# Reply to Anonymous Referee #1

September 11, 2015

**Referee's comment 1.** The objective is stated several places in the manuscript (p.3544,l.3; p.3547,l.1; p.3551,l.4; p.3555,l.10). It is not entirely clear why you chose to evaluate the MEWP model by comparing it to MGPWP. Also, on p.3551,l.4 you mention that the regional scale is of main interest, but this is also not clear in the introduction.

**Authors' reply:** The difference between MEWP and MGPWP lies in the assumption about the seasonal weather-pattern distribution: MEWP uses the exponential distribution (EXP) whereas MGPWP uses the GPD distribution. The reason for using the GPD is to allow distributions with heavier tails than that of the exponential distribution which is light tailed. There exists other distributions with heavy tails but 1) the exponential distribution is a particular case of the GPD (when the shape parameter $\xi \to 0$) and therefore by using MGPWP we just generalize MEWP; 2) the use of GPD distribution is justified by Extreme Value Theory. We will make this clearer in a revised version of the manuscript.

**Referee's comment 2.** p.3546,l.25: What does this inhomogeneous distribution of stations mean for the analysis?

**Authors' reply:** The fact that the network is denser in southern Norway, particularly along the coast, implies that this region weights more in the model evaluation (e.g. in the scores) than the very north of Norway. Since this is also the region with the highest population density, this means that more weight is given to highly populated regions in the model evaluation, which probably makes sense from a practical/political point of view.

**Referee's comment 3.** 3. p.355,l.20: Please state why you use L-moments, and not other estimation methods.

**Authors' reply:** First it should be noted that the choice of the L-moment method only affects the GPD case since for the EXP case, L-moments, moments and maximum likelihood estimators coincide. For the GPD case, we actually also tried the maximum likelihood (ML) method and it turns out that using L-moments or ML barely affects the global scores. The reason is that we are here in the case of a regional evaluation so slight differences in estimation that occur at local scale are smoothed out at the regional scale.

**Referee's comment 4.** p.3555,section 4.1: Please consider whether this section should be moved to chapter 3, as it does not present any actual results (at least not the first part). You could also split it.

**Authors' reply:** That's true, this subsection doesn't present any results. However section 3 is a general section explaining the theory, independently of the data. Section 4 relates to the application of this theory to our data and in particular section 4.1 explains which models are applied to the data. Therefore we think it is better to leave the structure of the paper as it is. However to make it clearer we propose to replace the section name "Results" by "Application of MEWP and MGPWP in Norway" and create a new subsection 4.2 named "Results" and containing sections 4.2, 4.3 and 4.4, which will be respectively numbered 4.2.1, 4.2.2 and 4.2.3.

**Referee's comment 5.** 5. p.3559,l.4 and l.19; Fig.5: which eight models? on p.3556 you describe 12 models, what happened to the other 4 models? In Fig.5 only eight models are shown, but on p.3559,l.19 you refer to models that are not shown (case (1,4) and (1,8)). Please refer to rows in Fig.5, as you did on l.17.

**Authors' reply:** We apologize for the confusion. There is indeed a mistake here. There are 12 models, not 8. Cases $(S, K) = (1, 4)$ and $(1, 8)$ are missing in Figure 5. That figure should be replaced by Figures 1 and 2 below. We will improve the text as suggested.

**Referee's comment 6.** p.3559,l.18: do you here mean (2,1) instead of (1,2) (ref.Fig.5)?

**Authors' reply:** We apologize for the confusion. Indeed, l.18 should be "there is a clear benefit obtained from the use of seasonal splitting (case $(S, K) = (1, 1)$ vs. $(2, 1)$) ..."

**Referee's comment 7.** p.3559,l.14-16: do you have a reference to damn safety regulations in Norway?

**Authors' reply:** A good reference is [Midttømme et al., 2011]. It will be added in next version of the article.

**Referee's comment 8.** p.3560,l.11-12: do you have a reference for the claim that the shape parameter is difficult to estimate?

**Authors' reply:** This is actually a classical result of extreme value theory. We can for example refer to page 528 of [Garavaglia et al., 2011] or to the upper right of page 350 of [Serinaldi and Kilsby, 2014].

**Referee's comment 9.** p.3563,l.15-20: I suggest to move this paragraph to section 4.4 and only briefly summarize it here in Conclusions.

**Authors' reply:** OK, we agree with that.

**Referee's comment 10.** p.3563,l.25: I believe the GPD can also be light-tailed (shape parameter $< 0$) (given that you use the definition of heavy-tailed to be distributions whose tails are not exponentially bounded). If I am not mistaken, this should also be corrected on p.3544,l.8. Perhaps the meaning of

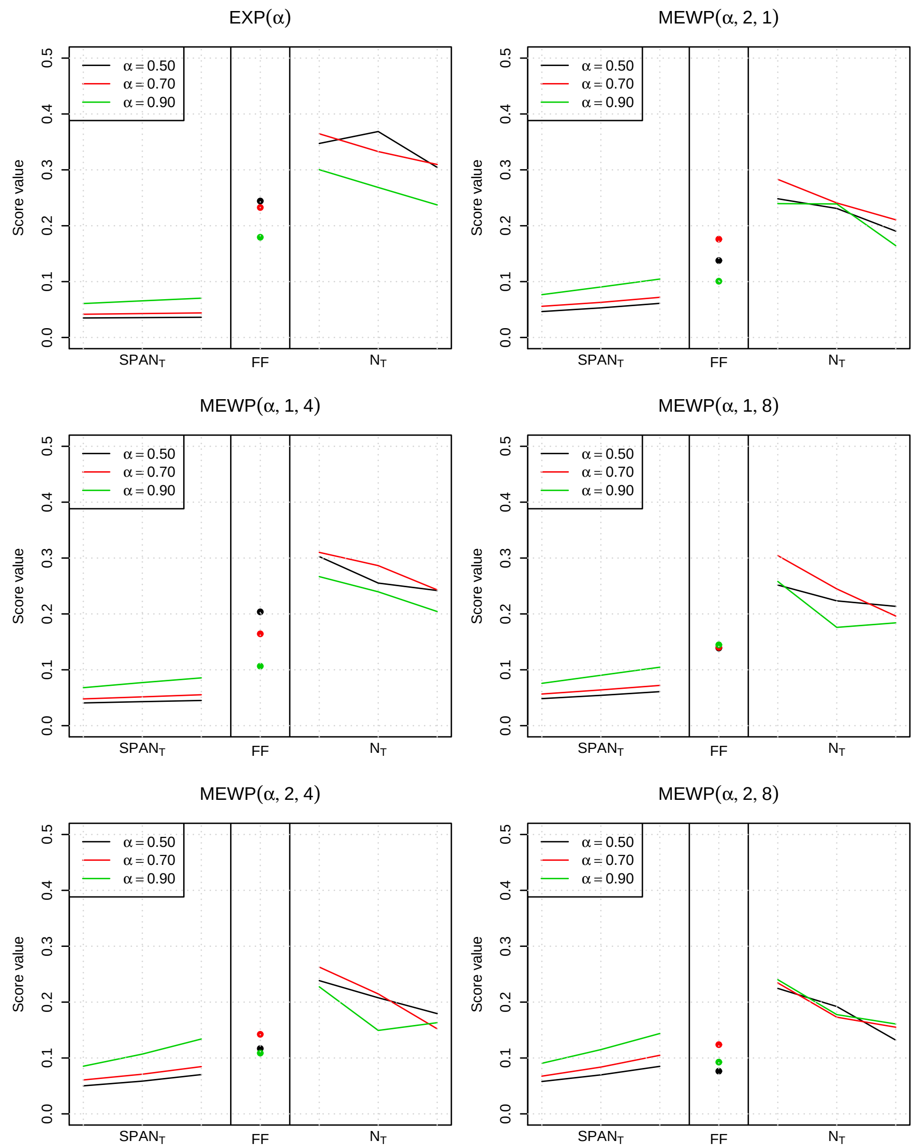


Figure 1: Scores of evaluation for MEWP models, for $\alpha = 0.5$, 0.7 and 0.9. Better scores have values closer to 0. Scores of $\mathrm{SPAN}_T$, for $T = 20, 100, 1,000$ year return periods, are the mean scores of (8), while scores of $FF$ and $N_T$, $T = 5, 10, 20$ years, are based on the density areas (9).

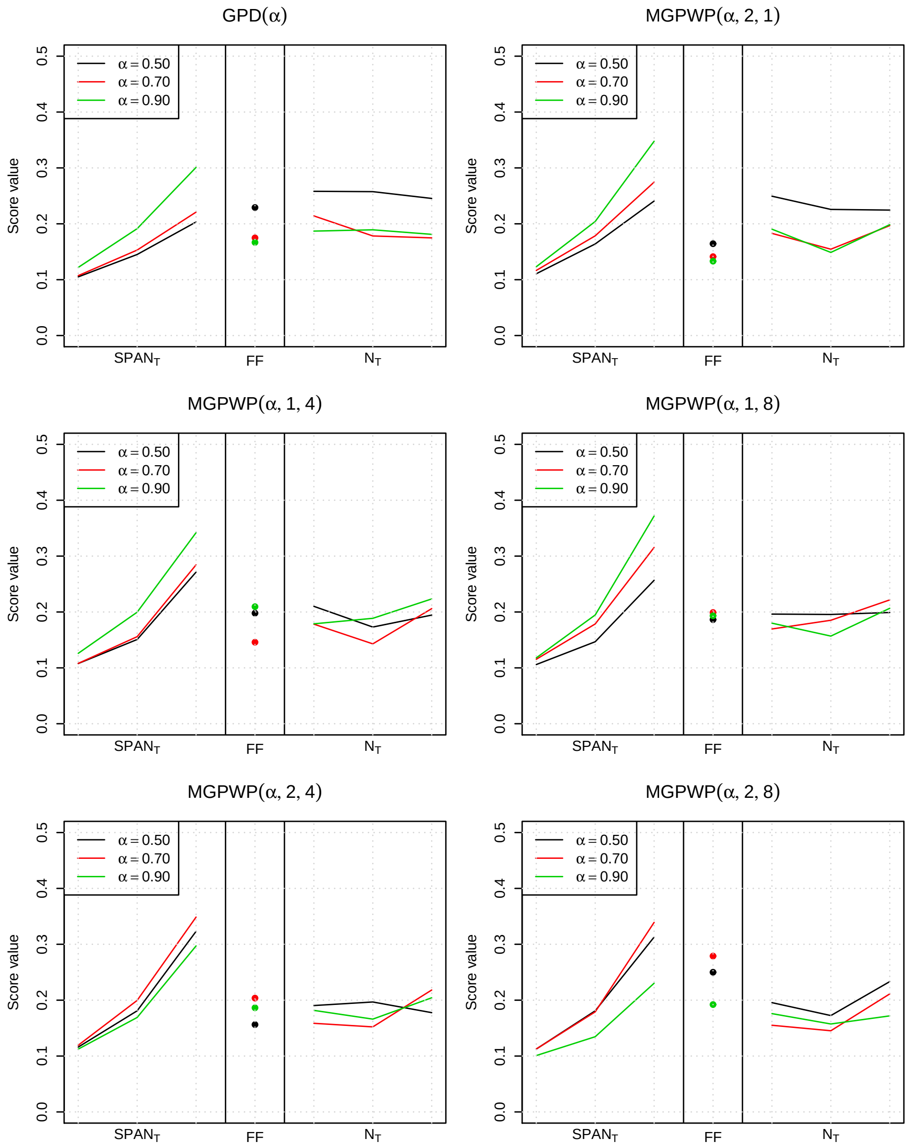


Figure 2: Same as Fig. 1 for MGPWP models.

the shape parameter on return levels should be clarified on p.3548.

**Authors' reply:** That's true, the GPD is only heavy-tailed only when the shape parameter is $> 0$. We will correct the article as suggested. Regarding the computation of return levels, to make it clearer we propose to add a section between 3.2 and 3.3 entitled "Computation of return levels" with the following text:

The T-year return level $r_T$ is the level expected to be exceeded on average once every $T$ years. It satisfies the relationship $F(r_T) = 1 - 1/(T\zeta)$ where $\zeta$ is the mean number of central rainfall events per year. When $F$ is EXP($\alpha$) or GPD($\alpha$), estimation of $r_T$ is obtained explicitly as

$$\hat{r}_T = \begin{cases} q_\alpha + \hat{\sigma}_\alpha \log\{(1-\alpha)T\zeta\} & \text{for EXP}(\alpha) \\ q_\alpha + \hat{\sigma}_\alpha\{[(1-\alpha)T\zeta]^{\hat{\xi}} - 1\}/\hat{\xi} & \text{for GPD}(\alpha) \end{cases} \quad (1)$$

where $\hat{\sigma}_\alpha$ and $\hat{\xi}$ are the parameters estimates of $F$ of Section 3.2. For the MEWP and MGPWP models, there is not an explicit formulation for $\hat{r}_T$ and it is obtained numerically by solving $F(\hat{r}_T) = 1 - 1/(T\zeta)$ in (7). Eq. 1 shows that in GPD($\alpha$) model, $\hat{r}_T$ is mainly influenced by the value of $\hat{\xi}$. For the MGPWP model, practice shows that for reasonable to large $T$ (typically $T > 50$ years), $\hat{r}_T$ is mainly influenced by the largest $\hat{\xi}_{s,k}$.

**Referee's comment 11.** You conclude that the best model is the most complex MEWP model (2 seasons, 8 WPs); what does this imply for the computation? Are there any drawbacks?

**Authors' reply:** Computation is not really an issue. Calculations are very easy (see Section 3.2). Estimating MEWP(2,8) takes less that 1/100 second for each station. The drawback of MEWP is that a WP classification is needed for its application. This is not the case in all countries worldwide although such a classification has been developed for some areas (France, Norway, West Canada and Austria; see [Brigode et al., 2014]). Once this classification has been built, estimating MEWP is almost as simple and fast as estimating an EXP.

**Referee's comment 12.** p.3577,Fig.11: Increase figure size, as the points are hard to see (especially on the right map).

**Authors' reply:** We made the points bigger, see Figure 3.

**Referee's comment 13.** Fig.1, Fig.4, and Fig.11: Which stations is outside the west coast of Sweden in the south? I can not find this station in other figures of Norwegian weather stations

**Authors' reply:** We are sorry for the confusion. This station is located in Rygge, which is about 50km south of Oslo. We misplaced this station in Figures 1, 4 and 11 of the article. The correct maps are shown respectively in Figures 4, 5 and 3 of this document.

**Referee's comment 14.** Fig.5: As mentioned under 5, you should point out which eight models are shown and/or why only these are shown.

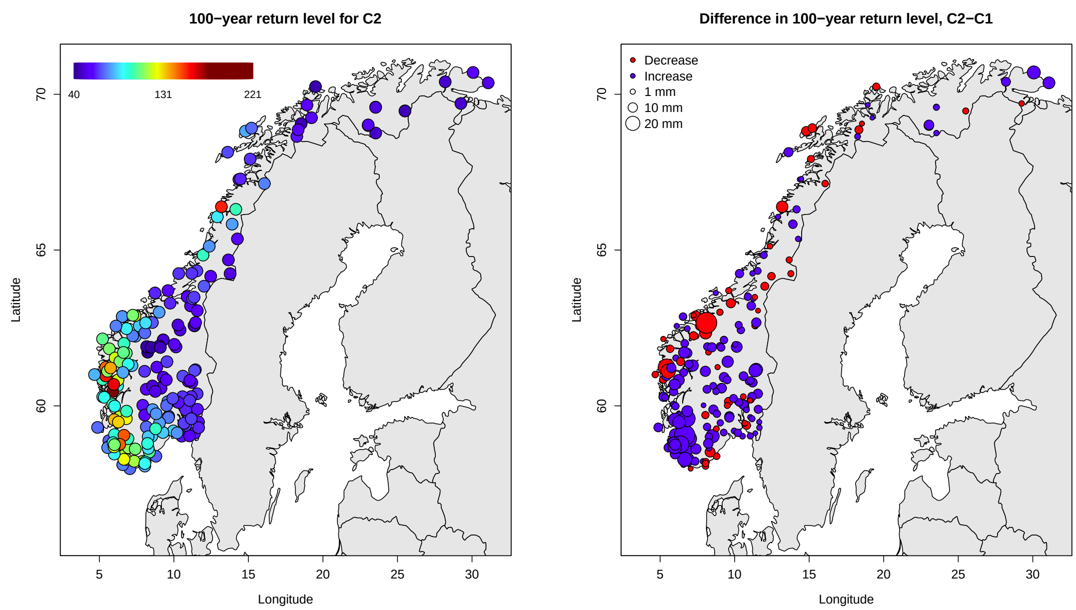


Figure 3: Left: map of 100-year return level estimated on $C^{(2)}$ (1979-2009) with MEWP$(0.5, 2, 8)$. Right: difference in 100-year estimated on $C^{(1)}$ and $C^{(2)}$.

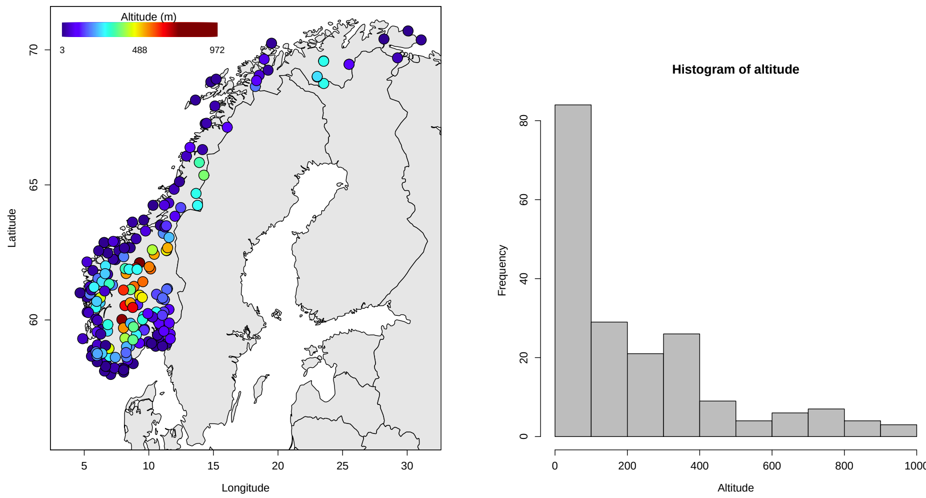


Figure 4: Left: location and altitude (m.a.s.l) of the stations. Right: histogram of altitude (m.a.s.l).

**Authors' reply:** As commented in 5., this is a mistake and the 12 models will be shown in the next version of the article (Figures 1 and 2 above).

**Referee's comment 15.** The literature on this field is becoming quite large, and some papers are very relevant to the current manuscript. I encourage the authors to add more references to former work on extreme value estimation; both in general, and for the region of interest.

**Authors' reply:** That's true, there are many articles on extreme value analysis of rainfall, including several in Norway. Let cite [Dyrrdal et al., 2014, Dyrrdal et al., 2015] for the case of Norway and [Papalexiou and Koutsoyiannis, 2013, Serinaldi and Kilsby, 2014] for studies at a more global scale. The introduction of [Serinaldi and Kilsby, 2014] provides a good review of studies on rainfall extremes.

We warmly thank the reviewer for the technical points 16 to 31 that will be corrected in the next version of the paper.

# References

[Brigode et al., 2014] Brigode, P., Bernardara, P., Paquet, E., Gailhard, J., Garavaglia, F., Merz, R., Mićović, Z., Lawrence, D., and Ribstein, P. (2014). Sensitivity analysis of SCHADEX extreme flood estimations to observed hydrometeorological variability. *Water Resources Research*, 50(1):353–370.

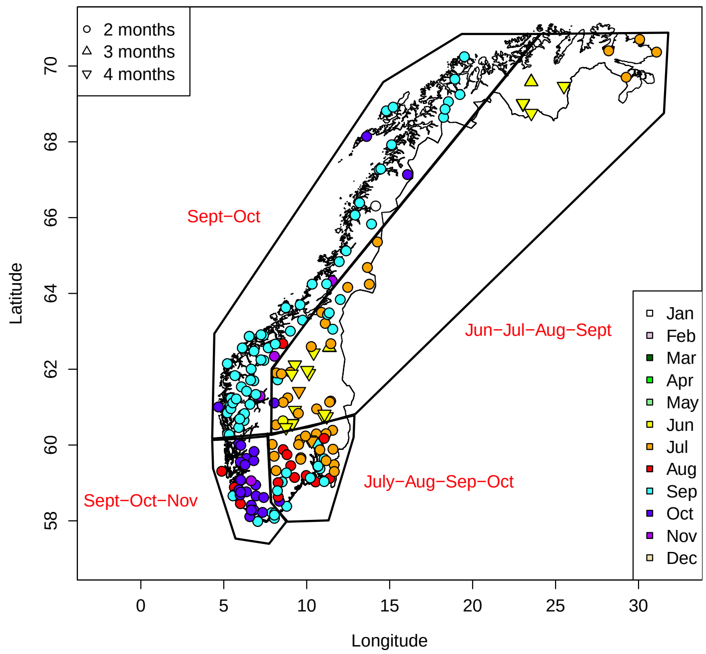


Figure 5: Length of the season-at-risk (shapes) and first month of the season (colors) for each station, with model MEWP(0.5, 2, 8). The local definition of seasons is used in Section 4.2, while the regional definition, with four regions, is used in Section 4.3.

[Dyrrdal et al., 2015] Dyrrdal, A. V., Lenkoski, A., Thorarinsdottir, T. L., and Stordal, F. (2015). Bayesian hierarchical modeling of extreme hourly precipitation in Norway. *Environmetrics*, 26(2):89–106.

[Dyrrdal et al., 2014] Dyrrdal, A. V., Skaugen, T., Stordal, F., and Frland, E. J. (2014). Estimating extreme areal precipitation in Norway from a gridded dataset. *Hydrological Sciences Journal*.

[Garavaglia et al., 2011] Garavaglia, F., Lang, M., Paquet, E., Gailhard, J., Garçon, R., and Renard, B. (2011). Reliability and robustness of rainfall compound distribution model based on weather pattern sub-sampling. *Hydrology and Earth System Sciences*, 15(2):519–532.

[Midttømme et al., 2011] Midttømme, G., Pettersson, L., Holmqvist, E., Nøtsund, Ø., Hisdal, H., and Sivertsgård, R. (2011). Retningslinjer for flomberegninger (Guidelines for flood estimation). *NVE Retningslinjer*, 4/2011.

[Papalexiou and Koutsoyiannis, 2013] Papalexiou, S. M. and Koutsoyiannis, D. (2013). Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, 49(1):187–201.

[Serinaldi and Kilsby, 2014] Serinaldi, F. and Kilsby, C. G. (2014). Rainfall extremes: Toward reconciliation after the battle of distributions. *Water Resources Research*, 50(1):336–352.