A GENERALIZATION OF THE LOG-LOGISTIC DISTRIBUTION FOR THE PROBABILITY OF DAILY PRECIPITATION

Roberto Moncho[1,2]

[1]Fundación para la Investigación del Clima, C/ Gran Vía, 22do, 6º Izda. Madrid, Spain.

[2]Departament of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia. N. 50, Dr. Moliner. 46100 Burjassot, Valencia, Spain.

*Corresponding author, e-mail address: robert@temps.cat

Introduction

The **daily precipitation in climate models** can be used for various environmental studies, but it has a **systematic error** that is difficult to correct. Probability of distribution is very high around the lower rainfall (most likely zero) and decreases sharply to high rainfall. For that reason, the commonly used models have **problems to fit all daily precipitation**.

The objective of this work is to present and evaluate an alternative model of probability, based on a generalization of the Log-Logistic Distribution. Some **test of goodness-of-fit** was done by comparing the Akaike Information Index, Anderson-Darling, Kolomogorov-Smirnov and the Normalized Mean Absolute Error (NMAE). The proposed model and the commonly used probability models (GEV, GPD, Gamma, Gumbel, Weibull, Exponential and Log-normal) have been applied to daily rainfall for **108 observatories from Spain**

Methodology

The proposed model is a **modification of log-logistic distribution** (**Fig. 1**) according to:

$$\pi(p \geq P) = \frac{1}{1+\lambda^{w+e^{-k}\lambda^{w}}} \qquad \text{where:} \qquad \lambda = \frac{P-P_o}{P_1}$$

where $\pi(p \geq P)$ is the *cumulative probability*, i. e., the probability that a station registers a precipitation equal to or higher than *P*; whilst *Po*, *P1*, *w*, and *k* are positive parameters. To compare the results, a **modified version of Gumbel Distribution** were used:

$$\pi(p \geq P) = \exp\left[-\exp\left(\lambda^{w}+k\right)\right]$$

The proposed model and commonly used models were adjusted for daily precipitation of every month of year, for **108 observatories from Spain**. To adjust all the functions we have used the maximum likelihood inference[1] (Profile Log-Likelihood). In addition, **some test of good-of-fit were applied**: NMAE, Akaike Index[2], Anderson-Darling[3] and Kolmogorov-Smirnov[4] tests.

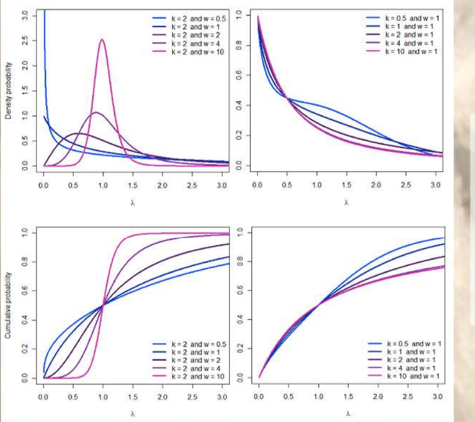
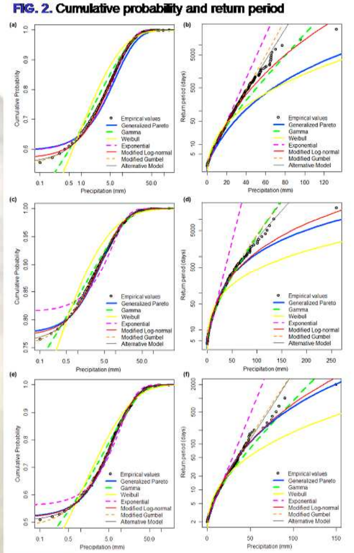


FIG. 1. Some examples of probability density curve (top) and cumulative probability curve (bottom), with w = 1 and k = 1, 2, 10; and with k = 2 and w = 1, 2, 10.

Results

FIG. 2. Cumulative probability and return period



Cumulative probability and return period were obtained according to the empirical values and several models fited, for all observatories. For example, **Fig. 2** show three observatories with different climates: (a) and (b) Tenerife - Los Rodeos (subtropical climate); (c) and (d) A Coruña (oceanic climate); and (e) and (f) Navacerrada (mountain climate).

In general, Modified Gumbel and Alternative Model (generalized Log-Logistic) **fits the high and low precipitation better** than other models.

The Modified Gumbel and the Alternative Model showed the best NMAE, and it was lower than 0.10, whilst the other models as Log-Normal III, GPD and Modified GEV presented a NMAE which ranges generally between 0.10 and 0.20 (**Fig. 3**).

FIG.3. Normalized Mean Absolute Error (NMAE) of the predicted precipitation was compared *versus* the observed prediction of empirical distribution function, for the set of 108 observatories, with the adjustment of: (a) Several models: Gamma, Weibull, Exponnetial, Log-normal III, Gen. Pareto Distr. (GPD), Modified Gen. Extreme Value (M. GEV), Modified Gumbel and Generalized Log-Logistic (GLL). (b) Comparison of modified Gumbel and generalized Log-Logistic, applied for the monthly precipitation.

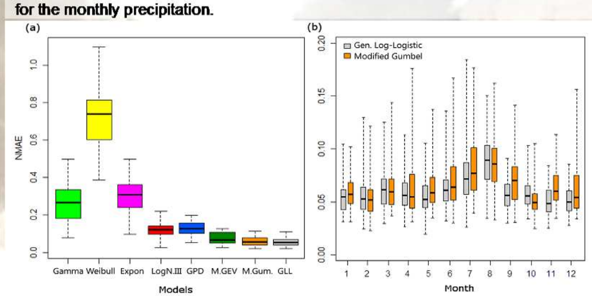


FIG. 4. Akaike Information Criterion of Log-Likelihood

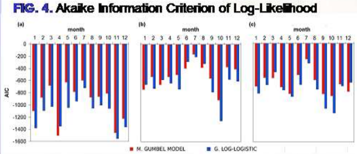
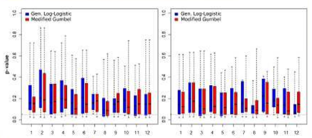


FIG. 5 . P-value of K-S and A-D test



Discussion

- The normalized **absolute error is similar for all scales** of precipitation (**Fig. 2**), i.e., the **generalized Log-Logistic** can accurately represent the **entire range of probability of daily precipitation.**
- In most cases, **this alternative model** shows **better results** than the modified version of Gumbel Distribution (**Fig. 3, 4 and 5**).
- The fact of using four parameters can lead to **difficulties in adjustment by computer algorithms**, as it requires initial values very close to the best estimates of the parameters.
- This model can be used to characterize and **correct the probability distribution** of any set of daily precipitation data
- The **parameters** fitted for a probability distribution of precip. should adequately **reflect climate characteristics (see ppt)**.
- The fit of the alternative model of probability for the different months of the year and different zones of Spain **suggests a good performance for different climates**.

Conclusions

- Most commonly used models were not able to adjust precisely enough the set of daily precipitation, included every frequency. It was therefore **necessary to use models with three or four parameters**.
- The **generalized Log-Logistic model** with four parameters **presented best results** than other models for the most observatories.
- To be precise, it obtained a **mean absolute error of less than 10%** in most of the stations analyzed, and it is similar for all scales of precipitation.
- The fit is worse for the summer months, due to the few days of precipitation; therefore, probably the **proposed model works best for rainy climates**.
- This model can be used **to characterize the climate patterns** of precipitation and **to correct numerical model outputs**.

Acknowledgments

- Department of Earth Physics of the University of Valencia.
- State Meteorological Agency of Spain (AEMET) and Hydrographics Confederations of Ebro (CHE) and Júcar (CHJ) for providing the data.
- Thank to Jose Angel Nunez, head of the Dep. of Clim. of AEMET delegation in Valencia, and Margarita Martín, AEMET delegate in the Basque Country.
- Support of Maddalen Mendizabal (Tecnalia) and especially for raising the issue of probability of daily precipitation.

References

1. Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J 2009. Structural and practical identifiably analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics 25: 1923–9.
2. Burnham, K.P. & Anderson, D. 2002. Model Selection and Multi-Model Inference. Springer-Verlag, NY,
3. Marsaglia, G,; Tsang, W.W. & Wang, J. 2003. Evaluating Kolmogorov's distribution. J. Statistical Software, 8/18
4. Scholz, F.W. & Stephens, M.A. 1987. K-sample Anderson-Darling Tests, Journal of the American Statistical Association, 399: 918–924

Photo of León Forado (2005)

Outstanding Student Poster Contest