 34 [mm] to 201 [mm] among the 773 stations. The ratio $\overline{MAX}/\overline{TOT}$ indicates the average proportion of the annual maxima in the annual total rainfall and shows how the intense values depart from the average ones.

**Table 1**
Some statistics of the daily rainfall series: Number of years spanned by the data; maximum rainfall [mm]; average annual total $\overline{TOT}$ [mm]; average annual maximum $\overline{MAX}$ [mm]; ratio $\overline{MAX}/\overline{TOT}$.

| | Min | $Q_{25}$ | $Q_{50}$ | $Q_{75}$ | Max |
|---|---|---|---|---|---|
| Number of years | 20 | 48 | 57 | 63 | 66 |
| Maximum rainfall [mm] | 50 | 99 | 124 | 184 | 551 |
| $\overline{TOT}$ [mm] | 523 | 840 | 1012 | 1250 | 2177 |
| $\overline{MAX}$ [mm] | 34 | 52 | 60 | 77 | 201 |
| $\overline{MAX}/\overline{TOT}$ | 0.03 | 0.05 | 0.06 | 0.08 | 0.16 |

As shown in Fig. 1, this criteria exhibits a clear geographical pattern with the most intense rainfall maxima being observed around the Mediterranean coast, in the south-east part of the Massif Central and in the low Rhône valley. The specificity of this region regarding extreme rainfall is well documented (see Neppel, 1998; Moliné et al., 2011), and a delimitation is proposed using the 0.9-quantile of $\overline{MAX}/\overline{TOT}$ within the whole dataset: zone B (with heavy Mediterranean precipitations) is defined as a region which roughly gathers stations exceeding this quantile ($\overline{MAX}/\overline{TOT} > 0.09$), and zone A is its complement.

As a recent study suggests that MEWP tends to underestimate the most extreme events in zone B (Neppel et al., 2014), the performances of the different rainfall extreme models will be computed separately in zones A and B.

## 3. Nonstationarity of rainfall extremes in France

According the last IPCC report (IPCC, 2014), some properties of rainfall extremes clearly exhibit a trend. For example, it has been confirmed that the number of heavy precipitation events has globally increased in land regions. Therefore, the nonstationarity of rainfall extremes cannot be ignored and must be discussed.

In Europe, the frequency and intensity of heavy precipitation events has likely increased (IPCC, 2014). However, for extreme precipitation, van den Besselaar et al. (2013) report a median reduction in 5- to 20-year return periods, which varies substantially according to the subregion and season. In Southern France, several studies (see, e.g., Neppel et al., 2011; Tramblay et al., 2013; Vautard et al., 2015) have been devoted to the analysis of the stationarity of rainfall extremes and leads to different conclusions. Pujol et al. (2007a,b) and Neppel et al. (2011) analyze regional trends in annual maximum series of precipitation in seven homogeneous climatological zones in southern France. Pujol et al. (2007a) study the regional evolution of daily peaks-over-threshold records and find an increase of the occurrence and intensity of extreme daily rainfall in the southern part of the Massif Central and Pujol et al. (2007b) find a significant increase of annual maximum series in the same region. Neppel et al. (2011) also observe a slight but significant trend of annual maximum series in the mountainous area around the Cévennes-Vivarais mountain range. However, for a large region in the South-East of France which contains our zone B, no significant trend is detected. Focusing on the southern region including the Cévennes-Vivarais, Tramblay et al. (2013) develop a non-stationary model and obtain quantile estimates that do not differ significantly from the stationary ones. In the French Mediterranean area, local upward trends in annual maxima of daily precipitation found by Soubeyroux et al. (2015) are usually not significant. Studying seasonal maxima of daily precipitation in the Cévennes-Vivarais range, Vautard et al. (2015) find local upward trends with a median increase of about 5% per decade but few of these local trends are significant. Finally, in a recent study, Blanchet et al. (2016) analyze these trends in a region around the Cévennes-Vivarais range which covers approximately half of the stations used in this study. They find significant trends in a sub-region of zone B including the Cévennes slope and part of the Rhône valley, but this corresponds to less than 5% of the total region studied in this paper.

Brigode (2013) investigates the impact of climate change in the context of the SCHADEX method, and analyzes the Clausius-Clapeyron relationship in France. While significant correlations