FITTING AN EXPONENTIAL DISTRIBUTION: EFFECT OF DISCRETIZATION

Roberto Fraile[1], Covadonga Palencia[1], Amaya Castro[1], Dario Giaiotti[2], Fulvio Stel[2]

[1]*Universidad de León, Departamento de Física, 24071 León, Spain,* roberto.fraile@unileon.es,

[2]*OSMER-ARPA FVG Regional Met. Obs., via G. Oberdan, 18/A , I-33040 Visco, Italy ,* dario.giaiotti@osmer.fvg.it

(Dated: April 27, 2007)

I. INTRODUCTION

Raindrop sizes follow an exponential distribution (Marshall and Palmer, 1948) of the type f(x) = λ exp($-\lambda x$). Parameter λ has two interesting features: it is easy to calculate, and its meaning is easy to interpret. Hailstone sizes are also commonly considered to follow this type of distribution (Fraile et al., 1992, Giaiotti et al., 2001).

Previous studies (Fraile and García-Ortega, 2005) recommend the use of the moments method to calculate the value of parameter λ because other methods based on the linear fitting of the logarithm generate systematic errors. It is not possible to carry out a climatic analysis of the precipitation in two different regions by comparing the values of parameter λ if it has been calculated using different methods. The error introduced by the different calibration techniques (Palencia et al., 2007) would be enhanced by the error due to the calculation method.

In order to calculate λ using the moments method it is necessary to know the size of each drop or hailstone. However, this is not always possible since it may be the case that the database contains exclusively the sizes of the hydrometeors grouped in classes. The aim of this study is to reduce the error resulting when applying the moments method to the size classes.

II. PRESENTATION OF RESEARCH

The authors considered this problem when attempting to study the evolution of parameter λ in hailstones and finding that the data storage method had changed from an old system (in groups every one or two millimeters) to the computerized storage of each individual size.

How can the moments method be applied to this type of measurements to calculate the parameter lambda of the exponential distribution? Let's start from a size classification like the one shown in Table I.

Lower limit	Upper limit	Nº of hydrometeors
d_0	d_1	N_1
d_1	d_2	N_2
d_2	d_3	N_3
...	...	...
d_{i-1}	d_i	N_i
d_i	d_{i+1}	N_{i+1}
...	...	...
d_{n-2}	d_{n-1}	N_{n-1}
d_{n-1}	d_n	N_n

TABLE I: Classes of hydrometeors.

Then, since the moments method provides a value for λ which is the inverse of the mean value of all the data, the most straightforward option would probably be the following (by analogy):

$$\lambda = \frac{\sum_{i=1}^{n} N_i}{\sum_{i=1}^{n} d_i' N_i}$$

where d_i' is the representative value for class i (from d_{i-1} to d_i).

The problem is the selection of d_i'. If we opt for identifying it with $(d_{i-1} + d_i)/2$, the value in the middle between d_{i-1} and d_i, a certain error is introduced, since the objects (drops or hailstones) also follow an exponential distribution in that class, and the mean value of all the objects in that class is NOT the central value between d_{i-1} and d_i. The following error is introduced:

$$\varepsilon = \frac{\delta}{2}\frac{1+\exp(-\lambda\delta)}{1-\exp(-\lambda\delta)} - \frac{1}{\lambda}$$

where the size of the class is $\delta = d_i - d_{i-1}$.

It can be seen that the error is directly related to the size δ of the class: the narrower the class, the smaller the error.

If we have data about the sizes of hydrometeors grouped in classes of size δ, and if we use as the representative value of the class the middle value to calculate parameter λ' of the exponential distribution, it is easy to determine parameter λ with the mean value of the sizes of each class:

$$\lambda = \lambda' \frac{1}{1-\varepsilon\lambda'} = \lambda' + \frac{\varepsilon\lambda'^2}{1-\varepsilon\lambda'}$$

It can be seen that since $\varepsilon > 0 < \lambda$', it must always be true that $\lambda > \lambda$'.

III. RESULTS AND CONCLUSIONS

In order to test the validity of the equation above 500 hailpads have been collected from the network described in Giaiotti and Stel (2006). A representative sample was selected, discarding those hailpads with less than 100 impacts or with impacts of sizes under 7 mm.

The value of λ_o was calculated using the moments method because the size of each hailstone is known. Next, the sizes were classified in intervals of one millimeter and the value of λ' was calculated taking as the representative value of the class the middle point. Fig. 1 shows the relative errors committed in this calculation according to the number of hailstones registered per square meter.

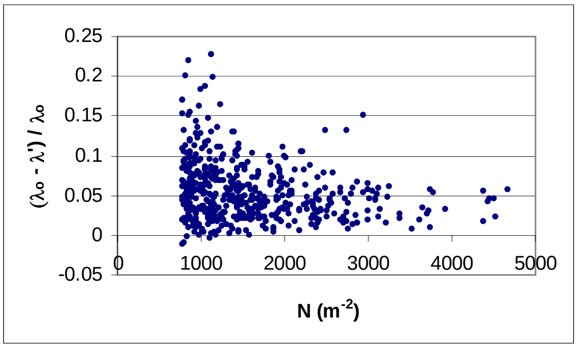

FIG. 1: Relative errors committed calculating λ'.

Fig. 1 illustrates the fact that in the case of small numbers of hailstones the error increases: with less than 2,000 hailstones per square meter the error frequently exceeds 5%. If instead of representing N on the axis of abscissas we had represented λ_o, the relative error could be seen to increase linearly with λ_o exceeding 5% whenever λ_o is over 1 mm^{-1}. For values of more than 1.5 mm^{-1}, the relative error always exceeds 15%.

The λ values have been calculated for the same hailpads used before. When the λ value was calculated with all the hailstone sizes the error is considerably reduced. Fig. 2 represents the two types of error in order to compare the relative errors of both methods of calculation.

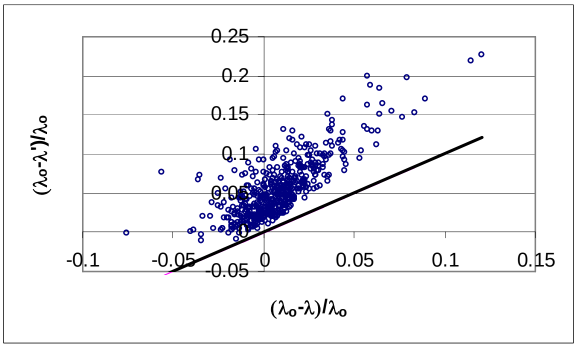

FIG. 2: Relationship between the relative errors calculating λ' and λ. The line $y = x$ is also represented.

Fig. 2 shows a number of interesting aspects:

- There is a relationship of proportionality between the two relative errors.
- The error in calculating λ exceeds 5% in only 4% of the hailpads.
- The relative error of λ is always smaller than in λ' (in Fig. 2 all the points lie over the line y=x)

The error in λ tends toward zero if the sizes of the hailstones fit perfectly an exponential distribution.

The whole process just described was repeated with the classes of 2 and 3 mm, and it was observed that the error in calculating λ' increased with the width of the class. This represents an experimental verification of the calculation of ε made in this study.

The results of the study imply that the correction of the value λ suggested in this study is needed if we want to minimize the error in this parameter. If we intend to determine λ from data distributed in classes, we may first calculate λ' and apply the correction to determine λ and compare the value obtained with other values of λ obtained by the moments method using all the data.

IV. ACKNOWLEDGMENTS

The authors are grateful to Dr. Noelia Ramón for translating the paper into English. We would like to thank all the volunteers who, since 1988, contribute to the Italian hailpad network maintenance. We are also in debt with Elena Gianesini for her work on hailpad data reduction.

V. REFERENCES

Fraile R., Castro A., Sánchez J. L., 1992: Analysis of hailstone size distributions from a hailpad network. *Atmos. Res.*, 28 311-326.

Fraile R., García-Ortega E., 2005: Fitting an exponential distribution. *J. Appl. Meteorol.*, 44 1620-1625.

Giaiotti D., Gianesini E., Stel F., 2001: Heuristic considerations pertaining to hailstone size distributions in the plain of Friuli-Venezia Giulia. *Atmos. Res.*, 57 269-288.

Giaiotti D., Stel F., 2006: The effects of environmental water vapor on hailstone size distributions. *Atmos. Res.*, 82 455-462.

Marshall J. S., Palmer W., 1948: The distribution of raindrops with size. *J. Meteorol.*, 5 185-166.

Palencia C., Berthet C., Massot M., Castro A., Dessens J., Fraile R., 2007: On the individual calibration of hailpads. *Atmos. Res.*, 83 493-504.