

## Response of authors to Referee # 1 comments

**The authors are grateful to the reviewer for her/his comments, which contributed to improve the quality of the paper.**

**Referee comment:** *The proposed method is well analysed and the paper is well written. I support the publication of this paper in HESS. I have a couple of comments that should be addressed before publication, but since they mostly involve additional discussion, the resulting revision should be minor.*

*Page 4670, line 12-13: this sentence seems to be inconsistent with line 20-23 at page 4669 where it is said that "there are two main deficiencies in using the OLS as a record extension technique for water quality data. First, it is not robust for the presence of outliers. Presence of outliers significantly affects intercept and slope estimates in the OLS..." and in the abstract "On the other hand, the Kendall-Theil robust line (KTRL) method has been proposed as an analogue of OLS with the advantage of being robust in the presence of outliers."*

**Authors' response:** In the revised version, the sentence in page 4670, lines 12-13 is modified as follows: "Their results showed that Kendall-Theil robust line (KTRL) is a very strong analogue to OLS regression with the advantage of being robust in the presence of outliers." Reference is made to Navitt and Tam's (1998) work, where they concluded that the KTRL is an analogue to OLS, while in the presence of outliers, unlike the OLS, the KTRL can provide accurate estimates of the population parameters.

**Referee comment:** *Page 4671, second last line: "period from  $n_1 + 1$  to  $n_1 + n_2$  years", not "period from  $n_1 + 1$  to  $n_2$  years".*

**Authors' response:** In the revised version of the paper, " $n_1+1$  to  $n_2$ " is replaced by " $n_1+1$  to  $n_1+n_2$ ".

**Referee comment:** *Page 4672: suggest to insert the assumptions of the OLS method.*

**Authors' response:** In the revised version of the paper, the OLS assumptions are included in page 4672, after line 19 as follows: "It should be emphasized that OLS has five assumptions (Helsel and Hirsch, 2002):  $y$  and  $x$  are linearly dependent; the data used to fit the model are representative; the variance of the residuals is constant; and the residuals are independent and normally distributed."

**Referee comment:** *Page 4672, equation (1): "for  $i = 1, \dots, n_1$ ", not "for  $i = n_1+1, \dots, n_2$ ".*

**Authors' response:** In the revised version " $i = n_1+1, \dots, n_2$ " is replaced by " $i = 1, \dots, n_1$ ".

**Referee comment:** *Page 4677, line 9: how is the sample sizes?*

**Authors' response:** The sample size was 100. This will be indicated in the revised version as follows: "Different mixture distributions (each of 100 samples) were generated containing between 100 and 80 percent of the main distribution and between 0 and 20 percent of the second distribution."

**Referee comment:** *Page 4682, second last line: "moments", not "momnets".*

**Authors' response:** In the revised version "momnets" is replaced by "moments".

**Referee comment:** *Subsection 4: I would suggest to shorten this subsection.*

**Authors' response:** The authors have decided not to shorten this section since we would like to be sure that all necessary information and discussions are provided.

**Referee comment:** *Figure 1: Why are the  $RMSE_{loc}/RMSE_{ols}$  and  $RMSE_{loc}/RMSE_{ctrl}$  not also shown?*

**Authors' response:** The  $RMSE_{RLOC}/RMSE_{LOC}$  was the only comparison considered because the RLOC is proposed as a robust version of the LOC technique.

The  $RMSE_{rloc}/RMSE_{ols}$  and  $RMSE_{rloc}/RMSE_{ktrl}$  comparisons were not considered because the RLOC is a technique that is able to maintain the variance in the extended records, while OLS or KTRL are not. In addition, for  $RMSE_{rloc}/RMSE_{ktrl}$ , both techniques are robust, and the comparison considered was to show the effect of the presence of outliers on the slope estimates. These are the reasons for considering only the comparison between the original technique LOC and its modified version RLOC.

**Referee comment:** *explain better the y-label, for instance "RMSEratio = RMSErloc=RMSEloc".  
Caption: "Relative efficiency of ... LOC slope estimator. The population is composed by a mixture of two normal distributions ( $N(10; 1)$  and  $N(11; 3)$ ); the x axis tracks the percentage of the second distribution in the population."*

**Authors' response:** In the revised version, Figure 1 caption is modified as follows:

“Figure 1. Relative efficiency of the RLOC slope estimator as compared with the LOC slope estimator; the population is composed of a mixture of two normal distributions ( $N(10,1)$  and  $N(11,3)$ ); the y-axis represents the RMSE ratio =  $RMSE_{RLOC} / RMSE_{LOC}$ ; the x-axis tracks the percentage of the second distribution in the population.”

**Referee comments:** *Figure 2: (this is just a suggestion) Caption ".Electric Conductivity (EC) and Chloride (Cl)..". What the numbers in the box-plot means? Why there are circles and stars? Maybe they can be remove.*

**Authors' response:** In the revised version, the numbers in the box-plots are removed, and outliers are represented by stars only.

Figure 2 caption is modified as follows: “Figure 2. Edko drain sampling locations and box-plots for Electric Conductivity (EC) and Chloride (Cl) records at three locations.”

**Referee comment:** *Figure 3, 4, 6 and 7: would it be useful to use mixed "points+lines" in the figures.*

**Authors' response:** We tried the “points + lines” (dash-dot) in the figures, but because the two lines (either those representing the LOC and RLOC, or those representing the OLS and KTRL) almost coincided, we found that keeping one line as dashed and the other solid was better than using “points + lines”.

**Referee comment:** *Figure 3-4:(this is just a suggestion) The figures could be merged into one figure showing BIAS (e.g. only for  $\rho = 0.5$  and  $\rho = 0.9$ ) and RMSE (e.g. only for  $\rho = 0.5$  and  $\rho = 0.9$ )*

**Authors' response:** In the Monte Carlo study, three correlation coefficients were considered along with four different  $n_1$  and  $n_2$  combinations. These create 12 different combinations, from which half of them were presented to show the correlation coefficient and sample size effects. Presenting only the four extreme combinations may also show the effects of both the correlation coefficient and sample size. However, we found that if four BIAS sub-figures and four RMSE sub-figures are combined into one figure (containing eight sub-figures), each sub-figure may be too small to be clear to the reader. Thus, we are suggesting that we keep Figures 3 and 4 unchanged in the revised version.

**Referee comment:** *Figure 6-7: (this is just a suggestion) The figures could be merged into one figure.*

**Authors' response:** Figures 6 and 7 will be merged into one figure in the revised version and related figure numbering within the text (Page 4686, line 24 to Page 4689, line 1) will be adjusted. Figures 6 and 7 become Figure 8 in the revised version, because of inserting two new figures based on the comments of Referee # 2 (Dr. Serinaldi). Figure 8 caption becomes: “Figure 8. BIAS and RMSE of the tested extension techniques in estimating Chloride (Cl) percentiles”

**Referee comment:** *Figure 5, caption: (just a suggestion) "Box plots of the Inter-quartile Range (IQR) ratio, RLOC technique."*

**Authors' response:** In the revised version, this comment is considered and Figure 5 caption becomes: "Figure 5. Box plots of the Inter-quartile Range (IQR) ratio, RLOC technique."