Hydrol. Earth Syst. Sci. Discuss., 6, C2309–C2314, 2009
www.hydrol-earth-syst-sci-discuss.net/6/C2309/2009/

Hydrology and Earth System Sciences Discussions

Interactive comment on "An evaluation of the canadian global meteorological ensemble prediction system for short-term hydrological forecasting" *by* J. A. Velázquez et al.

J. A. Velázquez et al.

francois.anctil@gci.ulaval.ca

Received and published: 8 October 2009

We thank Referee #1 for his objective review and important comments. Follows our answers:

General comments: The introduction has been improved in order to justify the use of a single storm. At P 4894 after L24 we have added:

"Other hydrological applications of M-EPS have been recently reported. For instance, Jaun et al. (2008) have studied an extreme flood event in August 2005, for a Rhine sub-catchment of 34500 km2, for which the precipitation had a return period over 10

to 100 years. They coupled the meteorological operational system COSMO-LEPS, which downscales the ECMWF-EPS to a resolution of 10 km, with a semi-distributed hydrological model, and concluded that their H-EPS is effective and provides additional guidance for extreme event forecasting in comparison to a deterministic forecasting system. These conclusions are confirmed by a second analysis reported by Jaun et al. (2009), based on a longer duration (two years). Another case study, over a 9-month period, is presented by Renner et al. (2009). It evaluates the performance of a H-EPS for various Rhine stations: catchments areas from 4000 to 160000 km2. Two meteorological ensembles are then confronted: the low resolution ECMWF-EPS and the high resolution COSMO-LEPS ensemble. Results showed that the increased resolution meteorological model provides higher scores, particularly in the short term precipitation forecasts. The authors concluded that there is a need for the downscaling of the ensemble forecast in order to obtain to a more representative scale for the sub-basins in the hydrological model. There are some difficulties by using the H-EPS for flood forecasting. Cloke et al. (2008) discuss some of these problems as, for example, the difficulties in assessing flood forecasts because of their rarity and the difficulties to compare consecutive floods because of the spatial and temporal non-stationarity of the catchments. The authors suggest that there is no other option than to analyse the performance of an EPS driven flood forecast on a case by case basis, and gradually, over the decades, to build up a database of several hundred of flood events on which to base a more thorough flood analysis. This paper also presents an extensive list of recent studies applying ensemble approaches for runoff forecasts with a variety of catchment areas, periods, hydrological models and meteorological EPS". We have also included additional information about the precipitation uncertainty by comparing the reliability of the forecasted streamflow against the observed discharge, and the reliability of the forecasted streamflow against the base-line simulation. This helped us to evaluate the uncertainty due to meteorological data. At P. 4903 after L26 we have added the following:

"We also evaluated the lack of reliability due to the M-EPS and to the hydrological model. The M-EPS ratio ïĄď drawn in Fig. 9 shows the flatness of the M-EPS rank histograms for all prediction horizons. The M-EPS may be responsible for part of the reported under dispersion. This is assessed in Fig. 10 that shows the 72h reliability diagrams for four sites and two approaches: the first one evaluates the reliability of the updated flow forecasting against the observed discharge, as in figure 8; the second one evaluates a no-updated forecast against a base line simulation (a simulation produced with observed precipitation). The first approach is used to evaluate the reliability of the ensemble forecast, including meteorological, hydrological and observational uncertainties, while the second one is used to evaluate reliability due to meteorological model only (e.g. Renner et al 2009). For all sites, there is an improvement of the reliability, especially for the site Du Lièvre T50, for which it could be inferred that the bias originates from the hydrological model more than in the meteorological forecast. In other cases, like Kénogami T15, both lines are very close to each other and present a more marked under dispersion. This is an indication that the meteorology is biased. In the case of the hydrological model, we could conjecture that part of the lack of reliability of the H-EPS is due to the fact that uncertainty in the initial conditions is not taken into account." This addition to the manuscript asked for the following supplemental comment in the conclusion (P4904 L23): "We have also distinguished the lack of reliability due to the M-EPS and to the hydrological model. Results showed that the meteorology is biased and may be responsible for part of the reported under dispersion. In the case of the hydrological model, we conjecture that the lack of reliability of the H-EPS is in part due to the fact that uncertainty in the initial conditions of the hydrological model is not taken into account." We also agree with the referee's comment that there is no need for post-processing, so we have removed this part of the conclusion.

Specific comments: P 4894, L 25 – 26: Global Environmental Multiscale model (GEM) The definition has been added

P 4897, L 20 – 27: "based on climate observations and CEHQ state variables"; do you mean "based on climate observations in order to estimate the initial conditions of

the Hydrotel model state variables? It is not clear from this paragraph how the initial conditions on the 12 October and following days are obtained. The reviewer is correct in that some information is missing, namely how we progressed form one day to the other between 11 to 31 October. The following sentence is thus added after the one identified by the reviewer: "For the test period, the model is again driven by the climate observations, but the state variables are left unconstrained".

P 4898, L 6 – 17: this paragraph should be improved. Why "routine" comparison? Improvements have been proposed to this paragraph.

L 9–10: do you refer to the reliability that is introduced later? "The performance corresponds between the predicted probability and the actual frequency of occurrence". Yes, we referred to the reliability; we wanted to stress the necessity of using a score which accounts for the reliability of the forecasts. Improvements have been proposed to this paragraph.

L 12: what are the needs in this particular study? Our needs (and selection) are expressed starting with the next paragraph. To make things clearer, the following was added line 13: "For the present study, these will be described next."

L13 – 17: In such a verification study, how could hedging occur? Hedging cannot directly occur in this study because the predicted stream flows are not interpreted by any forecaster. However, the use of an improper score may be equivalent to hedging. That is the reason why this issue was raised here. We believe that the text is detailed enough on that issue for not proposing any modification.

P 4899, L 4 – 10: In the equation (2), an expectation operator is missing. We thank the reviewer for picking up that mistake. The equation has been corrected. CRPS=E|X-x"t" |-0.5E|X-X'| How is the Monte Carlo approximation implemented, and why not to compute CRPS for an ensemble system like in Hersbach (2000)? We have exploited a gamma distribution from which 1000 samples were randomly drawn. To be more precise, the following replaces the actual text for lines 9 and 10. "where X and X' are independent vectors consisting of 1000 random values from a gamma distribution adjusted to the predictive function." As explained before in the text (line 3), an exact solution to eq. (1) only exists for Normal predictive distributions. Hersbach (2000) approximates the predictive distribution of the variable of interest by a discret distribution obtained from the finite set of values formed in the ensemble (equation 21). We have chosen instead to fit a gamma distribution, hence obtaining a continous approximation to the predictive distribution. Using a gamma distribution has some advantages for decision makers as all measurable sets of flow values have a positive probability of occurrence, which is not the case for the approximation used by Hersbach.

P 4899, L 11 – 16: "certain time": it is understood only at Section 3 that the average is made over the 17 values corresponding to a given lead time and outlet. We agree that the expression "for a certain time" is a bit vague. We are now using "for one time step".

P 4901, L 1 – 10: you mean you select 10 members out of 20 for a given forecast and forecast day, rank them together with the observed streamflow and repeat 200 times. "Quasi equiprobable" is being tested actually. The reviewer interpretation is correct. Also, we have not tested the equiprobability of our procedure. That is the reason why we have qualified our resulting members a quasi equiprobable. We should also stress that the members of the tested Canadian ensemble prediction system are also considered as quasi equiprobable.

P 4902, L 17: the sentence could be rephrased ("moments"). "moments" was change to "at times".

P 4903, L 26 – P 4904, L 2: the authors are true that uncertainty in the initial conditions of the hydrological model should be taken into account but the flatness of the histograms of the ensemble precipitation forecasts is not sufficient to exclude their role in the lack of reliability of the streamflow forecasts. More information about the verification of the precipitation is needed. We added more information at P. 4903 after L26 (see above).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 4891, 2009.

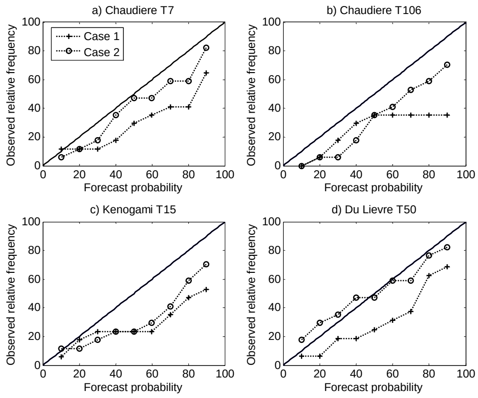

Fig. 1. Figure 10. H-EPS 72-h reliability diagrams. Case 1 evaluates the reliability of an updated flow forecasting against the observed discharge. Case 2 evaluates a no-updated forecast against the base line