Hydrol. Earth Syst. Sci. Discuss., 3, S1230–S1232, 2006
www.hydrol-earth-syst-sci-discuss.net/3/S1230/2006/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Skill and relative economic value of medium-range hydrological ensemble predictions" *by* E. Roulin

E. Roulin

Received and published: 23 October 2006

I am indebted to the reviewer for his or her comments. The use of the sample variance as a reference for measuring the skill of the medium-range flow predictions has the advantage of making the relative contributions of the different aspects to the Brier Skill Score easy to compare (Eqs 4 and 7). There is a range of alternative forecasts that would be skillful in this situation: persistence of the streamflow, persistence of the weather used as forcing to the hydrological model, ensemble streamflow predictions using weather series drawn from past climate at the same time of the year. This last alternative was investigated previously (Roulin and Vannitsem 2005) and it has skill but less than ensemble streamflow predictions based on ECMWF-EPS. Indeed, the skill of ensemble streamflow predictions using EPS compared to this alternative probabilistic

forecast was above 0.5 after three days during winter and remained significantly positive after nine days. In the present study, the hydrological ensemble prediction system is compared to different deterministic prediction systems and the reference to compute the skill scores is taken simply as the sample variance.

For completeness, and following the suggestion of the reviewer, the Brier Skill Score and its CR and LBR decompositions for the persistence forecasts have been included in Fig. 9: persistence has a relative resolution decreasing steadily and, as soon as after two days, the skill of persistence is worst than climatology due also to increasing biases.

The skill of medium-range flow predictions was analyzed separately for the hydrological winter and summer in a previous study (Roulin and Vannitsem 2005). In the present study, the climatological probability is used to define threshold values for the streamflow for the whole year. For high flows (Figs 9 and followings), the threshold is defined with the 95th percentile in order to include a sufficient number of events but this is still low compared to operational pre-alert threshold (Roulin and Vannitsem 2005). Note that the Brier Skill Score and other relative measures are estimated by comparison against the "sample climatology" which is not a real forecast alternative since it is known a posteriori. Pooling the forecasts for the entire year doesn't alter the comparison among the different systems.

The relative economic value estimated with a simple static cost-loss is a useful validation measure to help in choosing among different alternative systems. The two-stage dynamic problem requires more specific information than a single cost-loss ratio and the dynamic elements of decision are somewhat arbitrary. Nevertheless, the dynamic aspects associated with the medium-range forecasts are worth being addressed with regard, for instance, to the belief that persistent probabilistic forecasts have more value. A simpler two-stage dynamic problem to test this hypothesis would consist in taking a single action provided not only that a probability threshold is exceeded but also that a probability threshold, possibly different, has been exceeded in a previous

HESSD

3, S1230–S1232, 2006

forecast. With the variables defined in Figure 5, the expenses could be summarized with $E=\{0, L, 0, L, 0, L, C, C+L-L_1\}$. An even simpler static problem related to this view consists in pooling successive forecasts into lagged super-ensembles. These two approaches have been tested but not reported in detail in the paper; none has led to additional value or skill. This comment has been mentioned in the revised paper.

The analysis did not account for forecasts made when a river is exceeding a flood threshold and, as pointed out by the reviewer, this makes the economic value inflated. This problem is linked with the persistence of the streamflow already mentioned and, as a simple control, the value of an alternative forecasting system consisting of persistence is also included in Fig. 13 where it is shown marginal. For shorter lead times, the value of persistence is greater but doesn't modify the conclusions.

None of the deterministic forecasts has been "optimized" by selecting a streamflow threshold. The same threshold has been used for the deterministic and probabilistic forecasts. The ensemble forecasts are optimized by selecting the probability threshold that results to the highest value for each cost-loss situation. Concerning a fairer comparison with probabilistic forecasts, the Brier Skill Score of the deterministic forecasts could be improved by adding some random noise (Dr Kees Kok, KNMI, personal communication).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 3, 1369, 2006.