December 15, 2014

To: Jesus Carrera
Editor
Hydrology and Earth System Sciences

Re: Revised Manuscript "Hydrol. Earth Syst. Sci. Discuss., 11, C5366–C5371, 2014" -
Authors' replies to Reviewer No. 2

Extreme Value Statistics of Scalable Data Exemplified by Neutron Porosities in Deep Boreholes

by A. Guadagnini, S.P. Neuman, T. Nan, M. Riva, and C.L. Winter

Dear Editor:

We appreciate the efforts you and the Reviewers have invested in our manuscript. Following is an itemized list of the comments of Reviewer No. 2 together with our response to these. Comments are reported in blue and our responses in black font.

Sincerely,

Alberto Guadagnini, Shlomo P. Neuman, Tongchao Nan, Monica Riva, and C. Larrabee Winter

<u>Comments by Reviewer No. 2</u>

The paper presents a scaling analysis aimed at investigating the behavior of extensive sets of neutron porosity data collected in the field. Section 2 describes the dataset, consisting of data collected at six different wells located within three different geological environments, and presents their frequency distribution. Section 3 analyzes the frequency distribution of data increments as a function of lag distance; these are seen to follow Levy stable or normal-lognormal distributions at smaller lags, while becoming Gaussian at larger lags. Section 4 discusses the scaling behavior of sample structure functions of generalized order, showing a dual behavior of the scaling exponent, that is distinctly larger for smaller lags than for larger lags. Corresponding Hurst coefficients, estimated via the method of moments, are of order 0.7 and 0.1 respectively, showing persistence is associated with intra-layer variability, anti-persistence with inter-layer variability. The relationship between structure functions of different order satisfies Extended Self Similarity (ESS) and provides support to the unified theoretical framework, proposed by the authors in previous papers, which views data to be consistent with sub-Gaussian random fields subordinated to tfBm/tfGn (truncated fractional Brownian motion/ truncated fractional Gaussian noise). Section 5 presents the general scaling theory and the derivation of variogram parameters for the specific data set. Further elements of the theoretical framework are reported in Appendices A and B. Sections 6 and 7 of the paper deal with statistics of peaks over threshold (POTs), defined as such when exceeding the 95% quantile. Their frequency distribution follow a generalized Pareto distribution. The structure functions of their increments exhibit behavior similar to unfiltered field.
The paper is fully within the scope of HESS, and of interest to its readership. My concern is that the title reflects only partially the contents of the paper. In fact, the application of the methodology to a large dataset constitutes an important contribution in itself. The extension

The Reviewer has summarized with clarity the key points of the manuscript and we thank him/her for his/her very positive appreciation of our work. Also following the comments of Reviewer No. 1, we will modify the tile of the manuscript to de-emphasize the aspects associated purely with extreme values and make it in line with the overall content of the work. It now reads: "Scalable statistics of correlated random data and extremes applied to deep borehole porosities."

We can provide a qualitative comparison amongst all cases we have examined and/or are presented in the literature, in the sense that this pattern is common to a variety of data we have analyzed. We are not convinced that a quantitative comparison is appropriate as the specific results, in terms of parameter variability with lag can be data dependent. We will include appropriate reference to support the qualitative comparison amongst a variety of available cases presented in the context of the current literature. These are not limited to Earth and environmental sciences and include applications to financial, biological and other types of data.

Well 6 is associated with the Tabnak formation, which is the richest in terms of carbonate content amongst the three types of depositional environments we consider. Heterogeneity of carbonate rocks can be stronger that that displayed by sandstone-based rocks and we clarify this point in the revised manuscript.

We think the Reviewer is referring to the lower, not upper, cutoff scale, i.e., λ_l. Theoretically, the value of λ_l should be a fraction of the measurement scale. In our case, the measurement scale is smaller than the 0.15 m lag, which represents data resolution (in Well 6 data resolution is 0.07 m). When compared to the overall length scale spanned by each borehole (which is of the order of 10^3 m), λ_l can therefore be considered as negligible. As we noted in our previous work, which we reference in the manuscript), this assumption also enables us to eliminate instabilities in the Maximum Likelihood parameter estimation associated with small parameter values and does not influence the generality of the methodological approach. We add this clarification in the revised manuscript.

4. The structure functions associated with POTs behave similarly to the unfiltered field. An interesting exception are the (extremely low, almost zero) values of the Hurst coefficient associated with (large lag) intra-layer variability. Could the authors provide a physical interpretation of this effect?

As we state in our responses to Reviewer No. 1, our estimates of H at large lags are low, indicating antipersistence of data correlated over long distances; this is typical of alternating layers formed by generally diverse geomaterials. Our interpretation is consistent with depositional events similar to those invoked by Dashtian et al. (2011), which we reference in the manuscript.

We are not convinced that there is a clear and definite interpretation for the very low estimated Hurst coefficients in the case of the POTs. These findings might be interpreted as indicative of a degree of antipersistence (tendency of low and high values to alternate rapidly, in a rough rather than a smooth manner, across layers) between POTs of a property such as neutron porosity in two diverse layers which is even higher than that associated with unfiltered values of the property. We feel that at this stage of the research this interpretation is mostly speculative and would prefer not to elaborate further on this aspect.

5. Generalized Pareto distributions seem to describe the behavior of POTs. A short appendix describing their behavior and associated parameters would help in following the last section of the paper about POTs.

We will prepare such an Appendix in the revised manuscript. This will also accommodate a corresponding request from Reviewer No. 1.

6. I suggest to review the presentation of the background material (underlying theoretical framework), that is now split between Section 5 and Appendices A and B.

We are prepared to do so in our revised manuscript. This will also accommodate a corresponding request from Reviewer No. 1.