June 12, 1998.

Demetri,

I am sorry to say that this is one of the cases that I got myself and you in trouble by suggesting prematurely and without a careful reading the submission of this paper and especially to HESS. As you can see both reviewers do not recommend publication. I read it myself more carefully and I also feel that unless extensive additional work is done (that would be a different paper all together) I do not feel comfortable suggesting its publication. I have marked my comments on the manuscript (sometimes on a critical tone intentionally). I am surprised I saw it in a different light when you asked me for a recommendation. It was probably for lack of proper time to carefully consider it and also because I considered it part of the package with the other paper (which is original). The material from the other paper is simply used here and do not even count as original; the proper motivation and context of it is not even properly stressed in my opinion.

In any case, the major issues I have with the paper are:

1) originality
2) One single record which someone can even question its validity
3) Small-sample results must be verified against controlled experiments to isolate the effects of

small sample from those of inhomogeneities etc.

I feel very bad to have mislead you on this and I do not know what to say because I should have been more critical when I first looked at it° for a suggestion.

I have recommended rejection to the editor. I have to think (it's not absolutely obvious to me now) how the paper can be reworked for further submission. We can talk about this in the summer.

Well, after all anonymity helps (less worry!) but I trust that you understand my decision anyway.

All the best,

Efi

Review of the paper: "Effect of Sample Size on Design Rainfall Inferences"
by D. Koutsoyiannis and G. Baloutsos

The paper provides an extensive analysis of annual maxima of daily precipitation in Athens. Due to the presence of an unusually long data series it is found justified to use a three- parameter description, and the GEV distribution is found appropriate. A comparison to probable maximum precipitation is made and intensity-duration-frequency curves are derived using the GEV assumption. The paper is well written and well organised. There remains, however, a major problem in that no original research is presented. All analyses appear to be applications of well-known methods which, although being relatively new, may be considered more or less standard today. Hence publication is not recommended, unless the authors can expand their analysis and hereby contribute original findings.

Review Notes, Koutsoyiannis & Baloutsos; Effect of Sample Size on Design Rainfall Inferences

Preliminary comment: The draft which I reviewed will require substantial editing for English usage if it is to appear in an English-language journal.

Specific comments and points requiring clarification:

Abstract, p.1: Reference is made to "remarkable statistical properties" of the 136-year time series, although this statement is not adequately supported by the subsequent analysis.

Section 2, p. 5: I would like to see more information on the nature and results of the tests for time series homogeneity cited by the authors. Given the nature and objectives of this study, the possibility of non-homogeneity of data between the early (1860-1889) and later periods due to changes in rain gauge location is an issue of great potential significance. See, e.g., Potter (1981), *Monthly Weather Review* 109(9) for applications of homogeneity tests to precipitation gauge data.

Section 3, p.6: I don't recognize the functional form presented for the EV-II distribution as appearing in the text (equation 3), and was not able to reconcile it in a straightforward way with Johnson & Kotz (1970), although possibly for reasons relating to differing parameter conventions. The Handbook of Hydrology, Chapter 18 (1993) also employs a different sign convention on k in the GEV (equation 4).

Section 3, pp.7-8: The addition of the L-moment diagram will serve to clarify this discussion.

Section 3, p.8: Lower-tail fit may have statistical, but little hydrometeorologic significance.

Section 3, p. 9: The observation that the GEV provides a fit superior to either of the 2-parameter distributions is fully anticipated, since the GEV contains an extra parameter. A longer time series will, naturally, embody elements of an increasing variety of events generated by heterogeneous mechanisms, thus requiring additional parameters for successful fit. A useful addition to this portion of the analysis would be a prior evaluation of the *annual* distribution of daily precipitation totals, i.e., the initial distribution from which annual maximum values are extracted, as each extreme value distribution is associated with specific types of initial distribution.

Section 4: The discussion of the differing (and misleading) behavior of subsets relative to the longer data set is potentially the most useful aspect of this study, but as it stands, it is not clear whether the authors are describing (a) a location-specific anomaly, (b) an artifact of mixed

distributions and/or climatic nonstationarity or (c) a purely statistical phenomena. The study could be strengthened by some additional analysis of the possible causes of this phenomena, and the insight acquired could suggest the scope of applicability of the authors' observations to other locations or to other data series. Possibility (b) could be investigated easily by examining the dates of occurrence of each year's annual maxima for homogeneity. A more detailed examination, perhaps beyond the scope of this study, might focus on the atmospheric circulation pattern prevailing during the event. Possibility (c) is easily investigated via monte carlo simulation: do subsets of a long record, generated to conform to GEV distribution, in general exhibit behavior similar to the observed records? There is an existing literature on this, e.g., Wallis, Matalas and Slack 1974 (WRR 10(2): 211-219), which the authors should perhaps have made reference to.

Section 5, p.10: What is meant by "(Besides), the distribution function obtained by the complete 136-year series, apparently, is not the true population distribution."? The authors elsewhere assert that "... the (above) analyses provide evidence that the GEV distribution is a consistent probabilistic model for the annual maximum series of the daily rainfall depth ..." (p. 8). This should be clarified.

General Recommendations:

The authors discuss an issue of great significance and ongoing research interest in hydrology -- the relationship between small-sample distribution moment estimates (and corresponding quantile estimates) and those of the parent distribution. A substantial literature exists on this topic, and it seems to me that the authors' contribution to this literature might come from an exploration of the *causes* (e.g., climatic mechanisms) of the observed phenomena. I would therefore recommend to the editors of HESS that the article not be published in its present form. I believe this research would better serve the objectives of an international journal, and its authors, if it were re-submitted following a more thorough investigation of the hydrometeorology generating the observed statistics, and a more thorough review of the existing literature on small-sample properties.