We welcome the discussion on "Ensembles, uncertainty and flood prediction" by Dance and Zou (2010, hereafter indicated by DZ2010). The "Ensembles" workshop in Reading (UK) in September was well attended, interesting and led to many important discussions. The questions and scientific challenges presented in this paper were indeed formulated by a large body of participants at this workshop within a brainstorming environment. Many of the questions and challenges raised by DZ2010 are important, and have been the issue of an increasing number of research articles over the recent years/decade. Some of the questions raised here have been already answered and many are currently explored.

The context and discussions presented in DZ2010 would benefit from a closer survey of recent (and not so recent) publications in this area especially as this is a rapidly growing field and the workshop was ~9 month ago. See, for instance, the following papers and references therein: Sivapalan et al. (2003); Pappenberger and Beven (2006); Schaake et al. (2007); Thielen et al., 2008; Zappa et al. (2008); Cloke and Pappenberger (2009); Rotach et al., (2009); Schaake et al. (2010).

In addition, DZ2010 ignore several major national and international projects which deal exactly with the subject in question: flood prediction under uncertainty. In table 1 are some of the international initiatives in hydro-metereological modelling that would be key to the opinion paper by the authors (we acknowledge that there are also many other initiatives). On national level we would like to single out the NERC program STORMS (http://www.nerc.ac.uk/research/programmes/stormrisk/) as DZ2010 is
based on the outcome of a NERC workshop (NERC being the Natural Environment Research Council and a major funding body in the UK), which contains many of the challenges put forward by DZ2010 in its program objectives and deliverables (again there is also a large quantity of other national research initiatives).

These issues make the usefulness of the paper in its current form very limited.

Table 1: International initiatives in hydro-meteorological modelling that would be key to the opinion paper

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>COST 731</td>
<td>European COST Actions (<a href="http://www.cost.esf.org/about_cost">http://www.cost.esf.org/about_cost</a>) are a means to focus and co-ordinate existing research efforts supported by national funding agencies. By means of those actions scientists and students working in the same research field are connected each to other and can start collaborations. COST-731 is a network for scientists dealing with the propagation of uncertainty in end-to-end hydro-meteorological forecasting chains (Rossa et al., 2010). Three working groups (WG) deal with different aspects of this chain. WG-1 focuses on the propagation of uncertainty from observing systems (e.g. radars) into numerical weather prediction models (NWP). WG2: WG-2, co-ordinates research efforts on the propagation of uncertainty from observing systems and NWP into hydrological models (Zappa et al., 2010). WG-3 makes use of uncertainty information for issuing warnings and improving decision making (Bruen et al., 2010).</td>
</tr>
<tr>
<td>MAP D-PHASE</td>
<td>“MAP D-PHASE” is an acronym for Mesoscale Alpine Program Demonstration of Probabilistic Hydrological and Atmospheric Simulation of flood Events in the Alps (Rotach et al., 2009). The MAP D-PHASE initiative was an important element of the COST 731 Action, right from its initial planning (<a href="http://www.smr.arpa.emr.it/dphase-cost/">http://www.smr.arpa.emr.it/dphase-cost/</a>). This WWRP (World Weather Research Programme)-approved Forecast Demonstration</td>
</tr>
</tbody>
</table>
Project (FDP) D-PHASE was a follow-on project of the Mesoscale Alpine Programme (e.g. Bacchi and Ranzi, 2003) to demonstrate the societal impact of MAP by showcasing the progress achieved in high-resolution and probabilistic numerical weather prediction in complex terrain, along with the consequent benefits for hydrological forecasting. The heart of D-PHASE was a distributed end-to-end forecasting system geared to Alpine flood events which was set up to demonstrate the state-of-the-art in forecasting precipitation-related high-impact weather. A first insight into MAP D-PHASE with a focus on operational ensemble hydrological simulations is presented in Zappa et al. (2008) and Ranzi et al. (2009).

| HEPEX | The Hydrological Ensemble Prediction Experiment (HEPEX) was launched as a bottom-up process by scientists and users at an ECMWF workshop in 2004 (http://www.hepex.org). This international research activity is designed to address questions related to end-to-end forecast systems in order to build useful systems and to promote their rapid development and deployment. Schaake et al. (2007) present some of the key scientific questions associated with the major components of a probabilistic hydrological forecast system, including calibration and downscaling of ensemble weather and climate forecasts, hydrological data assimilation, and user issues. Additional science questions were defined at the third HEPEX workshop held in Stresa in June 2007 (Thielen et al., 2008; special issue: http://www3.interscience.wiley.com/journal/119817000/issue). Approximately, ten site specific testbeds, as well as four multidisciplinary testbeds have been activated, focussing on one or more clearly defined HEPEX science questions. These have the potential to develop data resources needed for community experiments to address all of the scientific questions, and are expected to include active user participation. A special Joint HEPEX/COST731 workshop on downscaling NWP products and propagation of uncertainty in hydrological modelling was held in Toulouse, June 2009, and several studies are available from the HEPEX website. |
The IAHS Decade on Predictions in Ungauged Basins (PUB) 2003-2012 is an initiative launched by the International Association of Hydrological Sciences (IAHS) in 2002, aimed at formulating and implementing appropriate science programmes to engage and energize the scientific community, in a coordinated manner, towards achieving major advances in the capacity to make predictions in ungauged basins(http://pub.iwmi.org/UI/Content/Default.aspx?PGID=0). The PUB scientific programme focuses on the estimation of predictive uncertainty, and its subsequent reduction, as its central theme. A general hydrological prediction system contains three components: (a) a model that describes the key processes of interest, (b) a set of parameters that represent those landscape properties that govern critical processes, and (c) appropriate meteorological inputs (where needed) that drive the basin response. (see Sivapalan et al. 2003 for more details).

The European Flood Alert System, its development, evaluation and communication products are described by Ramos et al. (2007), Thielen et al. (2009) and Bartholmes et al., (2009). It grew from a European Union funded research project and now provides flood warnings from 3 to 10 days in advance for large transnational river basins in Europe. It uses weather forecasts from the European Centre for Medium Range Weather Forecasts and from the German Weather Service as inputs to the LISFLOOD distributed hydrological model (De Roo et al. 2000) operating at a 5 km grid scale. Ramos et al (2007) reports that users of the system found the combination of deterministic and probabilistic forecasts, included on the charts, very useful and also appreciated the use of colour-coded threshold exceedences.

A lot of progress has been made by the scientific studies reported within Table 1 and elsewhere, some of which is cited in this commentary. Several claims pointed out by DZ2010 have been discussed and treated by the hydrologic community, even though challenges still remain in flood prediction and hydrologic ensemble forecasting. It is
important new achievements in these fields be put in perspective when published
With this in mind let’s examine some of the claims in detail

Uncertainty in initial conditions, boundary conditions and forcing data, Parameter Errors and Model structural errors
Sections 3 to 5 of DZ2010 address how the various sources of uncertainty in the models used can be represented and accounted for. This has been an area of hydrological research for decades (e.g. Beven and Binley, 1992, Molteni et al., 1996; Zhang and Lindstrom, 1997, Beven and Freer, 2001, Guinot and Gourbesville, 2003, Pappenberger and Beven, 2006; Mattot et al. 2009).
A large variety of methods exist. Many, but by no means all (e.g. Gupta et al., 1998, Bardossy et al., 2008), are Bayesian methodologies looking to condition the parameter space of the hydrological model (and possibly the parameters of the error distribution). The differences between the methods are in the degree of separation of, and level of belief shown about, the characteristics of the various sources of uncertainty.
For example consider the simplest regression framework where all the error is characterised by an additive term representing the difference between the model predictions and observations. The “levels of belief” shown about this additive term range from the Informal Likelihoods utilised in implementations of the Generalized Likelihood Uncertainty Estimation (GLUE) methodology (Beven and Binley 1992) to formal statistical likelihoods (e.g., Vrugt et al., 2003).
Other work has attempted to separate the sources of uncertainty. For example an error-in-variables framework can be utilised to estimate the parameters of a distribution conceptualizing the error in meteorological observations (e.g., Kavetski et al., 2002; Vrugt et al., 2008, Kavetski et al., 2006; Salamon and Feyen, 2009), or adding an additional stochastic terms to represent deviations in the model response from reality (presuming the remaining observational errors are accounted for separately). In these later cases the separation of the sources of uncertainty can only be achieved by placing informative prior beliefs on the relationship between the observed data and the unknown inputs and outputs of the model. DZ2010 are right to highlight this as an area where closer interaction between meteorologists and
hydrologists may lead to benefits, for example in the characterisation of the uncertainties associated with evapotranspiration.

It is also important to acknowledge the uncertainty due to the model structure, which requires using multiple model structures. The use of multi-model forecasts as described by DZ2010 is one approach to attempt the quantification of model structural errors - although the assumptions that multiple NWP (or hydrological) models are largely independent and encompass all possible structural uncertainties is questionable. In hydrology multi model approaches have proven to be a promising tool in forecasting. The TIGGE achieve does (http://tigge-portal.ecmwf.int/, http://tigge.ucar.edu/, http://wisportal.cma.gov.cn/tigge/) present an excellent opportunity for the hydrological community to investigate such approaches by having multiple forcing conditions. Results in this area have been presented by He et al., 2010 and 2009, Pappenberger et al, 2008). However, recent results by Hagedorn et al (2008a,b) indicate that a similar performance may be achieved by downscaling or calibration of weather forecasts through re-forecasts (see discussion on propagation of uncertainty below).

In hydrology several multi-model forecasting approaches have been presented for example so called Bayesian model averaging (Ajami et al., 2007; Vrugt and Robinson, 2007). Here, Bayes theorem is applied to weight structurally different rainfall-runoff models depending on there past forecasting capabilities. This is readily coupled with the Bayesian methodologies for representing the input and parameter uncertainties outlined above.

Techniques which do not directly focus upon conditioning the model to represent prediction error have also been utilised. For example various post-processing methods have been developed to represent the error between a point prediction (single value deterministic forecast) of the observed runoff at a given time with a view to estimating the total predictive uncertainty (Krzysztofowicz, 1999, Bogner and Kalas, 2008).

In Section 3 ZP2010 discuss the use for data assimilation to improve forecasts. This has been utilised in hydrology to improve predictions of discharge and water level in flood model (Young, 2002, Schumann et al., 2009, Madsen, 2005). As highlighted the implementation of data assimilation algorithms raises a number of issues.
Ensemble validation or verification

We fully agree with the authors that “greater cooperation between meteorologists and hydrologists is still needed to ensure that the new skill scores are fit for purpose.” (see also discussion in Pappenberger et al., 2008). Although we cannot falsify the statement that the “knowledge of properties of existing skill scores used routinely in NWP is not yet widespread in the flood prediction community”, one has to acknowledge a large body of publications in this area (e.g. Renner et al., 2010, Brown et al., 2010, Jaun and Ahrens, 2009, Velazquez et al., 2009, Olsson and Lindström, 2008, Thirel et al., 2008, Laio and Tamea, 2007, Roulin, 2007 to name just a few!).

The scientific questions raised by DZ2010 are very important and also build the foundation of the cross-cutting verification testbed of HEPEX (see http://www.hepex.org for more details and Brown et al, 2010). It also extends on these scientific challenges and asks for example: for methods which are appropriate for multivariate forecast; to propose methods to characterize timing error, peak error and shape error in hydrologic forecasts; to define an optimal set of benchmark; and to understand how to account for correlations in predictors and forcing variable. HEPEX is a community initiative and any reader is invited to contribute to these important science questions.

Propagation of uncertainty between models

We wholeheartedly agree with the authors that the modelling chain offers us a new and exciting opportunity to evaluate ensemble forecasts from a user perspective. Hydrological catchments are meaningful integrators over space, time and multiple forecast variables, in particular in contrast to meteorological skill evaluation of a single variable of a fixed domain with a fixed accumulation step.

Propagation of uncertainty and cascading the uncertainty is a key scientific challenge in any forecasting chain. Part of this uncertainty propagation is the post/pre-processing (downscaling) of ensembles and the authors can only be applauded for pointing out that this interconnection may have a significant impact on the forecasting cascade. A comprehensive review of various techniques is presented in Maraun et al (2010). Schaaake et al (2010) summarizes the recommendations of the first workshop on Postprocessing and Downscaling Atmospheric Forecasts for Hydrologic Applications held at Météo-France, Toulouse, France, 15-18 June 2009. They for
example formulate requirements for post-processing and downscaling highlighting multiple points such as the need to extract as much relevant information as possible from the weather and climate forecasts on different spatial and temporal scales. They further discuss the role of dynamic versus statistical downscaling and point out that the meteorological and hydrological community should rigorously compare statistical downscaling against dynamical downscaling and multi-model ensemble approaches. They reiterate the Requirements for re-forecasts of NWP models for optimal integration into a decision making framework (e.g. Fundel et al., 2010). In addition atmospheric modelling issues are stated for the example that it is important to find a good trade off between NWP model resolution and the number of members in the ensemble. It is not possible to give a full list of all results of this workshop and the interested reader is encouraged to read Schaake et al (2010) and the papers in this special issue.

DZ2010 point out that often only a one way coupling between the hydrological and meteorological model exists. In parts, such a decoupling may be justified by the focus of a hydrological model (e.g. forecasting). However, all meteorological forecasting systems also have a land surface scheme and ocean. One can ask whether a more active engagement by the user community of meteorological data in the design of these intrinsically coupled models would not be beneficial and thus address some of the important recommendations in this opinion paper such as the ensemble design.

Communicating uncertainty to users
Communication is certainly one of the key issues of any forecasting chain as any forecast (including any improvement) will be useless unless it can and is communicated. Visualisation is clearly an important part of it and Bruen (2010) as well as Cloke et al. (2009) summarize some of the current practise. We are certainly still a long way off answering the key scientific question on “how we can best present and visualize uncertain information” (DZ2010, p 3606), maybe because there is no universal answer. As the research by Norbert et al (2010) conclude that effective training and clear communication of ensemble prediction systems, while clearly necessary, are by no means sufficient to ensure effective use of ensemble prediction systems. Attention must also be given to overcoming the institutional obstacles to their use and to identifying operational choices for which ensemble prediction systems...
is seen to add value rather than uncertainty to operational decision making. Only then will the communication of uncertainty in hydro-meteorological forecasts not be a mission impossible (see papers by Ramos et al, 2010 and Demeritt et al., 2010 and special issue on Communicating weather information and impacts in Meteorological Applications, 2010 17(2))

Final Remarks
Parts of this answer are excerpts from the publications we quote, shortened and/or adapted to fit as an answer to DZ2010.

We recommend any potential reader (and the authors) who is interested in this topic to browse the references we propose, starting from Schaake et al. (2008); Rotach et al. (2009); and Cloke and Pappenberger (2009) and visit http://ensemble.nmpi.net/index.php?title=Main_Page). We also recommend any interested reader to join the international HEPEX (http://www.hepex.org) in which many of the questions are embedded into individual test beds.

We would like to make it clear that we welcome all discussion on this topic and believe that this science field contains many unanswered scientific questions. The concept of organizing international and national workshops around this topic can only be applauded as it raises awareness and will advance this scientific field. Internationally, HEPEX organizes several workshops (see webpage http://www.hepex.org) on specific as well as general topics in this field as well as various other organisations e.g. the International Commission on the Hydrology of the Rhine Basin (CHR, 2010 and 2006). The FREE workshop was an excellent example of a national workshop and we hope that more such initiatives are undertaken in the future and reported in the international literature.
References


He, Y., Cloke, H.L., Wetterhall, F., Pappenberger F., Freer, J., Wilson, M., 2009, Tracking the uncertainty in flood alerts driven by grand ensemble weather predictions, *Meteorological Applications*, 16(1), 91-10


doi:<http://dx.medra.org/10.1051/lhb/2009061>10.1051/lhb/2009061


Gupta, H. V.; Sorooshian, S. and Yapo, P. O. 1989, Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information Water Resources Research, 34, 751-763


Young, P. C. 2002, Advances in real-time flood forecasting Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences. 360, 1433-1450