Hydrol. Earth Syst. Sci. Discuss., 12, C4569–C4590, 2015
www.hydrol-earth-syst-sci-discuss.net/12/C4569/2015/

Hydrology and
Earth System
Sciences
Discussions
Open Access

# *Interactive comment on* "Accounting for three sources of uncertainty in ensemble hydrological forecasting" *by* A. Thiboult et al.

**A. Thiboult et al.**

antoine.thiboult.1@ulaval.ca

Received and published: 31 October 2015

The authors would like to thank the two anonymous referees for their careful reading and the interesting comments they provided that will contribute to the quality of the paper.

Shortly after initial submission, while working on the same database, we realize that the streamflow measurements of two of the 20 catchments where of dubious quality. Several clues led us conclude that they should not be longer included in the catchment pool. We sincerely apologize for any inconvenience this may have caused.

The two catchment that have been withdrawn from the initial submission are the ones that often behaved like outliers and were the most unreliable. Thus, the new results are more homogeneous. The two problematic catchments have been substituted by two new. Also, Figures 3, 4, 5, 6, 8, 9, 10, 11, 12, 13 were updated and are provided along this answer. Even if this does not affect any of the conclusion, the author suggest several modifications of the text:

- page 7193, line 17-19: " Exceptions can be occasionally observed for catchment 3, 17, and 20 where only one or two models outperform the ensemble." should be replaced by " Exceptions can be occasionally observed for catchment 3 and 17 where only one or two models outperform the ensemble."
- page 7194, line 5-6: we suggest to modify "For the first lead time, most of the catchments are close to reliability while there is a clear outlier for which accuracy skills do not match its corresponding spread. In fact, this low performing catchment exhibits a constant hydrological wet bias – partially explained by a meteorological forecast wet bias that over-forecasts precipitations by 15% – that is not captured by any of the models even if the global tendency is respected." to "For the first lead time, most of the catchments are close to reliability while there are two outliers for which accuracy skills do not match their corresponding spread. In fact, these catchments exhibits a constant hydrological bias – partially explained by an inaccurate meteorological forcing – that is not captured by any of the models even if the global tendency is respected"
- page 7194, line 20: "(outlier)" should be deleted.
- page 7194, line 20-23: "Data assimilation is particularly effective on catchments that present a systematic bias. For example, catchment number 11 that was problematic from the first lead time lies among the other catchments in terms of performance." should be deleted, even if we think that DA is particularly effective on catchment that have a systematic bias, but this assertion is no longer explicitly supported by the new Figure.

- page 7196, line 12-13: values should be replaced from "While they are almost identical with a value of 0.55 and 0.57 mm day-1 respectively for the day 3, G spread drops to 0.44 mm day-1 for day 9 while the use of the MEPS maintains the spread to 0.55 mm day-1" to "While they are almost identical with a value of 0.58 mm day-1 and 0.59 mm day-1 respectively for the day 3, G spread drops to 0.45 mm day-1 for day 9 while the use of the MEPS maintains the spread to 0.59 mm day-1."
- page 7197, line 18-19: "and only model 1 and 5 perform" should be replace to "and only models 1, 5, and 17 perform"
- page 7198, line 3: "catchment 19", should be replaced to "catchment 20".
- page 7199, line 1: "reducing the overdispersion with a sensible decrease in the ensemble spread from 0.65 to 0.54 mm day-1" should be replaced with "reducing the overdispersion with a sensible decrease in the ensemble spread from 0.72 to 0.57 mm day-1"
- page 7199, line 8-11: "The two outlier catchments that exhibit poorer reliability present an underdispersed forecast that is a bit more pronounced for the H system than the H system (see Fig. 9). This indicates that uncertainties used to define the EnKF perturbations are under-estimated." should be suppressed. We also suggest to replace by "As a matter of fact" by "Finally".

**The paper analyses different descriptions of uncertainty for hydrological ensemble forecasting and discusses their relative merit. It provides a valuable contribution to the research on probabilistic hydrological forecasting. However, different assumptions are made that may have a significant impact on the results and the general conclusions of the study. More elaborate discussions of the impact of these assumptions are needed (see specific comments below).**

## Specific comments

**Page 7183, line 15. The term 'open loop scheme' may not be familiar to all readers. It is explained later in Section 2.**

Indeed, a short definition will be added in a future version.

**Page 7185, line 7-8. Not clear why conversion to local time reduces the forecast horizon?**

The meteorological forecast retrieved from the database is issued from 12am UTC up to ten days ahead with a 6-hour time step. Since the hydrometeorological day starts at 6am EST (or 12 UTC) for the catchments of the study, the first 12 hours of forecast are not used. Consequently, only 9 and half days of meteorological forecast are available. Lastly, because the time step is daily, the remaining 12 hours of forecast of the last day have been discarded.

**Page 7185, line 10-14. Why first downscale and then aggregate to catchment rainfall? You could derive catchment rainfall directly from the ECMWF forecast.**

We believe that this is something that should be avoided. The raw resolution of the ECMWF is too coarse for this application and do not match systematically catchment size, as 6 of them are smaller than 1000km2. Without interpolation, it is possible that only one meteorological forecast grid point would fall within catchment boundaries, if any. Also, as most catchment straddle several initial ECMWF grid points, interpolation allow to take into account the contribution of each of these grid point. Finally, considering the influence of more than one meteorological grid point allows for smoothing out individual members occasional instability.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

**Page 7185, line 14-18. Pre-processing of meteorological forecasts is widely used in hydrological forecasting systems to improve forecast accuracy and reliability. Since this is not done in the study, the value of the rainfall forecast will most probably be underestimated.**

We agree on that point but it was not intended to investigate such pre-processing in this article as we deemed that it is a step that is sufficiently complex that it may require specific investigation. However, one can note that recent attempts at pre-processing meteorological forecasts have been performed without much success at improving streamflow forecasts, although the improvement of the meteorological forcing was indeed successful (e.g. Verkade, J. S.; Brown, J. D.; Reggiani, P. & Weerts, A. H. Post-processing ECMWF precipitation and temperature ensemble reforecasts for operational hydrologic forecasting at various spatial scales Journal of Hydrology, 2013, 501, 73-91). In the light of the comments of both reviewers, we realize that we should remind the reader and emphasize that no pre-processing is used and that the interpretation of the results should be done accordingly, in particular in Section 3.3. It is expected that a successful post-processing would enhance the MEPS capabilities to decipher meteorological uncertainty and would be eventually cascade these benefits through the hydrological systems, thus leading to better accuracy and reliability.

**Page 7187, line 20. The H operator has an index t in the equation. I would not expect H to be time varying.**

Indeed, the index will be removed in the future version.

**Page 7188, line 5-6. Different variants of the EnKF have been proposed in literature. Which method is applied here, and why?**

The EnKF has been implemented in its traditional form (Evensen, G. The Ensemble Kalman Filter: theoretical formulation and practical implementation Ocean Dynamics,

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

2003, 53, 343-367 ). We had the opportunity to developed our expertise by studying in detail the interactions between the filter and the different models (On the difficulty to optimally implement the Ensemble Kalman filter: An experiment based on many hydrological models and catchments, Journal of Hydrology, Volume 529, Part 3, October 2015, Pages 1147-1160 A. Thiboult, F. Anctil). This variant of the EnKF, if properly set, proved to be able to efficiently reduce the initial condition uncertainty and thus to fulfill the expectations we have from it, in regards with to the hydrometeorological setup that is presented here.

**Page 7188, line 21-22. How is reliability and accuracy evaluated in the tuning of the EnKF?**

The accuracy is assessed with the NSE that is computed on the median of the ensemble and the Normalized Root-mean-square error Ratio (NRR) is used for reliability assessment. Then, the 2-step criterion described page 7189 line 1-3 is applied. Results concerning the tuning of the EnKF can be found in the article cited in the previous answer.

**Page 7188, line 22-27. Only uncertainty in model forcing is assumed, and hence this uncertainty should compensate also for other model uncertainties such as parameter uncertainty. Model parameter uncertainty could be included in the EnKF. This would most likely improve the reliability of the EnKF since this would add uncertainty in the forecast period by propagation of parameter uncertainty.**

It is practically hard to untangle uncertainties through the use of EnKF only. EnKF, it its traditional form, can decipher the overall predictive uncertainty but does not distinguish between input-output, structural, and parameter uncertainty. By artificially and deliberately overestimating the input uncertainty, it is possible to compensate for the other uncertainties and achieve reliability for simulation and possibly for the first (and sometimes second) forecast day. This may be desirable only in a case where there is no other tool available to handle the other sources.

We did not consider dual parameter-state variable updating since the multimodel approach allows to take into account parameter uncertainty without the need to modify (update) time invariant values. Thus, model parameter uncertainty is treated outside of the EnKF through the multimodel approach. Moreover, it is shown that several dissimilar hydrological models bring much more diversity than traditional parameter uncertainty estimations, thus indicating that structural uncertainty, in some ways, encompasses parameter uncertainty (Poulin, A., Brissette, F., Leconte, R., Arsenault, R., and Malo, J. S.: Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin, J. Hydrol., 409, 626–636, doi:10.1016/j.jhydrol.2011.08.057, 2011. 7183)

**Page 7188, line 27-28. The definition of the state vector is not clearly described. The state vector is uniquely defined by the system model. Typically, for lumped, conceptual rainfall-runoff models it will consist of storages of the different conceptual reservoirs.**

By state vector we meant the ensemble of state variables that are updated. It is indeed a mistake as the definition of state vector is the one you gave. This will be changed.

**Page 7194, line 7-8. This illustrates the problem of not pre-processing rainfall forecasts cf. comment above.**

We are aware of the limitations that are induced by not pre-processing rainfall forecast and we are currently carrying out research on the subject. The sentence line 7-8 was initially written to explain the behavior of one of the catchment that has been withdrawn from the database but your remark about pre-processing remains valid.

**Page 7195. The results for the EnKF are due to an incomplete description of the uncertainty. It provides a good description of the initial uncertainty but this is quickly washed out of the system as forecast lead time increases. Use of a more elaborate description of the uncertainty in the EnKF would improve the reliability, e.g. by including model parameter uncertainty cf. comment above.**

This framework should be regarded as a possible way toward accurate and reliable forecast but we strongly concur that it is not the only one. More sophisticated version of the EnKF may indeed improve the description of uncertainties for longer lead times. In a different study, we tested direct state variable perturbations (Abaza, M.; Anctil, F.; Fortin, V. & Turcotte, R. Sequential streamflow assimilation for short-term hydrological ensemble forecasting. Journal of Hydrology, 2014, 519, 2692-2706). It contributes to maintain the dispersion a little longer (about 1 day longer) but the spread is at the end also maintained by the MEPS forcing spread. However, these results should not be compared in a very strict way with this paper since the hydrological models are different. A secondary objective of the article is also to show that with the traditional formulation of EnKF the spread (and the corresponding description of uncertainty) is not sufficient but can be compensated by the use of multimodel. Also, the combination of the multimodel and the traditional EnKF makes that is not necessary to resort to dual state variable-parameter updating. By keeping the parameters time invariant, the inner model dynamic is better preserved.

**Page 7196. I think the lack of pre-processing of the rainfall forecast ensemble can explain the small impact observed of using a probabilistic rainfall forecast.**

Despite the absence of pre-processing, the reliability is still better with probabilistic forecast. The hydrological ensemble spread is substantially larger and one could expect this spread to contribute more actively to reliability if the bias of the forcing would be removed.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

**Page 7198, line 20-23. Not clear why this would correspond to an optimal EnKF?**

Traditional EnKF accounts for input and output uncertainty explicitly. It could be optimal in a perfectly controlled environment where only the input and output are subject to uncertainty (in a synthetic experiment for example). Thus there shouldn't be any uncertainty in the structure / parameter / conceptualization. It is suboptimal in real cases as it has to account for other sources of uncertainty if used with a single model.

**Page 7198, line 26-28. Not clear.**

This assertion is closely related to the previous question. In the case where the different sources are not explicitly accounted for by dedicated tools, the EnKF has to compensate for them. One way to achieve reliability is to add perturbations to input. However, there is no obvious way to know by which amount the uncertainty on input should be overestimated to compensate for the other uncertainties. Thus, to ensure hydrological reliability, one needs to perform a "calibration" of EnKF hyper parameters (research of required noise magnitude) which is a fastidious step.

**Page 7198. Figure 12 is not referred in Section 3.5.**

Indeed, it was originally meant to be in the article but we finally decided to put it only as supplementary material as the main changes concern reliability and forgot to remove the corresponding figure from the manuscript.

**Page 7200, line 5-10. There seems to a contradiction here. First, it is stated that the EnKF does not provide a satisfactory uncertainty propagation. And then it is stated that the EnKF is the component that provides the most dispersion.**

It requires indeed some clarifications. It should have been specified that it is the component that provides the most dispersion, but only for the first lead times.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 7179, 2015.

| River name | Area (km2) | River length (km) | Average slope (%) | Mean ann. Q (m3/s) | Coeff. of variation | Mean ann. P (mm) | Mean ann. snow (cm) |
|---|---|---|---|---|---|---|---|
| Trois Pistoles | 923 | 52 | 0.52 | 18 | 1.81 | 1109 | 382 |
| Du Loup | 512 | 45 | 0.78 | 10 | 1.47 | 1050 | 378 |
| Gatineau | 6796 | 190 | 0.12 | 127 | 1.08 | 1023 | 332 |
| Dumoine | 3743 | 145 | 0.13 | 50 | 0.81 | 968 | 297 |
| Kinojévis | 2572 | 83 | 0.12 | 39 | 1.12 | 921 | 324 |
| Matawin | 1383 | 68 | 0.29 | 24 | 1.11 | 1025 | 328 |
| Croche | 1551 | 102 | 0.33 | 29 | 1.24 | 996 | 360 |
| Vermillon | 2650 | 145 | 0.20 | 39 | 1.10 | 957 | 312 |
| Batiscan | 4483 | 167 | 0.45 | 96 | 1.03 | 1162 | 381 |
| Saint-Anne | 1539 | 84 | 0.81 | 51 | 1.20 | 1412 | 502 |
| Bras du Nord | 643 | 77 | 0.82 | 19 | 1.21 | 1385 | 499 |
| Du loup | 767 | 57 | 0.78 | 12 | 1.27 | 1020 | 332 |
| Aux Ecorces | 1107 | 54 | 1.04 | 28 | 1.09 | 1236 | 450 |
| Métabetchouane | 2202 | 155 | 0.43 | 48 | 1.19 | 1168 | 420 |
| Péribonka | 1010 | 101 | 0.50 | 19 | 1.16 | 1000 | 376 |
| Ashuapmushuan | 15342 | 342 | 0.16 | 300 | 0.92 | 984 | 379 |
| Ashuapmushuan | 11200 | 232 | 0.12 | 227 | 0.88 | 1001 | 394 |
| Au Saumon | 586 | 69 | 0.65 | 8 | 1.36 | 877 | 334 |
| Mistassini | 9534 | 278 | 0.20 | 200 | 1.08 | 1004 | 409 |
| Valin | 761 | 59 | 1.06 | 24 | 1.13 | 1123 | 453 |

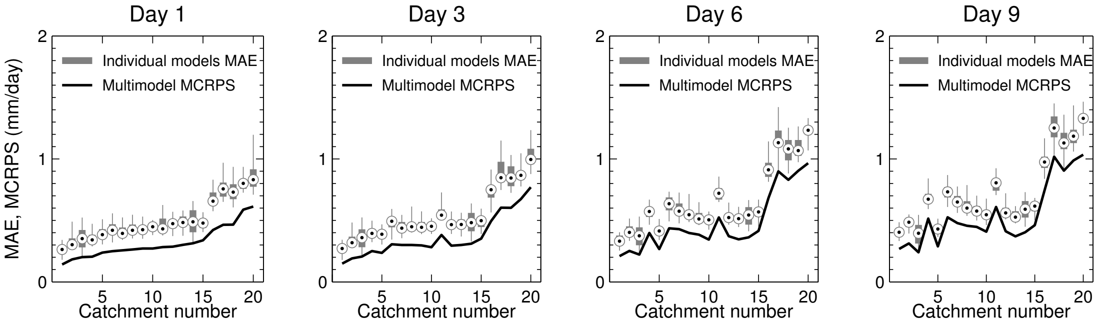

**Fig. 1.** Comparison of individual models MAE and multimodel MCRPS sorted by increasing multimodel MCRPS for the first day (version A vs. E).

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

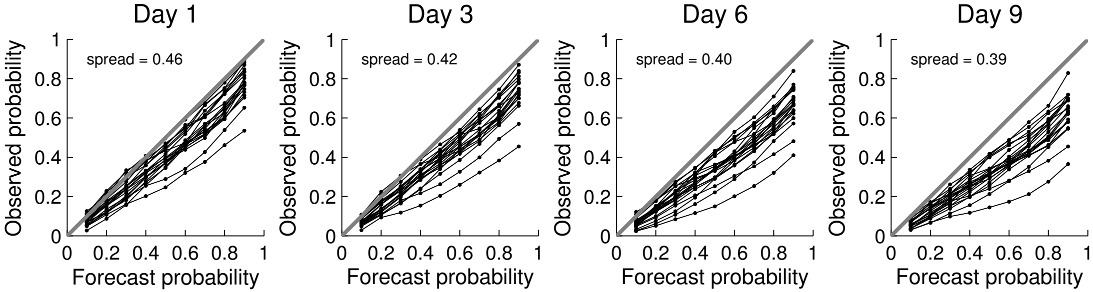

**Fig. 2.** Reliability of the multimodel ensemble (system E) for all individual catchments. The spread represents the square root of mean ensemble variance averaged over all catchments.

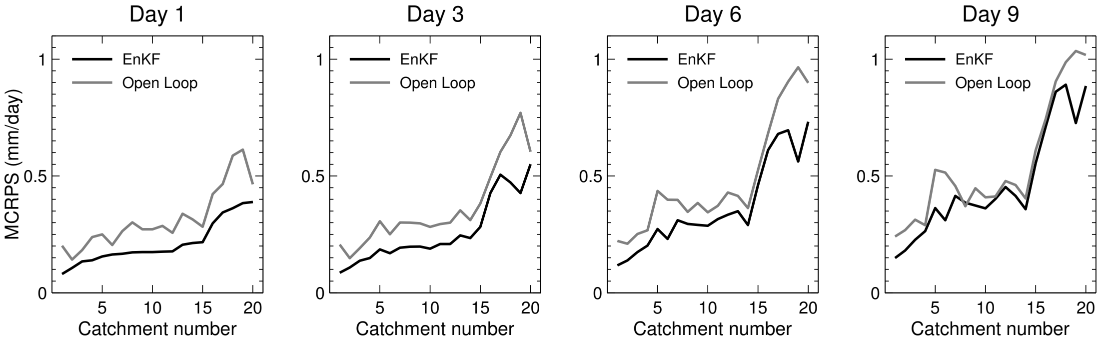

**Fig. 3.** Comparison of open loop and EnKF multimodel MCRPS sorted by increasing EnKF MCRPS (system E vs. G).

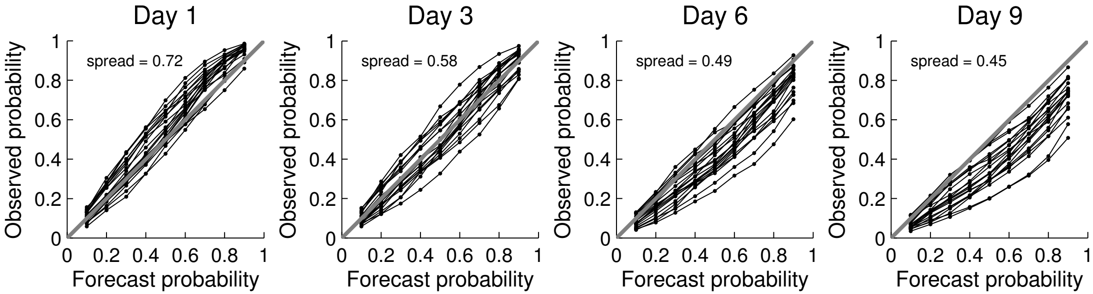

**Fig. 4.** Reliability of the EnKF multimodel ensemble (system G) for all individual catchments. The spread represents the square root of mean ensemble variance averaged over all catchments.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

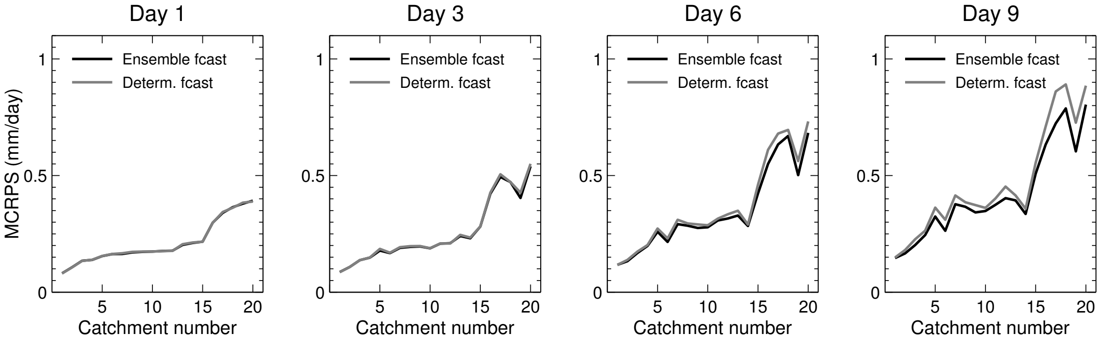

**Fig. 5.** Comparison of EnKF multimodel MCRPS with deterministic and ensemble meteorological forcing (system G vs. H).

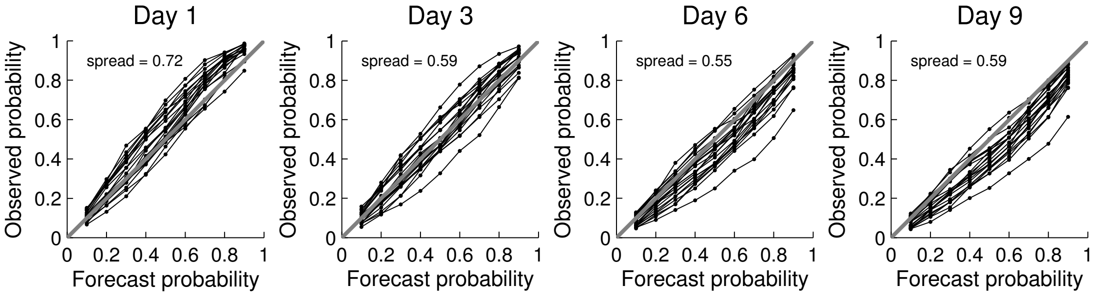

**Fig. 6.** Reliability of the EnKF multimodel ensemble with MEPS forcing (system H).

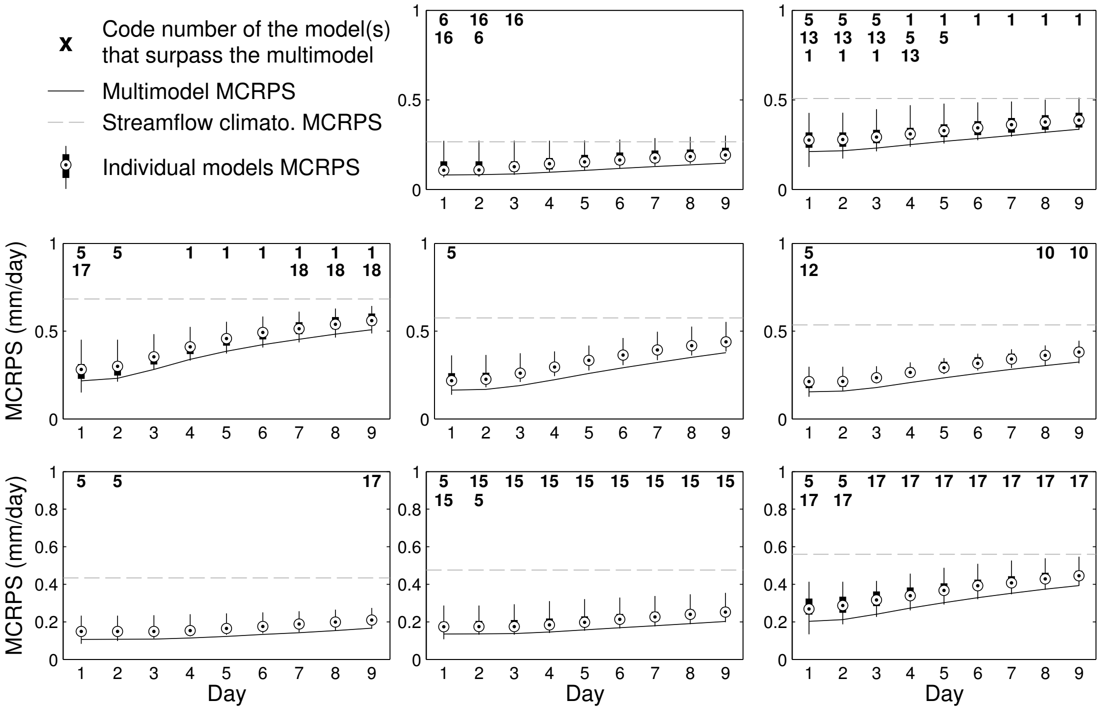

**Fig. 7.** Comparative examples of the MCRPS on 8 watersheds of the EnKF individual models and the EnKF multimodel, both using MEPS forcing (system D vs. H).

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

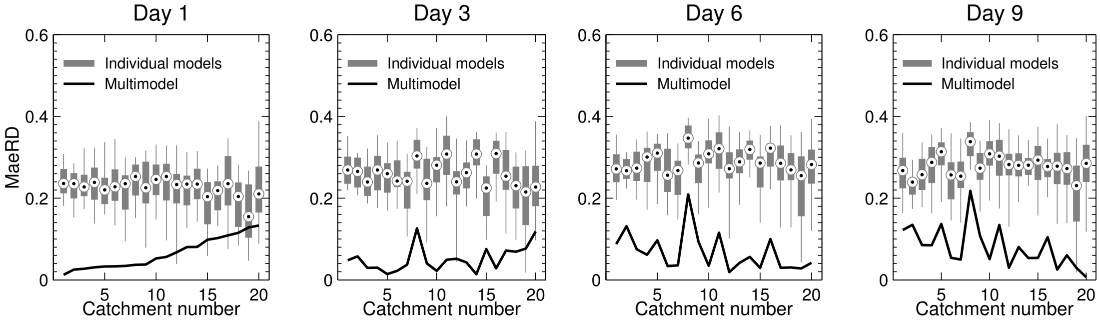

**Fig. 8.** Comparison of the deviation from perfect reliability of EnKF individual models and the EnKF multimodel, both using MEPS forcing sorted by increasing EnKF multimodel MaeRD for the first day (sys

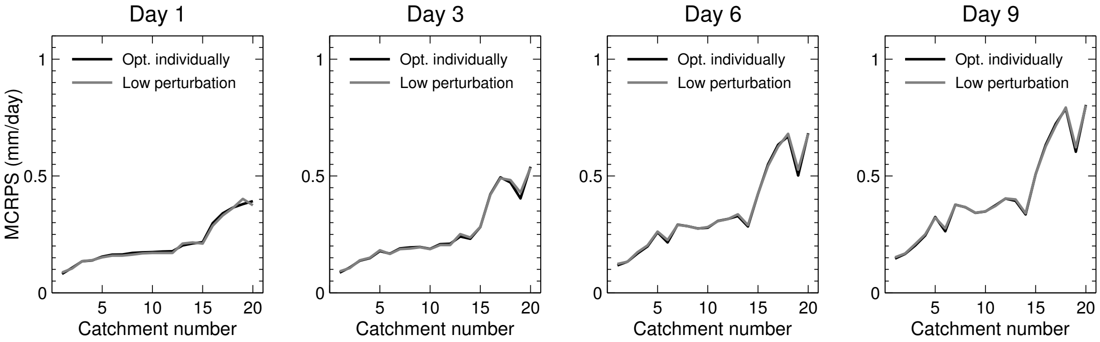

**Fig. 9.** Comparison of EnKF multimodel MEPS systems using either individually optimized EnKF perturbations or lower input-output perturbations (system H vs. H ).

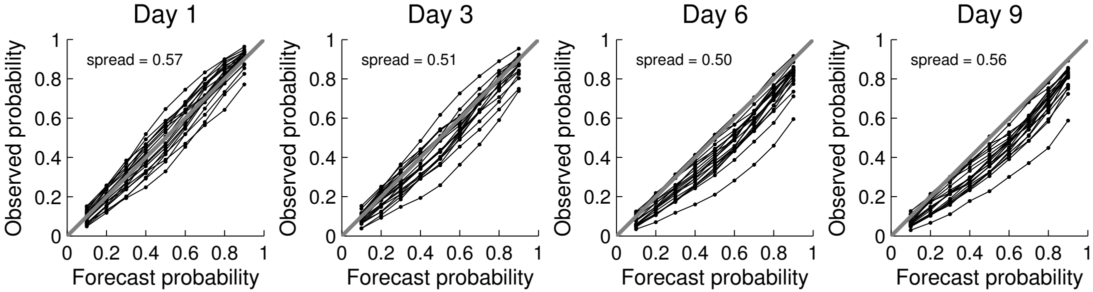

**Fig. 10.** Reliability of the EnKF multimodel ensemble with MEPS forcing and lower input- noutput perturbations (system H ).

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

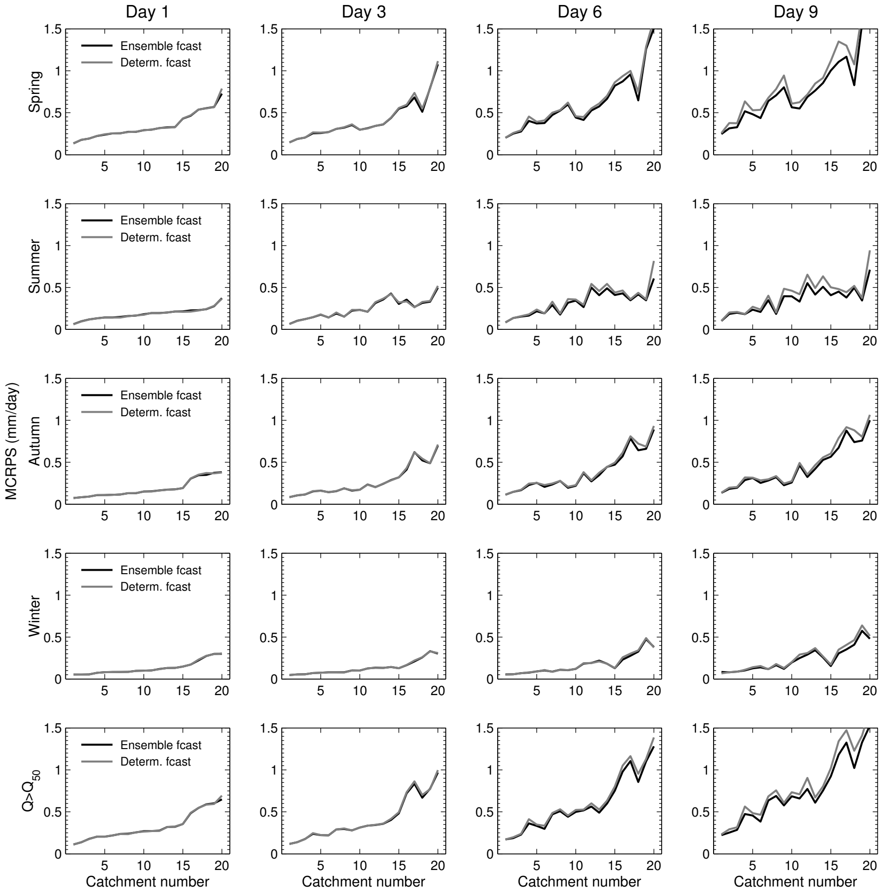

**Fig. 11.** Comparison of EnKF multimodel MCRPS per season with deterministic and ensemble meteorological forcing (system G vs. H).

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper