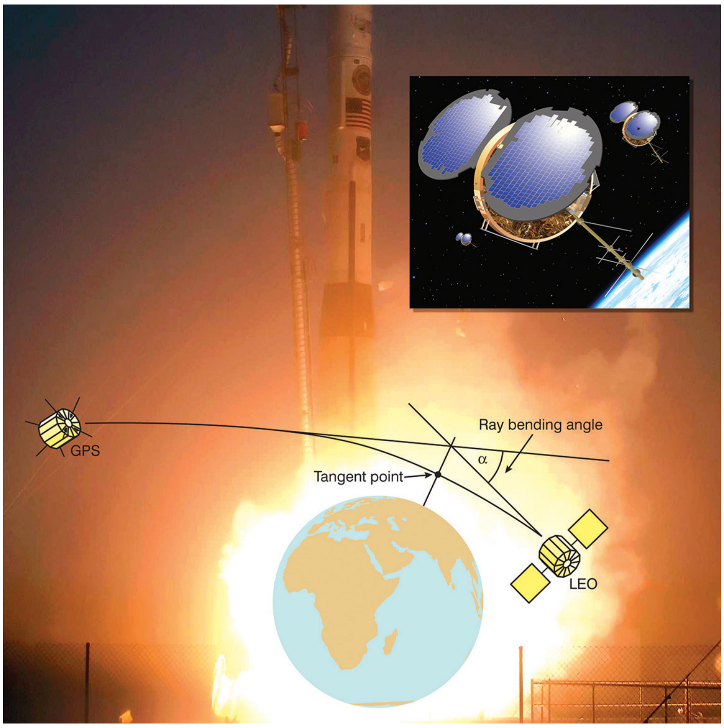
ECMWF Feature article

from Newsletter Number 111 – Spring 2007

METEOROLOGY

Ensemble streamflow forecasts over France



Courtesy of the Orbital Science Corporation.

www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/sbjdm2xu33

This article appeared in the Meteorolgy section of ECMWF Newsletter No. 111 – Spring 2007, pp. 21-27.

Ensemble streamflow forecasts over France

Fabienne Rousset Regimbeau, Florence Habets, Eric Martin, Joël Noilhan

In response to the increasing need for a better anticipation of severe hydrological events, the use of ensemble forecasts in hydrology is emerging as a key activity within the international scientific community. In 2004, a new international project was started, the Hydrological Ensemble Prediction Experiment (HEPEX), which brings together hydrological and meteorological communities to build a research project focused on advancing probabilistic hydrologic forecast techniques. This involves efforts on how to produce, disseminate, use and verify ensemble hydrological forecasts. Within the framework of HEPEX, inter-comparison experiments have been organized. Another European effort using hydrological ensemble forecasts is the development of the European Flood Alert System (EFAS) supported by the European Commission. The aim is to provide a harmonized Europe-wide long-term flood alert system in cooperation with national operational flood-forecasting centres. In addition to these international actions, the coupled hydrometeorological model SIM (SAFRAN-ISBA-MODCOU), developed for several years at Météo-France, is now the basis of on an ensemble streamflow forecast system.

The ensemble streamflow prediction system based on SIM has been running in real-time since 4 September 2004. It is forced by the ten-day meteorological ensemble forecasts from ECMWF. These forecasts are downscaled to the resolution of ISBA, which runs on a regular 8 km grid. A statistical analysis of the skill of the system was performed over nearly one year, using standard metrics for ensemble forecasts (e.g. Brier Score, Ranked Probability Score and Talagrand Histograms). It showed overall good results, for high flows as well as for low flows. In addition, a study of a few significant flood events from the past reveals that ensemble forecasts can be of great use to hydrological forecasters.

The SIM system

The SIM system is composed of a meteorological analysis system (SAFRAN), a land surface model (ISBA) and a hydrogeological model (MODCOU). It was first used in a forced mode in which the atmospheric forcing was derived from observations using SAFRAN. It has been validated over the long term for three large (regional scale) French basins and then for the whole of France. It has been shown that SIM is capable of reproducing the water and energy budgets, as well as the observed streamflows, aquifer levels and snow pack, particularly for basins with areas of over 1000 km2. In a recent study of the Seine basin, it was shown that the ability of SIM to reproduce the evolution of the aquifers and the partitioning between surface runoff and drainage leads to a good simulation of the main flood events of the Seine. Since 2003, the SAFRAN and ISBA modules have been used in an operational analysis mode at Météo-France, while the MODCOU module has been used in an experimental real-time analysis mode. This system provides a daily monitoring of the components of the water budget (such as soil moisture, snow pack, aquifer and river levels) over all France.

The real time ensemble streamflow prediction (ESP) system

The ensemble streamflow prediction (ESP) system developed at Météo-France is based on the SIM system. The idea is to use the ten-day ensemble forecasts from ECMWF instead of the SAFRAN analysis. An important goal of this ensemble prediction system is the ability to forecast the streamflow of large French basins, such as the Seine, and in particular the long-duration floods of these basins. The ensemble forecasts also provide valuable information during low flow periods, in which case the hydrographic system responses are strongly determined by the aquifers. The ten-day lead time of these ensemble forecasts is a major advantage: it provides information at a significant lead time. Such a lead time is not possible using the classical method, which consists in obtaining a several-day streamflow forecast from time extrapolation of the observed upstream flows.

AFFILIATIONS
Fabienne Rousset Regimbeau, Eric Martin, Joël Noilhan: CNRM Météo-France, 42 avenue Coriolis, 31057 Toulouse, France
Florence Habets: CNRS, UMR-SISYPHE, ENSMP, 35 rue St Honoré, 77305 Fontainebleau, France

doi:10.21957/sbjdm2xu33

The ten-day 51-member ensemble forecasts from ECMWF are used as an input of the ISBA and MODCOU modules, which produce 51 ten-day streamflow forecasts for about 900 river gauges over France (Figure 1). The initial conditions are provided by the real time SIM analysis. In particular, the soil wetness, the aquifer level and the amount of water stored in the rivers are initialised according to the real-time monitoring. The 00 UTC ECMWF ensemble forecasts are available daily in the Météo-France database, at a 1.5° resolution, with a 6-hour time step. In order to reduce the costs of calculation time and memory space, two adaptations were made.

- It was shown that the time step of ISBA could be increased from 5 minutes to 20 minutes without significantly changing the results (regarding the streamflows).
- It was determined that we only needed to use the ensemble forecasts of precipitation and temperature, while a climatology (based on the SAFRAN analysis, 1995–2003 mean) could be used for the other meteorological parameters used by ISBA.

A schematic of the SIM hydrometeorological coupled system and of the ensemble streamflow prediction system based on SIM is shown in Figure 2.

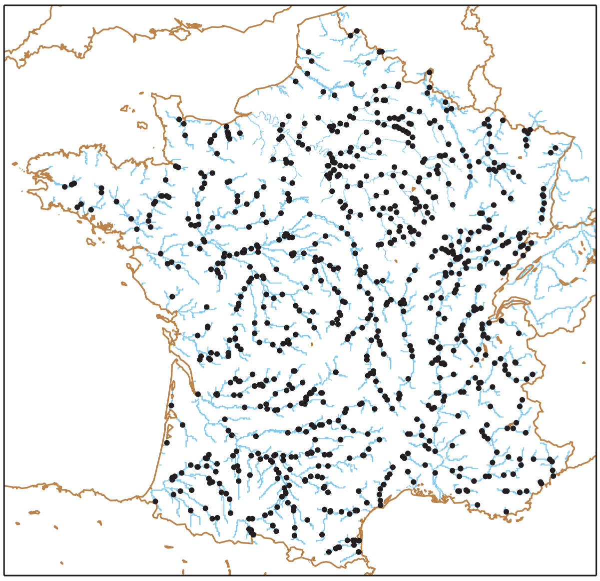


Figure 1 French hydrographic network used in SIM, and the position of the 900 river gauging stations over France.

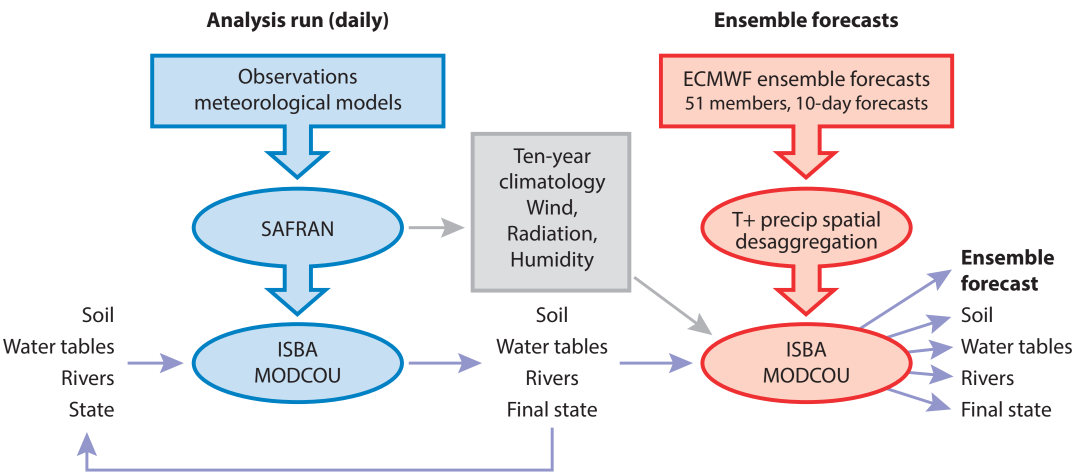


Figure 2 Schematic of the SIM hydrometeorological coupled system (left) and of the ensemble streamflow prediction system based on SIM (right).

Moreover, a spatial downscaling method was set up in order to adapt the 1.5° ensemble forecasts from ECMWF (temperature and precipitations) to the 8 km ISBA grid. We chose a simple method, based on that used by the SAFRAN analysis system, in order to maintain coherence between the operationally analysed SAFRAN and the ensemble forecasts. Figure 3 shows an example of this downscaling for one particular forecast period, with a map of the ECMWF data (Figure 3(a)) and a map of the data downscaled on the ISBA 8 km grid (Figure3(b)).

The downscaling is computed in two steps and uses an intermediate grid which consists of the SAFRAN zones (615 non regular zones, about 20×20 km each).

- Data from ECMWF is interpolated horizontally onto the SAFRAN zones using distance-dependent weights ($1/r^2$ interpolation).
- Altitude effects are then taken into account by imposing a vertical gradient within each zone which is computed as the difference between the ECMWF and ISBA grid-box average altitudes.

The vertical gradients are nearly the same as in SAFRAN: for the forecast temperature we use the ICAO gradient (-0.65 K/100 m), and for the forecast precipitation the value of the gradient is close to the one generally used by SAFRAN (0.7 mm year^{-1} m^{-1}). The gradient was calibrated over approximately one year (from 4 September 2004 to 31 July 2005). There is a good overall coherence between the accumulated precipitation fields forecast by the 24-hour ensemble mean (Figure 4(b)), and that analysed by SAFRAN (Figure 4(a)).

The ESP system has been running daily in real-time mode since 4 September 2004. It produces 51 ten-day streamflow forecasts for about 900 river gauges over France with a three-hour time step. Figure 5(a) shows an example of a graph that can be produced every day for each river station. This shows the ten-day streamflow forecast for the Seine at Paris on the 15 January 2006. The information on the forecast streamflow consists of a plot of the minimum and maximum values within the ensemble (as well as the interval including 80% of the forecast streamflow), the ensemble mean and the control (non-perturbed) run. The observed daily streamflow can be added afterwards.

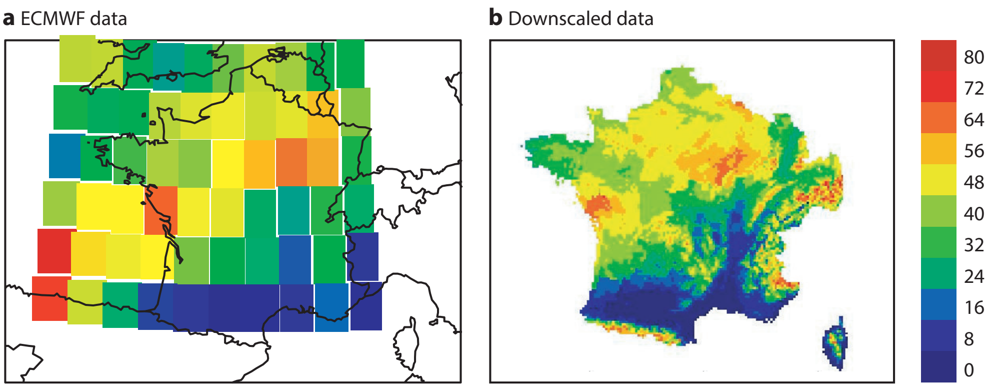


Figure 3 Result of the spatial downscaling of the ECMWF ten-day cumulative rainfall (mm), forecast for 17 October 2004, from the first ensemble member: (a) data from ECMWF on the 1.5° grid and (b) data after downscaling on the 8 km grid.

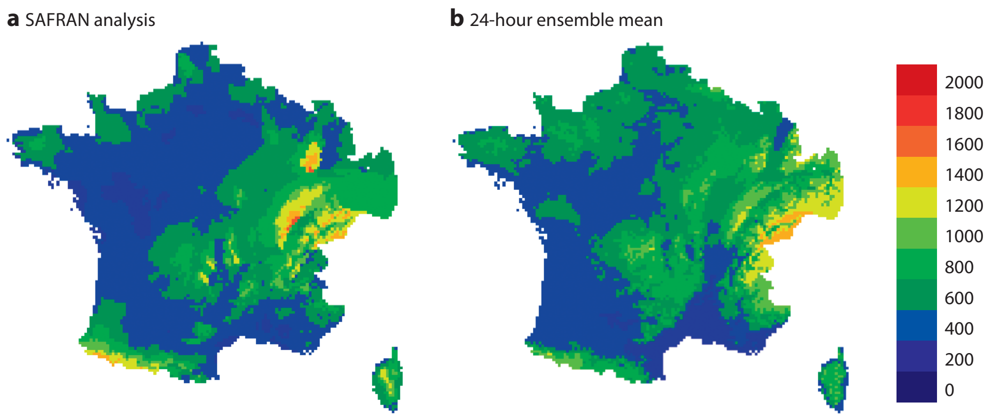


Figure 4 Result of the spatial downscaling of ECMWF cumulative rainfall (mm) from 4 September 2004 to 31 July 2005: (a) SAFRAN observation analysis and (b) 24-hour ensemble mean.

doi:10.21957/sbjdm2xu33

Figure 5(b) shows a different method of visualising the forecasts, which is similar to the products from EFAS. This chart shows the risk that the forecast streamflow will exceed (or remain below) a given threshold. Each line corresponds to one forecast day, and the lead times are shown in the columns. The percentage of ensemble members for which the forecast streamflow exceeds (or stays below) the threshold is indicated in each of the squares with a specific colour. For example, the chart in Figure 5(b) is for a high flow episode for the Seine river at the Paris gauging station using a threshold of 700 m3 s-1. This kind of chart allows a quick view of the forecast and provides an alert of a potentially dangerous event. It is of interest, in terms of the follow-up or evaluation of the forecasts, to see if the risk is confirmed from one forecast day to the next.

Figure 5(c) represents an example of the long-term behaviour of the system for the Seine river (at the Paris gauging station) from September 2004 to July 2005. It highlights the overall tendencies of the ensemble forecasts. In this example, the long-term tendency is well reproduced; however the forecast streamflow at the beginning of January 2005 is overestimated.

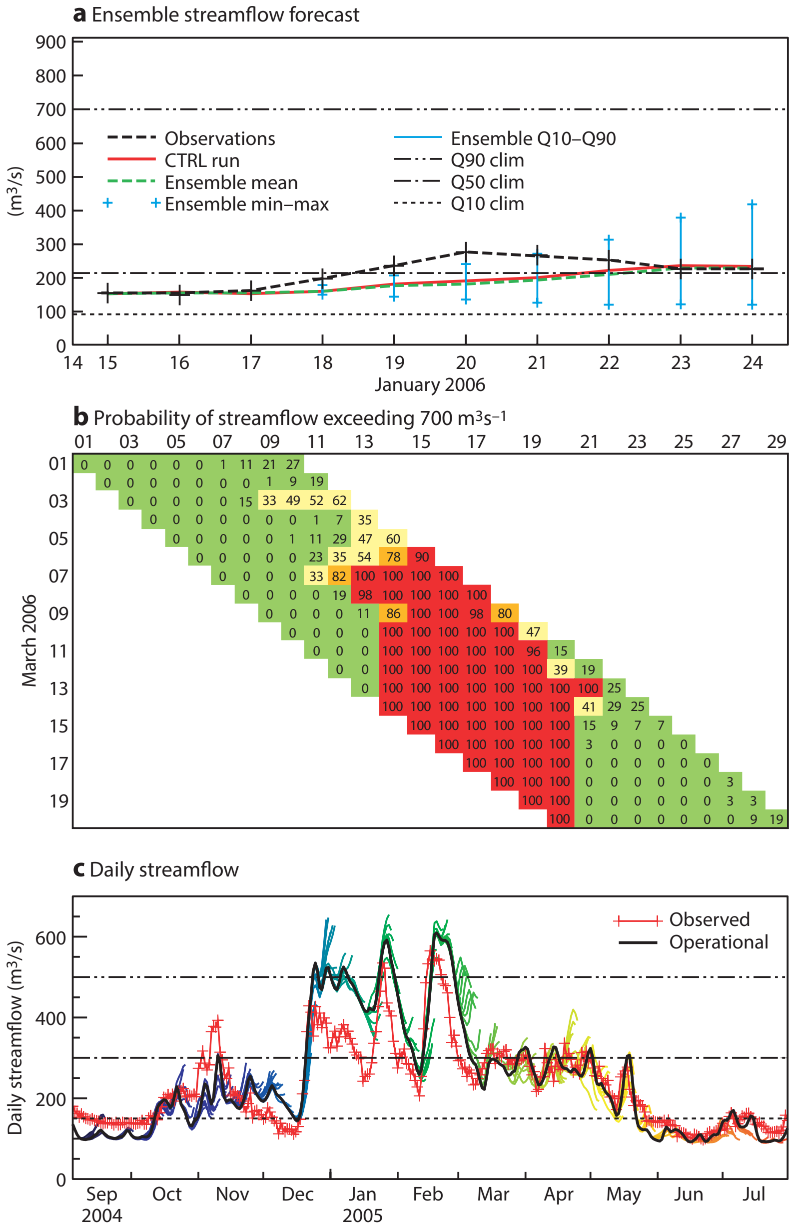


Figure 5 (a) Ensemble streamflow forecast for the Seine at the Paris gauging station, 15 January 2006. (b) Probability of the streamflow exceeding the 700 $m^3 s^{-1}$ level, for the Seine at Paris, forecasts from 1 to 20 March 2006. (c) Daily streamflow of the Seine at the Paris gauging station, from 4 September 2004 to 31 July 2005. Red line with crosses: observed streamflow. Black line: operational streamflow from the operational analysis run of SIM. Other coloured lines: ensemble mean of the ten-day forecast of streamflow starting every day.

Statistical analysis of the results of the ensemble streamflow forecast system

The probabilistic nature of ensemble forecasts implies a verification that is quite different from the verification of deterministic forecasts. As it is not possible to evaluate the quality of a single ensemble forecast against a reference, it is necessary to collect long time series of ensemble forecasts and observations, and to use statistical verification methods. We performed the statistical analysis of the ensemble streamflow prediction system over nearly one year, from 4 September 2004 to 31 July 2005.

Precipitation forecasts

First, we focused on the quality of a model input: the downscaled ensemble precipitation forecasts. To do so, we used the SAFRAN analysed precipitation as a reference. This preliminary study indicates that light precipitation events are slightly over-estimated a few days ahead, and the differences clearly become larger over a longer time period. In contrast, high precipitation events are underestimated, with the underestimation growing with the lead time. The spatial (over France) and temporal (over nearly one year) mean of the daily amount of precipitation ranges from 1.96 mm day^{-1} for the SAFRAN analysis, 2.11 mm day^{-1} for the one-day ensemble mean, and 2.40 mm day^{-1} for the ten-day ensemble mean. The mean square error of the ensemble mean (with the SAFRAN analysis being the reference) varies from 2.71 mm day^{-1} for one-day ahead to 3.92 mm day^{-1} for ten-days ahead. It is interesting to note that the mean square error is larger for the control run than for the ensemble mean, with 2.74 mm day^{-1} for the one-day forecasts and 5.04 mm day^{-1} for the ten-day forecasts. No annual cycle could be found in the error, probably because the 2004–2005 year is very dry, with a weaker precipitation annual cycle than for the other years.

The dispersion of the ensemble ranges from 0.53 mm day^{-1} (one-day forecast) to 3.36 mm day^{-1} (ten-day forecast). It seems quite weak for the first two or three forecast days, as these forecasts are really designed for medium-range forecasts. This fact is confirmed by the Talagrand Histograms (or rank histograms) which are clearly “U” shaped for the first few days of the forecast. This is less clear at the end of the lead time.

The Brier Score is a quite commonly used statistical metric; it is used to measure the performance of ensemble forecasts of dichotomous events (here, the daily precipitation amount exceeding a given threshold). The Brier Scores computed for thresholds ranging from 1 mm day^{-1} to 20 mm day^{-1} are quite good (i.e. near zero), with a slight increase (i.e. deterioration) with increasing lead time.

Finally, no basin appeared to have clearly better statistical scores, although there were generally slightly lower mean square errors for the northern-most basins (Seine, Loire, Brittany, Charente), and slightly larger errors for the mountainous basins in the south (Adour-Garonne, Rhône, South-East basins).

Streamflow forecasts

We also performed the statistical analysis for the streamflow forecasts. We focused on high streamflows as well as on low flow periods. As the study period was very dry over France, it provided a good opportunity to test the system for low flows and drought. It should be noted that information concerning low flows is just as important as that for floods for many industries as well as for the various French water management services. In this statistical study the forecast streamflows are compared to a reference run of SIM, and not directly to the observations. This allows the evaluation of only the quality of the ensemble without taking into account neither the hydrological model errors, nor the impact of the anthropogenic effects (especially the dam storage and outflow) that are more significant over a dry year. In the future, it is planned to compare ensemble streamflow forecasts in a more direct manner to the observations, once the improved hydrological-assimilation system based initialisation of SIM is available.

The Talagrand Histograms are generally similar in appearance for all of the river gauges, for example as seen in Figure 6 for the Seine river gauging station at Paris. The “U” shape is clear for the first few days of the forecast, which generally indicates a low ensemble dispersion (see Figure 6(a)). This seems to be mainly linked to the lack of dispersion of the ensemble precipitation for the one-day or two-day ECMWF forecasts. After three or four forecast days, the “U”shape vanishes and the histograms become rather flat, which indicates a good reliability. Sometimes there are particular character- of note: for the ten-day forecasts, the first rank is predominant which signifies a slight positive bias (see Figure 6(b)).

We computed the Brier Scores for different streamflow thresholds. We focused on high flows by defining the thresholds using the 90th and 50th percentile of daily streamflows from the whole existing climatology, and we also studied low flows by computing the Brier Scores for streamflow remaining below the 10th percentile from climatology. The results are generally good with low values of Brier Scores which slightly increase (i.e. deteriorate) with the lead time.

In order to evaluate the system compared to the climatology (the climatology being computed either with the observed streamflow or with the streamflow analysed by SIM from 1981 to 2004), we computed the Brier Skill Scores. When the Brier Skill Score is close to one, the ESP system provides better forecasts

doi:10.21957/sbjdm2xu33

than the climatology. For example, for the Seine river at the Paris gauging station, the Brier Skill Score for the threshold discharge of 300 $m^3 s^{-1}$ ranges from 0.99 (one-day forecast) to 0.8 (ten-day forecast). For all of France (with the climatology being the observed streamflows), the percentage of the 900 river gauges for which the Brier Skill Score is over 0.5 is very close to 100% on the first forecast day for the three thresholds. For ten-day forecasts, the percentages are 82% (>90th percentile), 92% (>50th percentile) and 98% (<10th percentile).

Figure 7 shows the repartition of the Brier Skill Scores as a function of the surface area of the catchments for one-, five- and ten-day forecasts. It reveals that the Brier Skill Scores are closer to unity (i.e. the best possible) for large basins, while the lowest Skill Scores are generally found for the smallest basins. Thus, the Brier Skill Scores tend to decrease (i.e. deteriorate) with increasing lead time. This shows that the system obtains good statistical scores for the majority of the river gauges, both for high as well as for low flow events.

Other statistical measures for evaluating the potential of the ensemble streamflow prediction system are the False Alarm Ratio and the Hit Rate. This kind of information may help ensemble prediction system end-users establish a decision-making system based on the ensemble forecasts, follow their costs and understand the consequences of false alarms or missed forecasts. The idea is to compute a contingency table for streamflow events which exceed a certain threshold. The event is considered as “forecast” when it is forecast by at least 50%, or 70%, or 90% of the ensemble members. Figure 8 shows the results for the Seine river at the Paris gauging station for the threshold discharge of 300 $m^3 s^{-1}$. The Hit Rate is quite high (over 0.75) in all of the cases. The False Alarm Rate stays under 0.2 at every lead time, and the ratio of cases which were not forecast stays below 0.35. It is notable that the absence of precipitation is generally well forecast by the ensemble, and consequently the low flows (< 10th percentile) and their duration is also well predicted, with in particular very low False Alarm Rates and good Hit Rates.

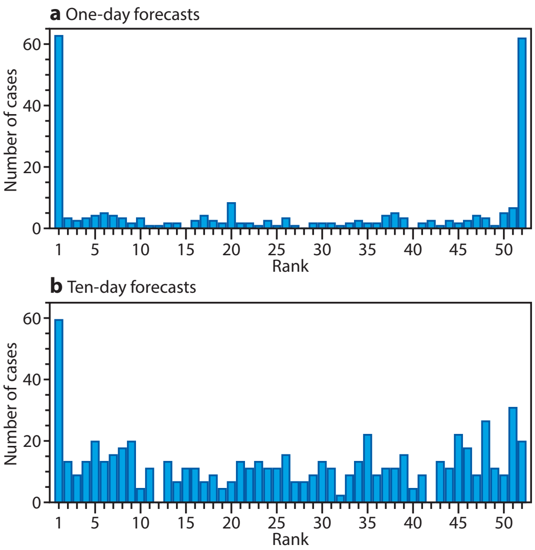


Figure 6 Talagrand Histograms for the Seine at the Paris gauging station, computed from 4 September 2004 to 31 July 2005; with the reference being the operational SIM streamflow: (a) one-day forecasts and (b) ten-day forecasts.

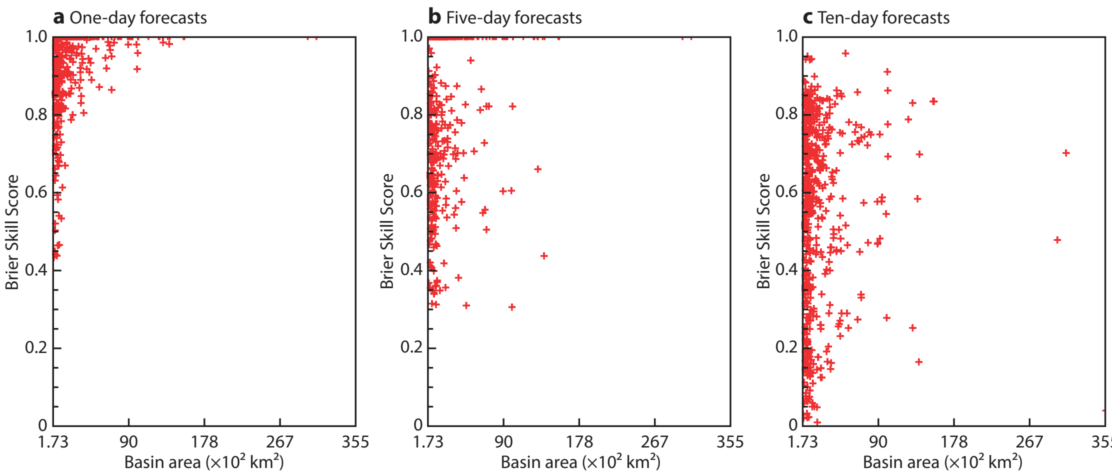


Figure 7 Brier Skill Score, computed from 4 September 2004 to 31 July 2005 using the climatology of 1981–2004 observed riverflows, for a streamflow exceeding the 90th percentile, for 900 river stations over France, following the basin area (km^2): (a) one-day forecasts, (b) five-day forecasts and (c) ten-day forecasts.

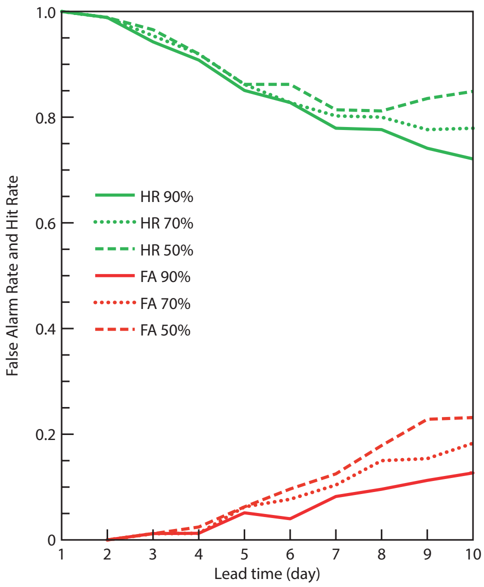


Figure 8 False Alarm Ratio (FA) and Hit Rate (HR) for the Seine at Paris, computed from 4 September 2004 to 31 July 2005 for a streamflow exceeding 300 m^3 s^{-1}, following the forecast lead time (from 1 to 10 days).

Focus on particular flood events

In addition to the statistical study, we focused on a few flood events in the recent past. Here, the results for the March 2001 Seine flood and the September 2006 flood in south-eastern France are presented.

The Seine basin is relatively large (44,000 km^2), with a slow hydrological response and a strong influence of the aquifers (the three main aquifers are explicitly simulated by MODCOU). Figure 9 shows the daily streamflow for the Seine at Paris from 5 March 2001 to 10 April 2001. This long duration flood is well analysed by SIM, as both the observed and the analysed streamflows are very close. The ensemble forecast seems meaningful, generally close to the observed or analysed streamflow, and the dispersion is neither too wide nor too narrow. The increase of the flow is well reproduced by the ensemble. The date of the flood peak also clearly appears in the forecasts, several days ahead, and it is coherent with both the observations and the analysis.

Figure 10 shows the ensemble forecasts for the Herault river at Gignac for the 2006 flood which occurred at the end of September (no observations are available at the current time for this river). This basin is quite small with an area of 1,400 km^2, and there is a very rapid time response to precipitation. Flash floods occur often following strong convective precipitation events. Figure 10(a) shows the forecast issued on 15 September (nine to ten days before the flood), and Figure 10(b) shows the forecast issued on 21 September (three to four days in advance). Figure 10(a) highlights the fact that even if the previous flood is not totally over (i.e. the streamflow is still decreasing), the ensemble forecast is able to give an early warning of the next event. Even if the flood peak is first forecast with an error of one or two days (Figure 10a)) and is underestimated (Figure 10(b)), the information given by the ensemble forecast can be of use for flood warning or water management agencies.

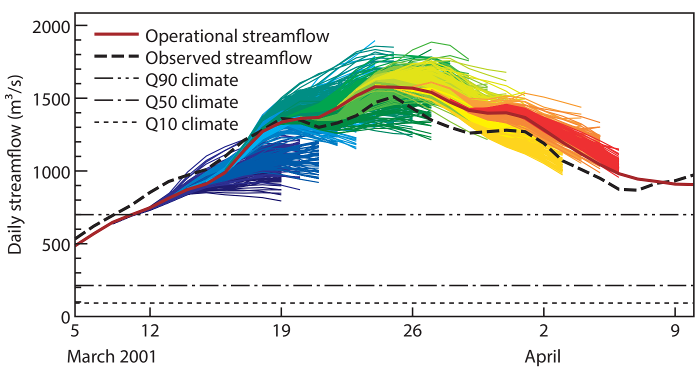


Figure 9 Daily streamflow of the Seine at Paris for the March 2001 ten-year flood. Black dashed line: observed streamflow. Red solid line: streamflow analysed by SIM. Coloured dotted lines: ensemble forecasts from 8 to 27 March 2006. Horizontal black lines: 10^{th}, 50^{th} and 90^{th} percentile of daily streamflow computed from 1974–2006 observed data.

 doi:10.21957/sbjdm2xu33

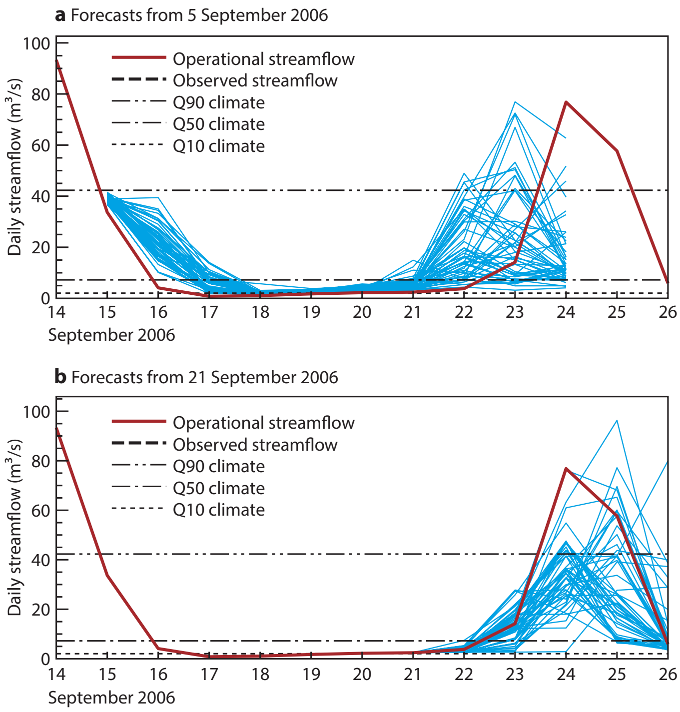


Figure 10 Daily streamflow for the Herault river at Gignac, for the end-September 2006 flood. Red solid line: streamflow analysed by SIM. Blue lines: ensemble forecasts from (a) 15 September 2006 and (b) 21 September 2006. Horizontal black lines: 10th, 50th and 90th percentile of daily streamflow computed from 1989–2006 observed data.

Further developments

The analysis of the downscaled ensemble precipitation forecasts showed good results overall; however there was a slight overestimation of light precipitation and an underestimation of heavy precipitation. Also, there was a lack of dispersion of the ensemble for the first few forecast days. Moreover, the statistical results for the ensemble streamflow forecasts are quite good, with satisfying values for the classical statistical scores (e.g. Brier Score, Ranked Probability Score and Talagrand Histograms). The system gave good results for both high and low flow (the study period being particularly dry) events. In addition consideration of a past flood events revealed that ensemble forecasts can be of great use to hydrological forecasters. This is true in particular for early flood warning for large catchments like the Seine which are characterized by relatively long-duration floods, as well as for smaller basins which have much faster hydrological responses.

In this preliminary study, no ensemble member was associated with perturbations of the physics of SIM or of the hydrological state. The streamflow dispersion was only due to the scatter in the precipitation forcing. Sampling uncertainties related to these hydrological elements is a possible method for improving the ESP system. Moreover, some improvements in the initialisation of SIM are necessary in order to more directly compare the ensemble forecasts and the observations. This is to be accomplished through the development of a hydrological assimilation system. The contribution of the short-range high-resolution meteorological ensemble forecast from Météo-France will also be examined, in particular for the forecasting of the flash floods for relatively small southern French basins.

Further reading

Rousset, F., F. Habets, E. Gomez, P. Le Moigne, S. Morel, J. Noilhan & **E. Ledoux,** 2004: Hydrometeorological modeling of the Seine basin using the SAFRAN-ISBA-MODCOU system. *J. Geophys. Res.,* **109,** D14105, doi:10.1029/2003JD004403.

Habets F., P. Le Moigne & **J. Noilhan,** 2004: On the utility of operational precipitation forecasts to served as input for streamflow forecasting. *J. Hydrol.,* **293,** 270–288.

HEPEX web site: http://hydis8.eng.uci.edu/hepex/

EFAS web site: http://efas.jrc.it/