ABOUT TIME OF OCCURRENCE OF RAINY DAYS FOR MEDITERRANEAN AND (SUB)-ALPINE AREAS

Agnese C.[1], Baiamonte G.[1]*, Cammalleri C.[2], Cat Berro D.[3], Ferraris S.[4], Mercalli L.[3]

[1] Dept. SAGA, University of Palermo, Italy.
[2] U.S. Dept. of Agriculture, ARS, Hydrology and Remote Sensing Lab., Beltsville, MD, USA
[3] Società Meteorologica Italiana, Castello Borello, Bussoleno, Italy
[4] Dept. EIAFA, University of Torino, Italy
* giorgio.baiamonte@unipa.it

Abstract
The daily scale is the maximum time-scale at which both clustering and intermittency characters of rain are still evident. These features can be viewed as a natural tendency to the persistence in time of the two opposite atmospheric states, i.e. rainy and dry (not rainy) states. In this work, the probabilistic structure of inter-arrival times, T, is analysed with the aim to account for the statistics of both clustering of rainfall and persistence of dry periods in a single distribution. A discrete probability distribution, the three-parameter Lerch distribution, has been applied to some inter-arrival time series derived from daily rainfall data recorded in Sicily and in Piedmont and in Aosta Valley, representing two different climatic environments. Statistical analyses have been carried out by considering two “seasons” of six months each: a growing season (GS) from April to September and a dormant season (DS) from October to March. Parameters have been estimated by using the maximum likelihood method and the performance of model fitting was evaluated by means of a chi-square test based on Monte Carlo procedure.

Keywords: Rainfall daily time-series, intermittency, inter-arrival times distribution, Lerch probability distribution.

Introduction

A peculiar character of rain is the intermittency that manifests by alternating wet and dry periods. The daily scale is the maximum time-scale at which this behaviour manifests. Usually, the probability of the length of wet and dry periods are analysed separately. An alternative to this approach is represented by probabilistic analysis of inter-arrival time series, T, which is constituted by the time intervals between a rainy day and the immediately preceding one. On the hypothesis that Ts are independent and identically distributed random variables, a renewal process is generated; as it is known, a relevant property of renewal process is that it probabilistically re-starts at each arrival time. Analyses of observed T distributions have highlighted some peculiarities: seasonality, high variability, high frequency associate to T=1 and monotonically decreasing frequency with slowly decaying tail; in order to account for these properties, a probability mass function belonging to the Hurwitz-Lerch-Zeta (HLZ) family could be choose to statistical reproduce empirical observations. In this work, the three-parameter Lerch distribution is fitted to some empirical T distributions, derived from daily rainfall time-series recorded in two climatically-different environments: the Sicily Island and the region of Piedmont and Aosta valley.

Data

For the analysis, 55 daily time-series recorded in Sicily and 21 daily time-series recorded in Piedmont and Aosta Valley have been used. All the series refer to the period from 1975 to 2005. Sicily’s yearly rainfall regime is characterised by a strong seasonality: a “growing” (dry) season, GS, from April to September, and a “dormant” (wet) season, DS, from October to March, were assumed for the following analyses. During the GS, rainfall is usually lower than atmospheric demand, whereas the converse is true during the DS. Rainfall regime in Piedmont and Aosta Valley is virtually opposite to that Sicily; indeed, the “growing” season is generally wetter than “dormant” season. Rain gauges location and altitude are showed in the maps of Fig. 1.

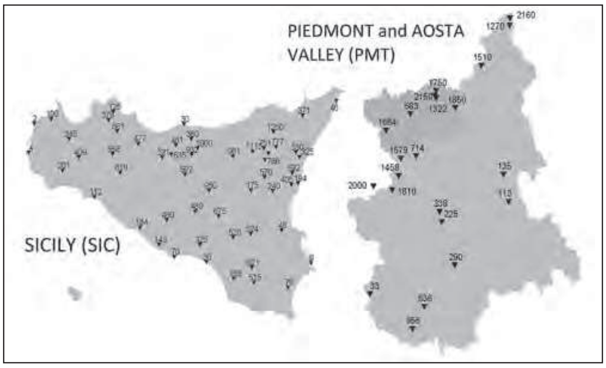

Fig. 1 - Location and altitude of rain gauges used in this study.

Procedures

The three-parameter Lerch probability distribution belong to the family of HLZ family, whose charac-teristics are described in detail by several researchers (Gupta et al., 2008; Aksenov and Savageau, 2005); accordingly, the probability that inter-arrival time is equal to k, results:

$$P[T=k]=\frac{w^k}{(a+k)^s\,\Phi(w,s,a)}$$

where w, s and a are the parameters of Lerch distribution and $\phi(,)$ is the Lerch function of argument (,):

$$\Phi(w,s,a)=\sum_{k=1}^{\infty}\frac{w^k}{(a+k)^s}$$

Firstly, each daily time-series was split in two sub-series, referring to the GS and the DS periods, respectively. Then, empirical T frequencies were fitted by three-parameter Lerch distribution, by using the maximum likelihood estimation (MLE).
Finally, goodness of fit was evaluated by calculating the p-value

Tab. 1 - Results of performed goodness-of-fit tests.

	# p-value > 0.05	% p-value > 0.05
SIC GS	47	85%
SIC DS	55	100%
PMT GS	19	90%
PMT DS	21	100%

Tab. 2 - Sample mean for some quantiles and performance indicators RMSE and ME.

	SIC GS			SIC DS		
quantiles	mean	RMSE	ME	mean	RMSE	ME
0.8	12.8	1.2	0.79	5.0	0.6	0.70
0.9	22.6	3.1	0.66	9.3	0.7	0.81
0.95	41.0	6.6	0.82	13.7	1.0	0.82
0.975	62.8	10.6	0.73	17.9	0.9	0.91
0.99	85.7	11.5	0.55	23.2	1.4	0.87
0.995	96.9	13.7	0.35	27.5	2.1	0.84
	PMT GS			PMT DS		
quantiles	mean	RMSE	ME	mean	RMSE	ME
0.8	5.9	0.8	0.74	7.0	0.7	0.85
0.9	9.3	0.6	0.92	12.6	0.7	0.94
0.95	13.0	0.7	0.93	18.9	0.8	0.97
0.975	16.3	0.6	0.97	26.0	0.9	0.98
0.99	21.6	1.2	0.90	36.8	2.9	0.91
0.995	26.2	2.6	0.82	44.0	3.3	0.92

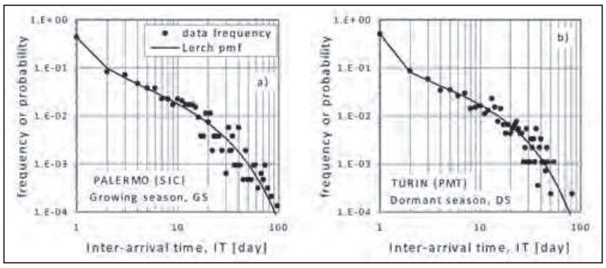

Fig. 2 - Two examples of fitting data with Lerch pmf.

associated with that chi-square statistic obtained by generating 2000 replicates via Monte Carlo procedure (Martinez-Rodriguez et al., 2011).

Results and discussion

In Table 1, test outcomes are synthetically presented, evidencing that excepting for few cases, a good agreement is obtained between the observed T frequencies and estimated probabilities. Figure 2 shows two examples of data fitting with Lerch distribution: the first (2a) refers to Palermo GS, the second (2b) to Turin DS. As can be observed, the fitting is excellent in both cases in the whole range of variability of T.

Figure 3a shows that estimated coefficients of variation, CV, are very close to the sample ones excepting for some GS inter-arrival series in Sicily; a similar pattern is observed in fig. 3b, where it is

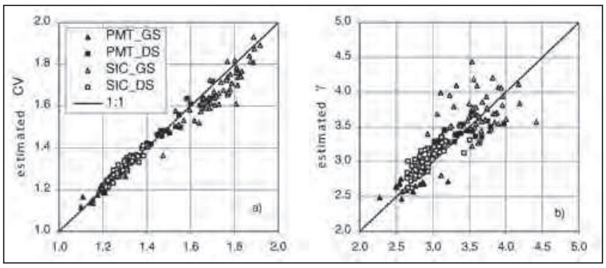

Fig. 3 - Comparison between a) sample and estimated coefficient of variation CV and b) sample and estimated skewness γ.

reported the comparison between sample and estimated skewness, γ. The analysis for the mean is not reported, because MLE fitting for Lerch distribution implies the exact equivalence between empirical and theoretical values.

Furthermore, Table 2 report the sample mean of some quantiles (0.8, 0.9, 0.95, 0.975, 0.99 and 0.995) as well as two performance indicators. The first indicator is the RMSE, defined as following:

$$RMSE(q) = \sqrt{\frac{\sum_{i=1}^{N} \left(q_{i,obs} - q_{i,est}\right)^2}{N}}$$

in which q is a quantile, observed (obs) or estimated (est) and N is the number of records. Values of RMSE (in days) are very lows with few exceptions for extreme quantiles of Sicily GS series. The second indicator is the ME (Modelling Efficiency), defined:

$$ME = 1 - \frac{\sum_{i=1}^{N} \left(q_{i,obs} - q_{i,est}\right)^2}{\sum_{i=1}^{N} \left(q_{i,obs} - \overline{q_{obs}}\right)^2}$$

where is the sample mean of any quantile. Values of ME are often in the range 0.7÷0.9, indicating a high performance of the fitted distribution.

Conclusions

The preliminary results here reported indicate the capability of 3-parameter Lerch distribution to describe observed frequencies of inter-arrival times in two different environments. That seems to suggest the possibility to accurately estimate the probability of the length of both rainy and dry periods with a parsimonious number of parameters.

References

Aksenov, S.V., and M.A. Savageau, Some properties of the Lerch family of discrete distributions, arXiv:math/0504485v1 [math.PR], April 2005.

Gupta, P.L., R.C. Gupta, Seng-Huat Ong, and H.M. Srivastava, A class of Hurwitz-Lerch Zeta distributions and their applications in reliability, *Appl. Math. Comput.*, 196, 521-531, 2008. doi:10.1016/j.amc.2007.06.012.

Martinez-Rodriguez, A.M., A.J. Saez-Castillo, and A. Conde-Sanchez, Modeling using an extended Yule distribution, *Comput. Stat Data Analys.*, 55(1), 863-873,2011. doi:10.1016/j.csda.2010.07.014.