Modelling maximum precipitation in a mountainous area of Greece under global warming

DIONYSIA PANAGOULIA[1] & CHRYS CARONI[2]

1 *School of Civil Engineering, Dept of Water Resources & Environmental Engineering, National Technical University of Athens, 5 Heroon Polytechniou, 15780 Zografou, Athens, Greece*
dpanag@hydro.civil.ntua.gr

2 *School of Applied Mathematical & Physical Sciences, Dept of Mathematics, National Technical University of Athens, 9 Heroon Polytechniou, 15780 Zografou, Athens, Greece*

Abstract We investigated the fit of generalized extreme value (GEV) distributions to maximum precipitation over the Mesochora mountainous catchment in central-western Greece under present and future climate scenarios. Precipitation was modelled as a stochastic process coupled with atmospheric circulation models. Automated objective classification based on optimized fuzzy rules was used to classify observed circulation patterns (CPs) and ECHAM4 General Circulation Model-generated CPs for 1×CO_2 and 2×CO_2 climate scenarios. The GEV distribution was fitted by maximum likelihood, allowing for non-stationarity over time in its location and scale parameters. The stationary model was adequate for historical data on annual daily maxima for 1972–1992 and also for 1×CO_2 for the period 1961–2000. However, the 2×CO_2 series for 2061–2100 required a cubic time trend in location to obtain a satisfactory fit ($P<0.0001$ by likelihood ratio test). This series declined to a minimum around 2080, followed by an increase to a maximum around 2092, and subsequently a further decline.

Key words generalized extreme value distribution; maximum likelihood; global warming; annual maximum precipitation; likelihood ratio tests; non-stationarity; Greece

INTRODUCTION

The analysis of extremes in hydro-meteorological data, such as the annual or monthly maxima in precipitation and discharge series, is fundamental for the design of engineering structures (Maidment, 1993). These maxima can be modelled asymptotically using the generalized extreme value (GEV) distribution (Jenkinson, 1955). The assumption of independent and identically distributed data in the series with constant properties through time (stationarity) may need to be modified to reflect climate change. There is mounting evidence that hydro-climatic extreme series are not stationary, owing to natural climate variability or anthropogenic climate change (Jain & Lall, 2001; Milly *et al*., 2008). The modelling of non-stationarity within the framework of the GEV distribution requires extended models with covariate-dependent changes in at least one of the distribution's three parameters (location, scale and shape) (Coles, 2001).

In recent studies along these lines, parameter estimates were obtained by the maximum likelihood method (ML) (Wang *et al*., 2004) or generalized maximum likelihood (GML) (El Adlouni *et al*., 2007; Cannon, 2010). In the studies of Wang *et al*. (2004) and El Adlouni *et al.* (2007), dependence of the GEV parameters on covariates was modelled by linear or log-linear parametric models. In the study of Cannon (2010), the parameters were specified via a function of covariates conditioned on a probabilistic extension of the multilayer perceptron neural network that allows unspecified interactions between multiple covariates. Both El Adlouni *et al.* (2007) and Cannon (2010) analysed annual maximum precipitation data recorded at Randsburg, California, and the Southern Oscillation Index (SOI) was taken to be the covariate process.

In this paper we take a somewhat different approach, simulating climate change via precipitation modelling as a stochastic process coupled with atmospheric circulation. An automated and objective classification of daily circulation patterns (CPs) based on optimized fuzzy rules was used to classify both observed CPs and ECHAM4 General Circulation Model (GCM)-generated CPs for 1×CO_2 and 2×CO_2 climate scenarios (Panagoulia *et al*., 2006a,b, 2008). From the resulting daily precipitation we calculated the annual daily maximum series over the Mesochora mountainous catchment in central-western Greece. The objective of this study is to investigate the fit of generalized extreme value (GEV) distributions to the series of historical data for 1972–1992, 1×CO_2 data for 1961–2000 and 2×CO_2 data for 2061–2100. To this end, we fitted the GEV distribution by maximum likelihood, allowing for non-stationarity over time in its parameters of location $\mu(t)$ and scale $\sigma(t)$, and compared different models by likelihood ratio tests.

PRECIPITATION DATA SERIES

The mountainous Mesochora catchment (Fig. 1) is the most upstream sub-catchment of the Acheloos River catchment, which lies in the western-central mountain region of Greece and has an area of about 633 km^2. A reservoir (useful capacity of 228 hm^3) has been constructed at the outlet of the catchment and a hydropower plant with installed capacity of 160 MW. Mean annual discharge is 23.2 $m^3 s^{-1}$. Precipitation stations are installed within and around the catchment, mostly in the lower half over a range of elevations from 780 to 1160 m. Daily measurements of precipitation were available at 12 stations over the period 1972–1992.

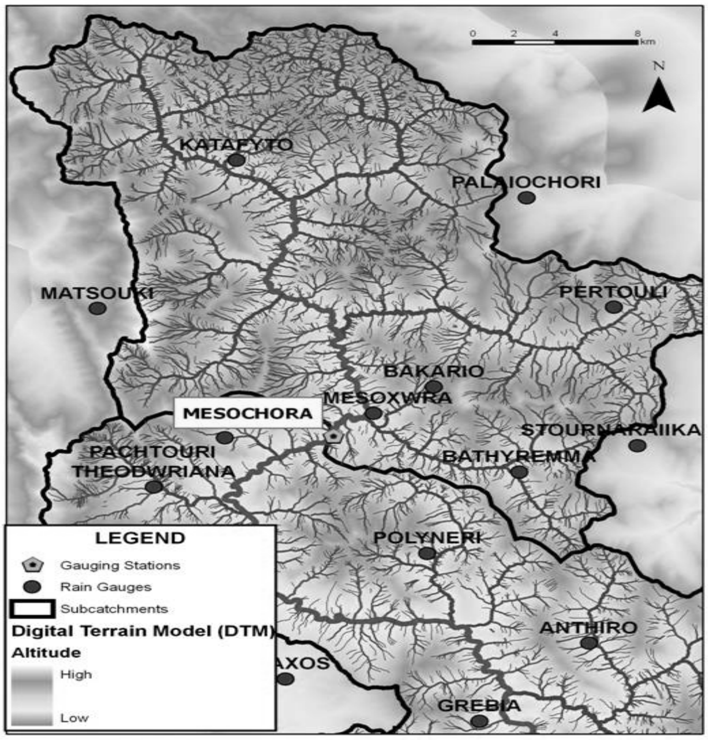

Fig. 1 Map of Mesochora catchment showing the available raingauges.

The precipitation variability at the stations was determined by conditioning on atmospheric circulation patterns (CPs). These were classified via the fuzzy-rules based approach combined with the simulated annealing algorithm (Bárdossy *et al.*, 2002; Panagoulia *et al.*, 2006a). Pressure data were obtained from the NMC grid-point data set for different windows over Europe with a grid resolution of 5° × 5°. It was found that the 700 hPa data in the window provided the best results, with an optimal number of 12 CPs based on the automated objective optimization procedure (Panagoulia *et al.*, 2006a).

Space–time intermittence, the occurrence probability of dry days, the rainfall amounts on wet days, and the clustering of wet and dry day occurrence that has significant impact on the persistence of CPs were taken into account in the stochastic modelling of daily precipitation, adopting the methodology of Stehlik & Bárdossy (2002). The observed daily precipitation series for all available periods were used to estimate precipitation coupling parameters to describe the stochastic links between CPs and point measurements. Precipitation time series were simulated using these estimated parameters (Panagoulia *et al.*, 2006b). A spatial correlation function using time series cross-correlations was assessed for extrapolation of point precipitation data to the entire area. For this purpose, external drift kriging was used (Ahmed & de Marsily, 1987).

Beyond the CP-dependent observed daily precipitation over the Mesochora catchment for the period 1972–1992, downscaling was carried out for ECHAM4 GCM-generated precipitation. The analysis was based on daily values in the relevant sector 20°–65°N, 20°W–50°E over the 700 hPa pressure field for 1×CO_2 and 2×CO_2 climate scenarios in the corresponding periods 1961–2000 and 2061–2100. The geo-potential pressure heights (the 700 hPa pressure) for both scenarios were classified by applying the same method as described above for the observed data. With the estimated parameters of the stochastic precipitation model for the observed data and the classified GCM-CPs, precipitation time series were generated representing two climate scenarios. The series of historical data and the two climate scenarios are shown in Figs 2–4, with trends indicated by lines obtained using the lowess smoother in the Minitab package (www.minitab.com).

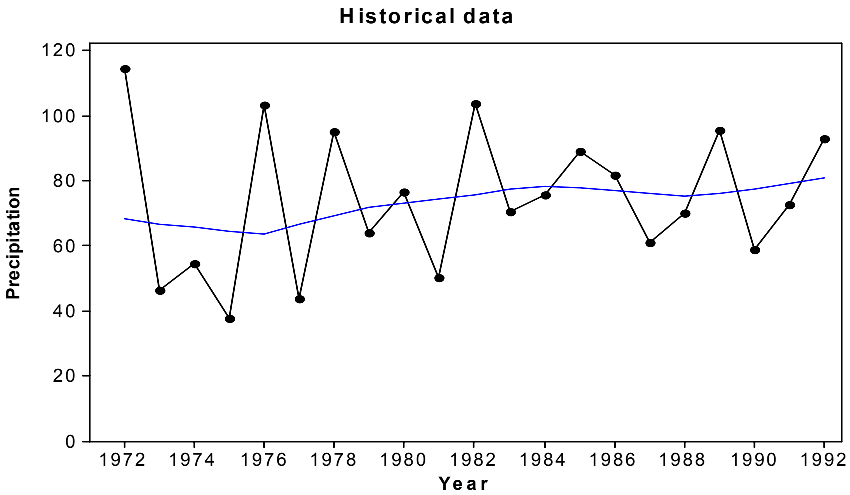

Fig. 2 Historical precipitation data for 1972–1992 with trend fitted by lowess smoother.

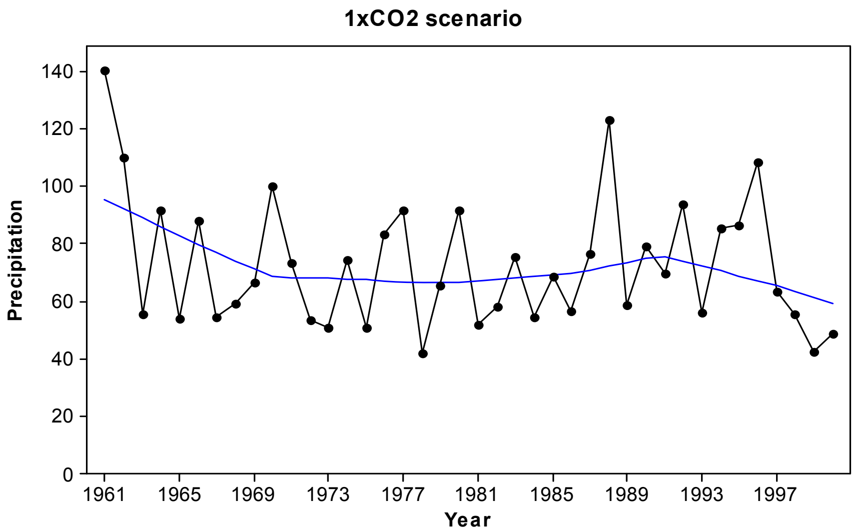

Fig. 3 1×CO_2 series for 1961–2000 with trend fitted by lowess smoother.

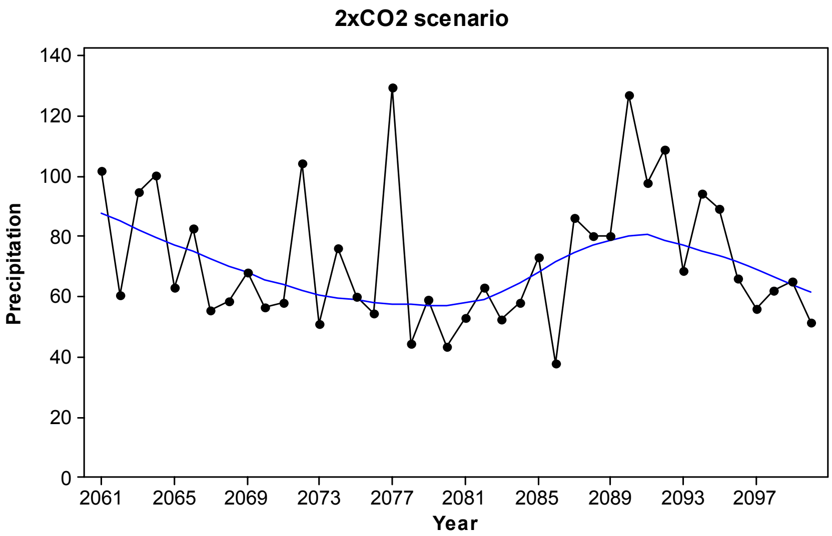

Fig. 4 2×CO_2 series for 2061–2100 with trend fitted by the lowess smoother.

FITTING THE NON-STATIONARY GEV

The distribution function of the GEV distribution is:

$$F(x;\mu,\sigma,\xi)=\begin{cases}\exp\left\{-\left[1+\dfrac{\xi(x-\mu)}{\sigma}\right]^{-\frac{1}{\xi}}\right\} & \xi\neq 0\\ \exp\left\{-\exp\left[-\dfrac{x-\mu}{\sigma}\right]\right\} & \xi=0\end{cases}$$

where, μ, σ and ξ are the location, scale and shape parameters, respectively. The range of the distribution is given by $1+\dfrac{\xi(x-\mu)}{\sigma}>0$ if $\xi \neq 0$; otherwise x is unrestricted. Non-stationarity in a time series can be modelled by allowing some or all of these parameters to be functions of time (or other covariates). In fact, we followed the usual practice of permitting μ and σ to vary but keeping ξ constant (Nogaj *et al.*, 2007). We fitted the GEV distribution by maximum likelihood and used likelihood ratio tests to examine the significance of linear, quadratic and cubic terms in time in $\mu(t)$ and $\sigma(t)$. Fitting was carried out in the R programming language, using the *gev* routine from the *ismev* package (available from cran.r-project.org/web/packages/ismev, accessed 6/01/11). The ML method is sometimes criticized because it can give rise to physically unacceptable estimates; the GML method of Martins & Stedinger (2000) avoids this by, in effect, constraining ξ within the range (–0.5, 0.5) (Katz *et al.*, 2002). However, as seen below, the unconstrained ML estimates of this parameter from our data fell well within this range anyway.

RESULTS

Historical data, 1972–1992

The first data series consisted of the CPs-dependent observed annual daily maximum precipitation for the period between 1972 and 1992. Stationary and non-stationary GEV distributions were fitted. The change in minus twice the log likelihood as a result of allowing the location to be a linear function of time rather than constant, was 2.97, with P = 0.085 from the chi-squared distribution with one degree of freedom. Likewise, a linear term in scale did not improve fit

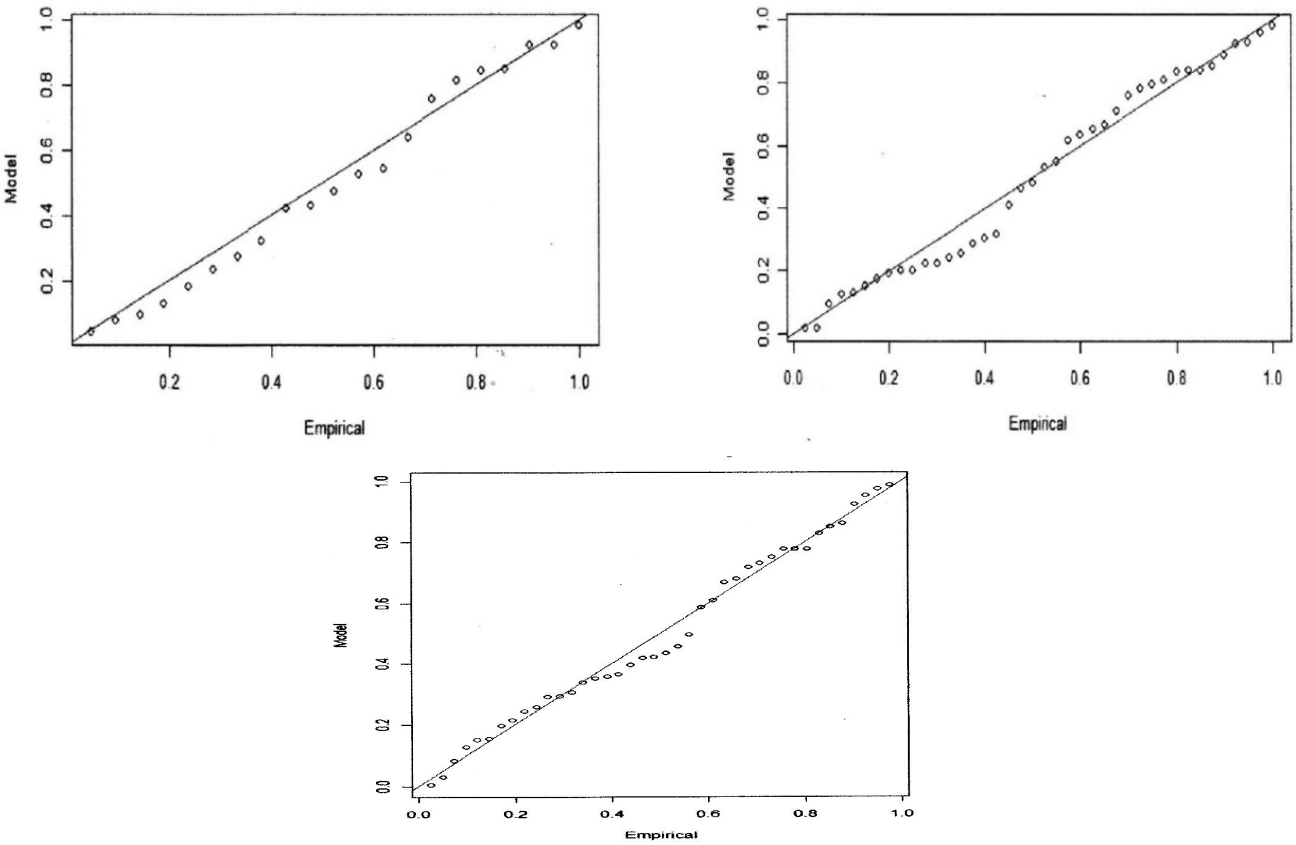

Fig. 5 Probability plots of fit of stationary GEV to historical series (upper left) and $1\times CO_2$ series (upper right); plot of residuals from fit of GEV with cubic term in location parameter to $2\times CO_2$ series (lower).

Table 1 Parameter estimates for the GEV distribution fitted to the three series of data. Estimated standard errors are shown in parentheses.

Parameter	Term*	Historical 1972–1992	$1\times CO_2$ 1961–2000	$2\times CO_2$ 2061–2100
Location μ	Constant	67.06 (5.27)	61.45 (2.89)	62.81 (2.86)
	Linear			11.06 (8.63)
	Quadratic			4.80 (2.17)
	Cubic			–12.44 (7.99)
Scale σ	Constant	20.87 (4.02)	15.52 (2.27)	15.45 (2.13)
Shape ξ	Constant	–0.31 (0.21)	0.14 (0.16)	0.03 (0.15)

*Following the recommendation in the documentation of the *ismev* package, each term was centred and scaled to unit variance before fitting.

significantly (P = 0.15), nor did higher order terms in either parameter. The probability plot of the fit was satisfactory (Fig. 5) and it was concluded that the stationary GEV distribution provided an adequate description of the historical data. Parameter estimates for all analyses are shown in Table 1.

$1\times CO_2$ series, 1961–2000

A similar analysis was applied to the ECHAM4 GCM-generated scenarios of daily CPs for the $1\times CO_2$ climate scenario for the period 1961–2000. There was no indication that linear terms (P = 0.43 for location, P = 0.67 for scale) or higher order terms were required. The probability plot again showed that the stationary GEV distribution provided a good fit to the series (Fig. 5).

2×CO_2 series, 2061–2100

A third analysis was carried out along the same lines of the ECHAM4 GCM-generated scenario of daily CPs for the 2×CO_2 climate scenario for the period 2061–2100. In contrast to the historical data and the 1×CO_2 series, the stationary model did not provide an adequate fit to the 2×CO_2 series. It was found that the cubic term in time in the location parameter was statistically significant (X^2_3 = 15.9, P < 0.0001), although once again the scale parameter appeared to be stationary (P = 0.92). The probability plot of the residuals was well behaved (Fig. 5).

CONCLUSIONS

In this study we investigated non-stationarity (time dependence) in the location and scale parameters of the GEV distribution fitted to annual daily maxima of precipitation over a mountainous catchment in Greece under present and future climate scenarios. The results in Table 1 show that the shape parameters did not differ significantly from zero. Thus the Gumbel distribution (GEV with ξ = 0) appears to provide an adequate fit to all three series. Non-stationarity was detected only for the 2×CO_2 series for the period 2061–2100. A cubic term in time in the location parameter was necessary to describe this series' decline to a minimum around 2080, followed by an increase to a maximum around 2092, and subsequently a further decline. Recognizing the small length of data series and the uncertainties of the precipitation modelling, the non-stationary GEV model presented in this paper is an important tool for use to take into consideration the temporal evolution of the climate. Such a tool is of great significance for the design of hydraulic structures under climate change.

REFERENCES

Ahmed, S. & De Marsily, G. (1987) Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. *Water Resour. Res.* **23**(9), 1717–1737.

Bárdossy, A., Stehlík, J. & Caspary, H. J. (2002) Automated objective classification of daily circulation patterns for precipitation and temperature downscaling based on optimized fuzzy rules. *Climate Res.* **23**, 11–22.

Cannon, A.J. (2010) A flexible nonlinear modeling framework for nonstationary generalized extreme value analysis in hydroclimatology. *Hydrol. Processes* **24**(6), 673–685.

Coles, S. (2001) *An Introduction to Statistical Modelling of Extreme Values*. Springer, London, UK.

El Adlouni, S., Ouarda, T. B. M. J., Zhang, X., Roy, R. & Bobée, B. (2007) Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resour. Res.* **43**, W03410, doi:10.1029/2005WR004545.

Jain, S. & Lall, U. (2001) Floods in a changing climate: does the past represent the future? *Water Resour. Res.* **37**(12), 3193–3205.

Jenkinson, A. F. (1955) The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quart. J. Royal Met. Soc.* **81**(348), 154–171.

Katz, R. W., Parlange, M. B. & Naveau, P. (2002) Statistics of extremes in hydrology. *Adv. Water Resour.* **25**(8-12), 1287–1304.

Maidment, D. R. (1993) *Handbook of Hydrology*. McGraw-Hill Professional, New York, USA.

Martins, E. S. & Stedinger, J.R. (2000) Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resour. Res.* **36**(3), 737–744.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P. & Stouffer, R. J. (2008) Stationarity is dead: whither water management? *Science* **319**, 573–574.

Nogaj, M., Parey, S. & Dacunha-Castele, D. (2007) Non-stationary extreme models and a climatic application. *Nonlin. Processes Geophys*. **15**, 305–316.

Panagoulia, D., Bárdossy, A. & Lourmas, G. (2008) Multivariate stochastic downscaling model for generating precipitation and temperature of climate change based on atmospheric circulation atmospheric circulation. *Global Nest J.* **10**(2), 263–272.

Panagoulia, D., Grammatikogiannis, A. & Bárdossy, A. (2006a) An automated classification method of daily circulation patterns for surface climate data downscaling based on optimised fuzzy rules. *Global Nest J.* **8**(3), 218–223.

Panagoulia, D., Bárdossy, A. & Lourmas, G. (2006b) Diagnostic statistics of daily rainfall variability in an evolving climate. *Adv. Geosciences* **7**, 349–354.

Stehlik, J. & Bárdossy, A. (2002) Multivariate stochastic downscaling model for generating daily precipitation series based on atmospheric circulation. *J. Hydrol.* **256**(1-2), 120–141.

Wang, X. L., Zwiers, F. W. & Swail, V. R. (2004) North Atlantic Ocean wave climate change scenarios for the twenty-first century. *J. Climate* **17**, 2368–2383.