

Interactive comment on “Data compression to define information content of hydrological time series” by S. V. Weijs et al.

Anonymous Referee #1

Received and published: 12 April 2013

In this paper the authors try to introduce the AIT to hydrologists, and to present the results of application of same compression algorithms to hydrological data. They try to explain how these algorithms could reflect the information content of these data.

Although the bibliographical review is needed to understand the paper because the AIT could be new for the readers of the journal (hydrologists), it seems to be too long. Some paragraphs of this part could be omitted for brevity.

The authors don't show clearly the usefulness of the results obtained in hydrology, especially what is the supply of AIT compared to the data mining techniques used in hydrology.

The paper is interesting, the methodology is clearly described; however, the most im-

C766

portant results are left as a question for future researches. It seems also that the authors let the important results (application with hydrological model) to a future paper. It would be better that the two papers (present and future paper) were published in the same journal as part 1 and part 2.

Quantization: the authors have chosen a uniform quantization with a precision equal to 8 bits. The question is why they have chosen 8 bits and how they have done this choice? This choice could affect the results (by inducing a loss of precision). It's possible that the quantization smooths the data with large range (the case of rainfall data for humid region). Is it better to choose precision proportional to the range of data for example? I recommend to do other experiments with different precision and compare results. It's possible also to use a non-uniform quantization (μ -law quantization for example) that could give less loss of information.

In experiment B, why you use the same quantization schema of Q for Q_{er}, Why you don't use the limits (min and max) of Q_{er} for quantization.

Section 3.2 : I recommend to authors to give more references about the compression algorithms used and the references of their source code.

Section 3.3 It seems better to explicit the different method used to generate the data. Readers could not understand the meaning of sine1, sin100...etc in tab 2.

Section 5 : discussion and conclusion need to be more related to the experiment done by the authors.

in page 2033 line 17 please say that the DKL is the KullBack-Leibler divergence and give a reference

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 2029, 2013.