University of
Reading

Ensemble flood forecasting: current status and future opportunities

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1002/wat2.1432

Publisher: Wiley

Ensemble flood forecasting: current status and future opportunities

Wenyan Wu[1], Rebecca Emerton[2], Qingyun Duan[3], Andrew W. Wood[4], Fredrik Wetterhall[5], David E Robertson[6]

(1) University of Melbourne, Australia, wenyan.wu@unimelb.edu.au, https://orcid.org/0000-0003-3907-1570 (2) National Centre for Atmospheric Science, University of Reading, UK, rebecca.emerton@reading.ac.uk, https://orcid.org/0000-0003-0707-3993 (3) Hohai University, China, qyduan@hhu.edu.cn, https://orcid.org/0000-0003-0707-3993 (4) National Center for Atmospheric Research, USA, andywood@ucar.edu, https://orcid.org/0000-0002-6231-0085 (5) European Centre for Medium-Range Weather Forecasts, UK, fredrik.wetterhall@ecmwf.int, https://orcid.org/0000-0001-5331-9064 (6) CSIRO Land and Water, Australia, david.robertson@csiro.au, https://orcid.org/0000-0003-4230-8006

Abstract

Ensemble flood forecasting has gained significant momentum over the past decade due to the growth of ensemble numerical weather and climate prediction, expansion in high performance computing, growing interest in shifting from deterministic to risk-based decision-making that accounts for forecast uncertainty, and the efforts of communities such as the international Hydrologic Ensemble Prediction Experiment (HEPEX), which focuses on advancing relevant ensemble forecasting capabilities and fostering its adoption. With this shift, comes the need to understand the current state of ensemble flood forecasting, in order to provide insights into current capabilities and areas for improvement, thus identifying future research opportunities to allow for better allocation of research resources. In this paper, we provide an overview of current research activities in ensemble flood forecasting and discuss knowledge gaps and future research opportunities, based on a review of 70 papers focussing on various aspects of ensemble flood forecasting around the globe. Future research directions include opportunities to improve technical aspects of ensemble flood forecasting, such as data assimilation techniques and methods to account for more sources of uncertainty, and developing ensemble forecasts for more variables, for example flood inundation, by applying techniques such as machine learning. Further to this, we conclude that there is a need to not only improve technical aspects of flood forecasting, but also to bridge the gap between scientific research and hydro-meteorological model development, and real-world flood management using probabilistic ensemble forecasts, especially through effective communication.

Origins of reviewed ensemble flood forecasting studies
2001-2019
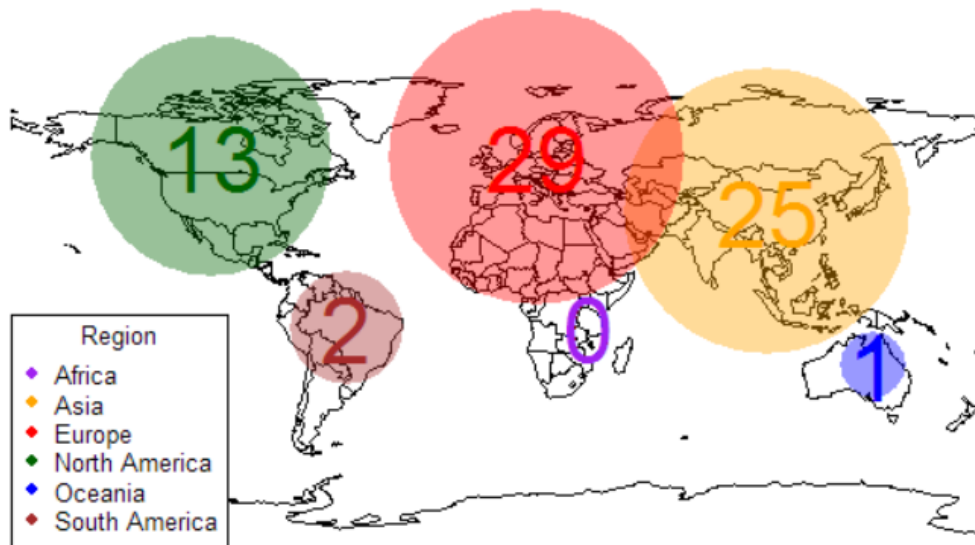
INTRODUCTION

Background information

Flood modelling and forecasting is key to managing and preparing for extreme flood events. According to the international disaster database, flooding occurs more frequently than all other types of natural hazards across the globe, and accounts for 39% of all disasters arising from natural hazards since 2000, with >94 million people affected worldwide every year (Guha-Sapir et al., 2018). The World Meteorological Organisation (WMO) estimates that while economic losses caused by flooding have increased over the past fifty years, loss of life has significantly decreased due to improved monitoring and forecasting of hydro-meteorological hazards (WMO, 2015).

There exists a wide range of techniques used for forecasting flood events, the majority of which use precipitation data as input, be this from upstream observations, or precipitation forecasts from numerical weather prediction (NWP) systems or rain field extrapolations. Precipitation and temperature forecasts are key in predicting river flow and flooding , particularly for lead times of more than 1-2 days (Cloke and Pappenberger, 2009). Many of the flood forecasting approaches, however, provide a deterministic forecast, i.e. a single forecast scenario of the future state, with little or no defensible estimate of the uncertainty in the forecast. Over the past 2 – 3 decades, meteorological NWP has moved from deterministic forecasting to probabilistic forecasting using ensemble prediction systems (EPS), and in the past decade, ensemble forecasting has also become increasingly available for hydro-meteorological flood forecasting purposes across the globe (Emerton et al., 2016; Pappenberger et al., 2016). Ensemble flood forecasts can provide an indication of the potential of upcoming extreme events, and the risk of damages associated with them. Despite a range of scientific and institutional challenges associated with moving to operational ensemble forecasting for flood management (Arnal et al., 2019), the benefits of ensemble forecasts over deterministic forecasts have been argued by a number of recent

analyses (Rossa et al., 2011; Siddique et al. 2017; Verkada and Werner, 2011; Pappenberger et al., 2015).

Ensemble Forecasting

In hydrological forecasting, the four key sources of uncertainty are the input data, model structure, model parameters and evaluation data for model validation (Kauffeldt, A., 2014; Shaw et al., 2011; Wood and Lettenmaier, 2008; Pagano et al., 2013), and their superposition (Zappa et al., 2011). For forecasts beyond 2-3 days ahead, the largest uncertainty stems from the meteorological forecast input (He et al., 2009; Emerton et al., 2016) (except for special circumstances, such as seasonal forecast involving snow melting). Ensemble forecasts allow for representation of the forecast uncertainty by producing a range (or ensemble) of possible forecast outcomes, each of which is equally probable. This is done by applying small perturbations to the initial state of the numerical weather prediction (NWP) model. Due to the chaotic nature of the atmosphere (Lorenz 1963), these small perturbations lead to significantly varying outcomes with increasing lead time. In addition, many forecasting centres include stochastic physics in their EPS, applying perturbations to the parameterised physical processes, to additionally represent the uncertainty associated with these parameterisations (Buizza et al., 1999).

In flood forecasting, probabilistic river flow predictions are typically produced by using an ensemble precipitation forecast to force a hydrological model (Cloke and Pappenberger, 2009; Pagano et al., 2013; Clark and Hay, 2004; Cloke et al., 2013). The ensemble nature of the forecast provides usable probabilistic information about flood events on much longer lead times than has previously been possible for deterministic forecasts (Emerton et al, 2016), and is useful in catchments where no other input data is available (Cloke and Pappenberger, 2009). Even so, many the latest operational ensemble flood forecasting systems do not additionally account for uncertainty in the hydrological model; as such, the spread of the forecasts is typically under-dispersive (Wang et al., 2009), not accounting for the full range of uncertainty throughout the modelling chain. This remains a grand challenge in ensemble flood forecasting and an active area of current research.

Cloke and Pappenberger (2009) investigated the scientific and technical drivers behind the shift from deterministic to ensemble flood forecasting, and highlighted seven challenges in this shift, from computing resources and limited sample sizes of extreme events to communicating uncertainty. As ensemble prediction systems have continued to increase in both number and quality, and decision-making for flood events has begun to move from deterministic to probabilistic (Hall et al., 2008; Bruen et al., 2010), we discuss current research activities in the area of ensemble flood forecasting and their evolution over the past decade. We further highlight current knowledge gaps and opportunities for future research in order to improve ensemble flood forecasting and its use in water management.

Aims and organisation of paper

This paper examines the development trajectory of ensemble flood forecasting and its current research status based on evidence in literature and discusses future research opportunities from a "flood" point of view. It will also serve as an introduction of ensemble techniques for flood forecasting, promoting the uptake of ensemble flood forecasting by flood practitioners.

In total, 70 papers up to the year 2019 were identified using keyword search on web of science. The first paper appeared in year 2001. The keyword was specifically selected to include "flood" in the title to

include studies related to flood management as opposed to only focusing on studies on hydro-meteorological forecast, where ensemble forecasting has been a research focus for many years. The review process, including paper selection process and criteria used to review the papers is described in Appendix A. The list of the papers selected, and the information of the case study system and data used are included in Appendix B.

REVIEW OUTCOMES

BOX 1: Overview of selected papers

The publication years of the selected papers are summarised in Figure B1 and the origins of these studies are summarised in Figure B2. There is an remarkable increase in publications in the research field over the past decade. The first publication fitting the search criteria appeared in 2001. In the following five years, there was only one publication in 2003. Since 2007 however, the number of papers published started to increase. Although varied across the years, the general increasing trend in the number of publications is evident. This observation agrees with the increasing awareness of flood risk and the needs to consider uncertainty in flood forecasting in recent years (Cloke and Pappenberger, 2009; Rossa et al., 2011) and the fact that many countries started to provide global ensemble NWP products around that time (Pappenberger et al., 2016). Despite the overall increase in the number of studies over the years, most of these studies were conducted in countries located in the northern hemisphere or in the developed world. European countries account for nearly half of the papers reviewed, followed by Asia and North America; while there is no study from Africa. This is generally in line with the origins of NWP ensemble products, which are the ultimate input for ensemble flood forecasting models.
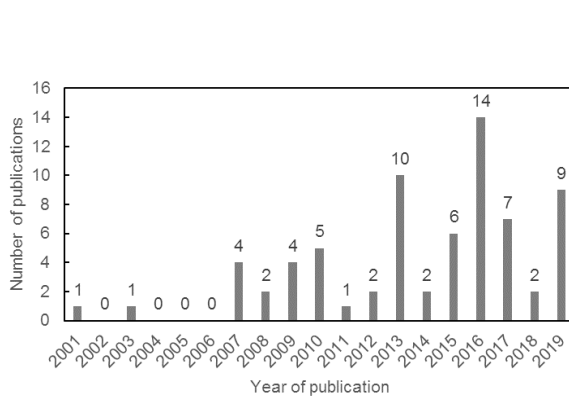
Figure B1 Number of publications in each year   Figure B2 Origins of selected studies
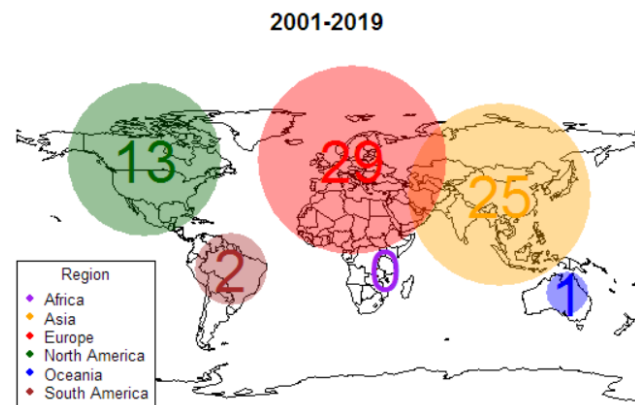
---

Study focus

The study focus of the papers was reviewed and the papers were categorised into those focusing on ensemble modelling, ensemble utilisation, and ensemble application, as shown in Figure 1. Over half of the studies focus on the application of ensemble input for flood modelling (using either climate or hydrological models), including the development of ensemble flood models, the comparison of different

ensemble products and feasibility studies examining the suitability of specific EPS to be used at specific locations of interest. This is especially the case for papers published during the early half of the review period (i.e. prior to 2010), when there are only 2 papers on technical topics related to ensemble modelling (i.e. data assimilation) and 3 papers on the application of ensemble forecasts to solve real-world problems (i.e. developing flood warning systems). During the second half of the review period, the overall number of papers in the field has increased significantly, especially on ensemble modelling and application. During this period, papers focusing on topics related to ensemble modelling have extended to cover ensemble generation and post-processing of ensemble forecasts; while the number of papers focusing on ensemble flood forecasting for the development of flood warning systems has increased from 3 to 10, and the application topics have been extended to cover reservoir operation and flood risk analysis.
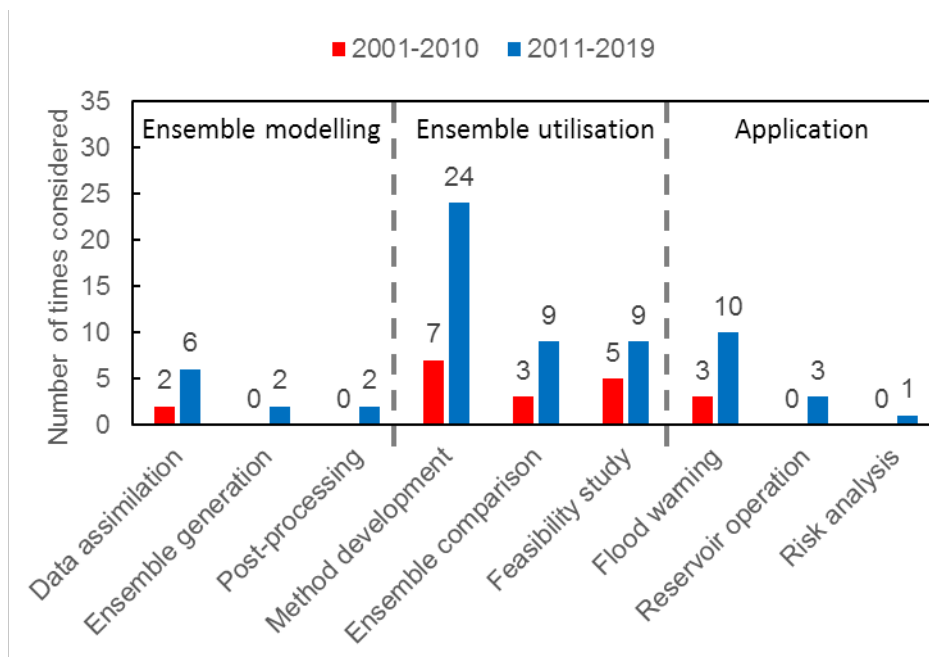
Figure 1 Study focus of papers selected

Cases study and datasets

The case study systems and datasets used in the papers reviewed were examined to understand the origins of the case studies (Figure 2), the types of flood simulated (Figure 3), and the spatial scale (Figure 4), temporal resolution (Figure 5) and lead time (Figure 6) of these flood events.

As can be seen in Figure 2, prior to 2010 most of the case study systems are from the northern hemisphere, including 9 (or 53%) from Europe and 6 (35%) from Asia. There is one case study system (6%) each from Africa and Oceania during this period. Although still dominated by the northern hemisphere, the geographic distribution of the case study systems changed in the second half of the study period, with Asia leading with a total of 22 systems (41%) being investigated, followed by 17 case

studies (31%) from Europe and 11 (20%) from North America. Further to this, there are also case studies from South America, and a global study.

2001-2010

2011-2019
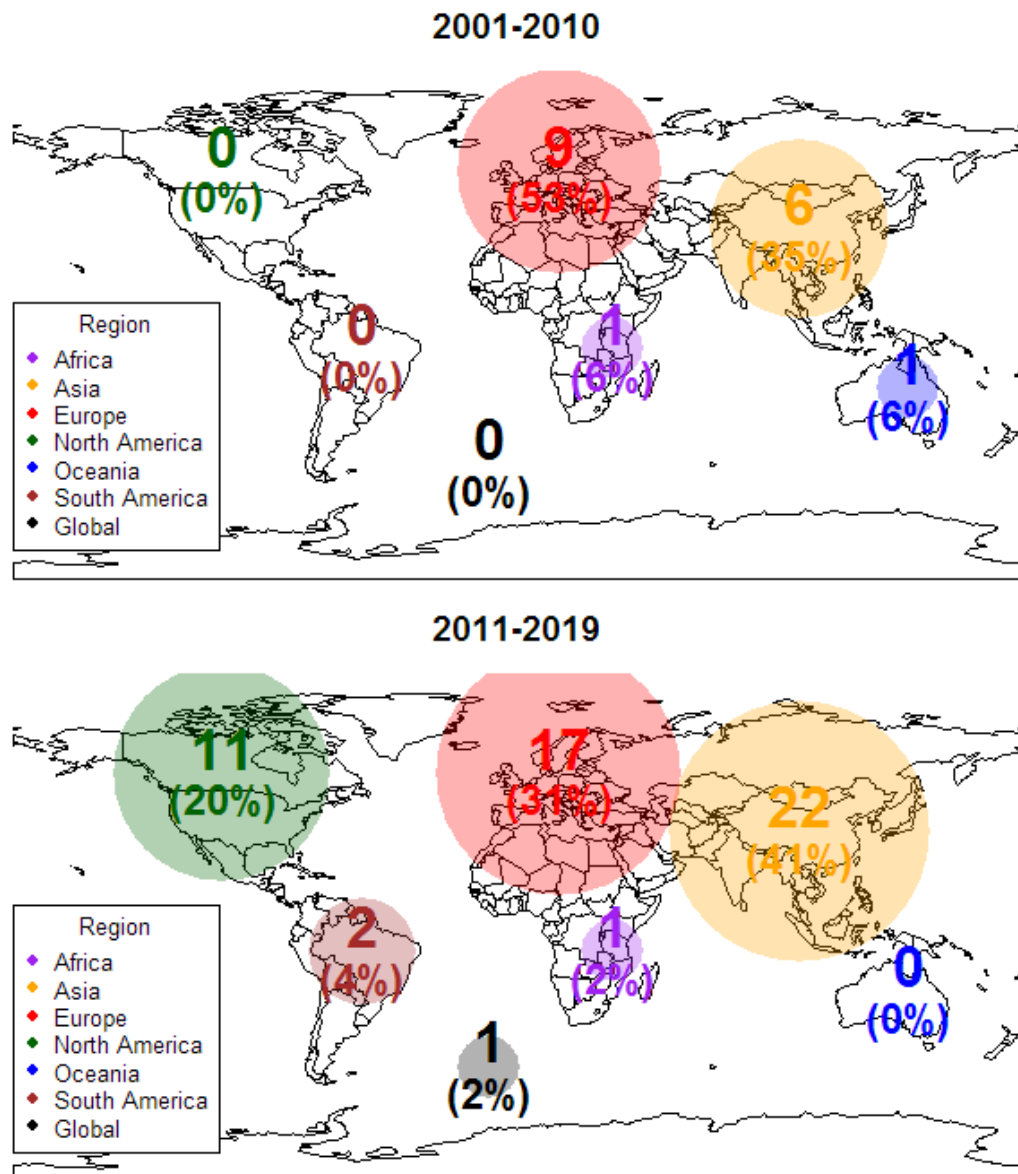
Figure 2 Locations of case studies investigated. The whole numbers represent the number of case study systems investigated in each region during each time period. The percentages in parenthesis represent the percentage of cases study systems in each region over the total number of case study systems considered during each time period.

During the early stages of research on ensemble flood forecasting, most of the studies focused on floods in rural catchments, as shown in Figure 3, with rural flood events accounting for 94% of those investigated between 2001 and 2010. More recently, the scope of ensemble flood forecasting has extended to include other types of floods, such as flooding in urban areas (an increase from one case

study between 2001 and 2010 to 11 between 2011 and 2019), flash floods (an increase from zero to 14) and compound floods (defined as a flood caused by a combination of events, such as tidal flooding and intense rainfall") (an increase from zero to 2). The increase in the flash flood studies is likely to be related to a special issue on the topic published in 2016 (Braud et al., 2016)

The change in the types of flood investigated also resulted in changes in the spatial scale of flooding investigated. As shown in Figure 4, although floods occurring on catchment or basin scale are the focus in both time periods, floods on a local scale have become more popular in the past decade. This is mainly due to increased research on urban floods, which affect a relatively small area. It should also be noted that one of the studies on catchment/basin-scale floods during the second half of the review period investigated river basins from around the world, (Figure 2). In addition, compound floods were not considered until more recently, beginning in 2017, due to increased awareness of compound events in recent years (Leonard et al., 2014; Zscheischler et al., 2018); this is also an emerging area for flood modelling.
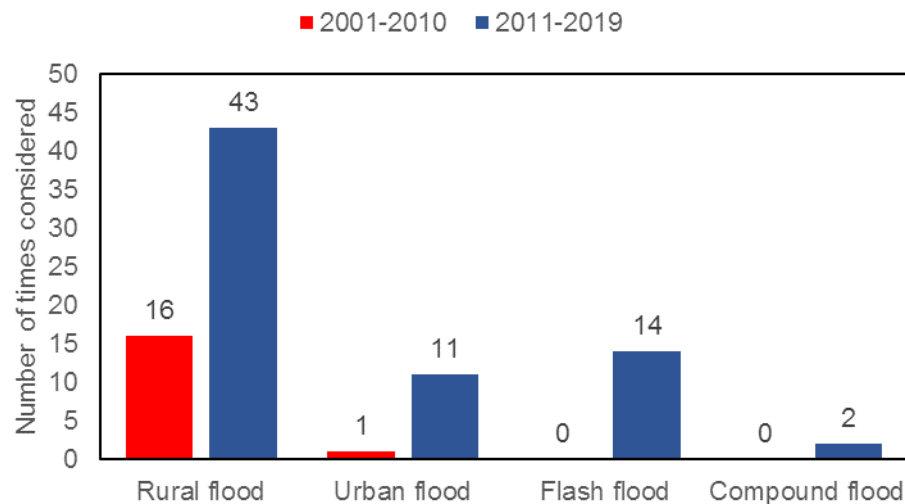

Figure 3 Types of floods investigated (Note: A flood event could be double counted, as an event could be a flash flood event occurred in an urban catchment. Therefore, the total number in the figure exceeds the total number of papers reviewed)
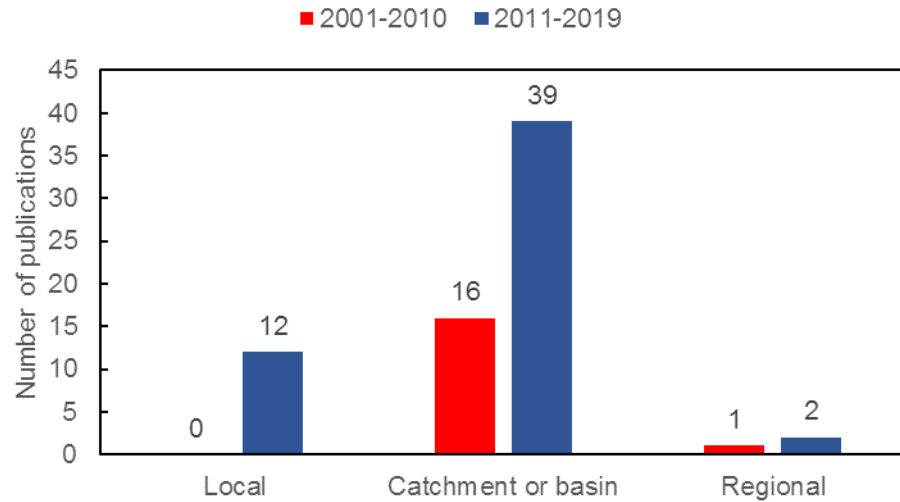
Figure 4 Spatial scale of floods investigated

Compared to the types and spatial scales of floods investigated, there are less significant changes in the temporal resolution of the flood events (i.e. the frequency of flood forecasts produced) investigated over the years. For both time periods before and after 2010, sub-daily flood data (e.g. hourly, 3-hourly or 6-hourly) were used in most of the studies (compared to daily flood data). There is a slight shift from daily data to sub-daily flood data after 2010, which is probably again due to the shift in the types of floods investigated and improved data availability in recent years. Urban floods have a relatively shorter duration (i.e. the length of the flood event) and event onset time (i.e. the response time between the storm and the beginning of the flood event) and therefore lead time (some are classified as flash floods with a lead time less than 6 hours). As a result, urban floods are often investigated with sub-daily data. In addition, very short (i.e. lead time less than 24 hours) to medium range (i.e. lead time of 3-15 days) flood forecasts are the focus of ensemble flood forecasting for both pre- and post-2010 periods, as shown in Figure 6. Only 5 studies considered lead time longer than two weeks, with 4 of these studies were conducted after 2010.
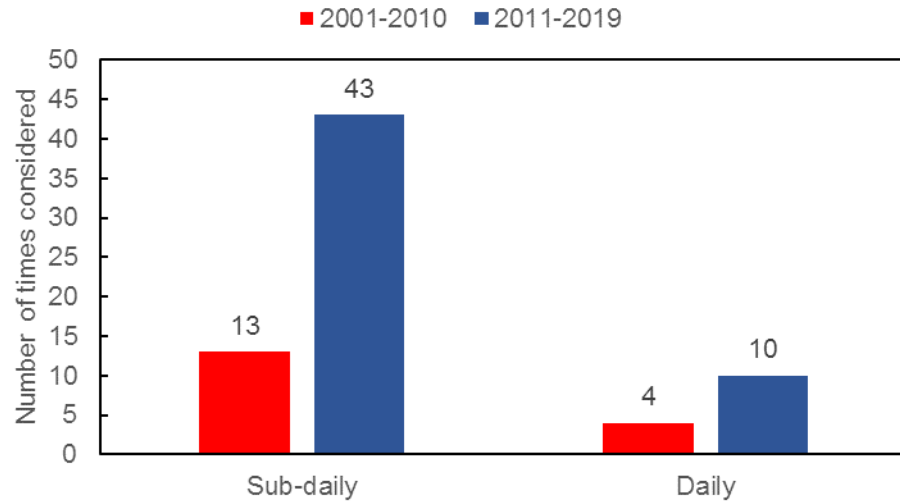
Figure 5 Temporal resolution of floods investigated
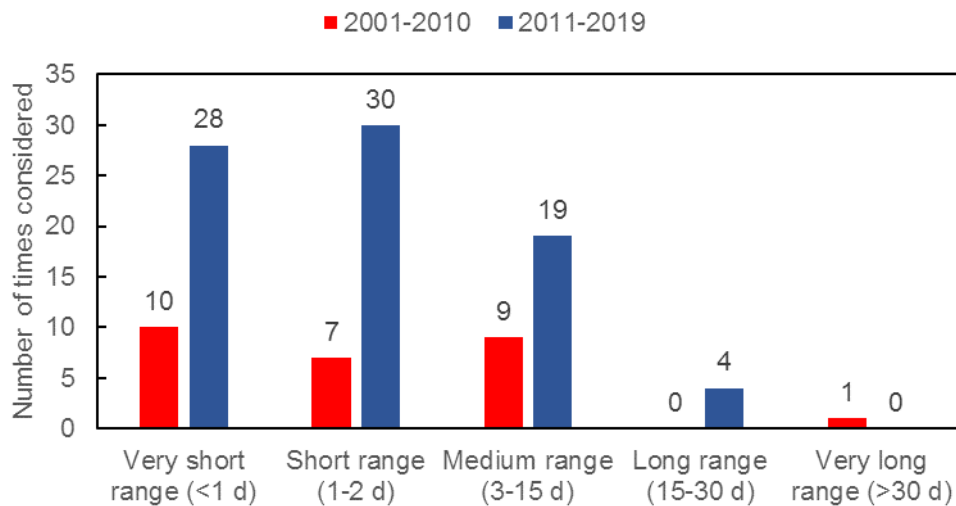
Figure 6 Forecast lead time considered

Flood modelling

Various flood models have been used in the studies to forecast different types of flood variables, as summarised in Figures 7 and 8. Hydrological models simulating rainfall-runoff relationships are the most commonly used flood models in the studies investigated, especially during the first half of the review period with hydrological models accounting for 76% of the studies reviewed. Consequently, river flow is the most common flood variable predicted (Figure 8). During the second half of the review period, although hydrological modelling is still a popular approach, a variety of other flood models were also considered, such as hydraulic models and data-driven models. This also related to the increase in the type of flood variables being considered, - with flood water levels, rainfall (as a surrogate of flood severity) and flood inundation area.
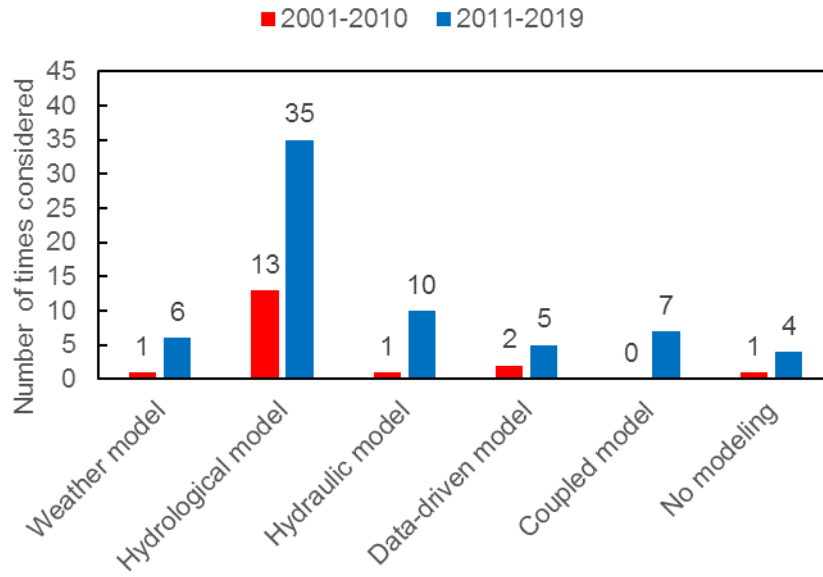
Figure 7 (Surrogate) Flood models used
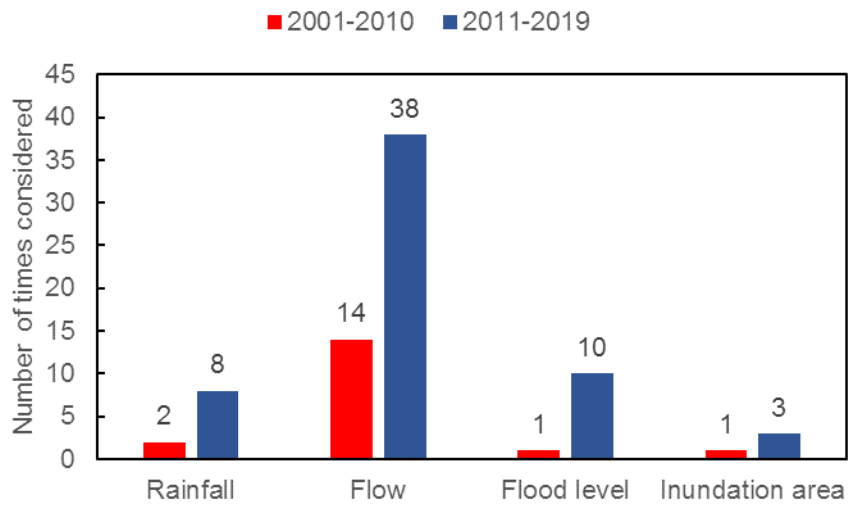
Figure 8 (Surrogate) Flood variables predicted
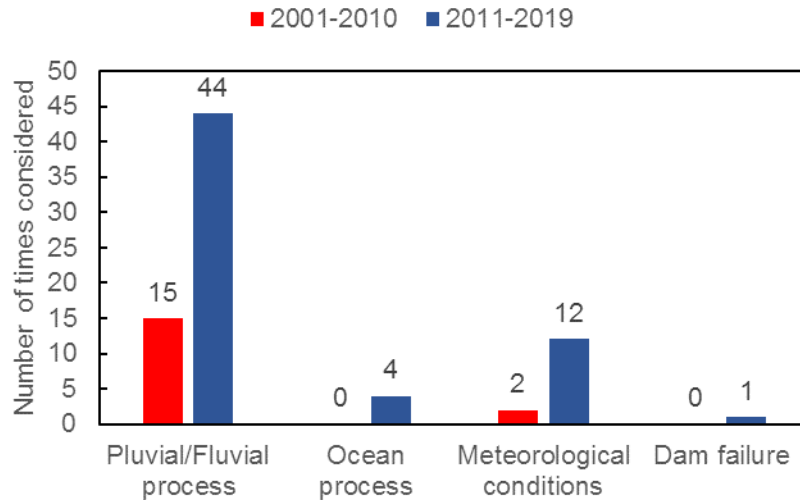
Figure 9 Flood driving processes considered

Fluvial or pluvial flooding driven by rainfall are the most common flood processes considered during both review periods (Figure 9). In the second review period, research was extended to include other flood driving mechanisms, including ocean processes such as storm tide and storm surge, and meteorological conditions such as temperature, air pressure, humidity and wind, which are typically used in machine-learning models to forecast floods affected by these variables (e.g. snow melt driven floods), and catastrophic events such as dam failure. The ocean process is often included using a hydraulic model; whereas meteorological conditions are often incorporated through a numerical weather model or a data-driven modelling approach. It should be noted that in the papers reviewed, there is only one study that considered floods caused by dam failure, and floods due to groundwater events or glacial lake are not considered.

Ensemble modelling

As discussed previously, ensemble forecasts can be generated using a variety of different methods. In the studies reviewed, four key methods were typically used, including multi-model-based ensembles, initial/boundary condition perturbations, parameterisation of models, and statistical approaches to account for error, as shown in Figure 10. Initial or boundary condition perturbation is the most commonly-used method for ensemble generation. This is due to the use of ensemble NWP products, many of which are generated using this method, in order to account for the chaotic nature of the atmosphere (Lorenz, 1963).

Research focussing on evaluation of ensemble forecasts adopts a variety of methods, as shown in Figure 11. Ad hoc methods based on visual inspection or the ensemble mean are the most commonly used methods, followed by threshold-based methods such as estimating the probability of detection compared to observed data. Whereas distribution-based methods accounting for all ensemble members were only used in 3 studies before 2010. Although this number has increased to 22 for the period after

2010, ad hoc methods are still the dominant methods used in the second half of the review period. In a small number of studies (4 for pre-2010 and 11 for post-2010 periods) only ad hoc methods were used.

In both review periods, the majority of studies were based on ensemble forecasts using ensemble sizes ranging from 10 ensemble members to 100, as shown in Figure 12. This is directly related to the ensemble NWP products used. There are, in total, 4 studies using forecasts with over 100 ensemble members. All of these studies focused on analysis of input uncertainty, with two studies also considering uncertainty from inputs generated using different climate models. In the second half of the review period, the number of ensemble members was not mentioned in 10 studies.
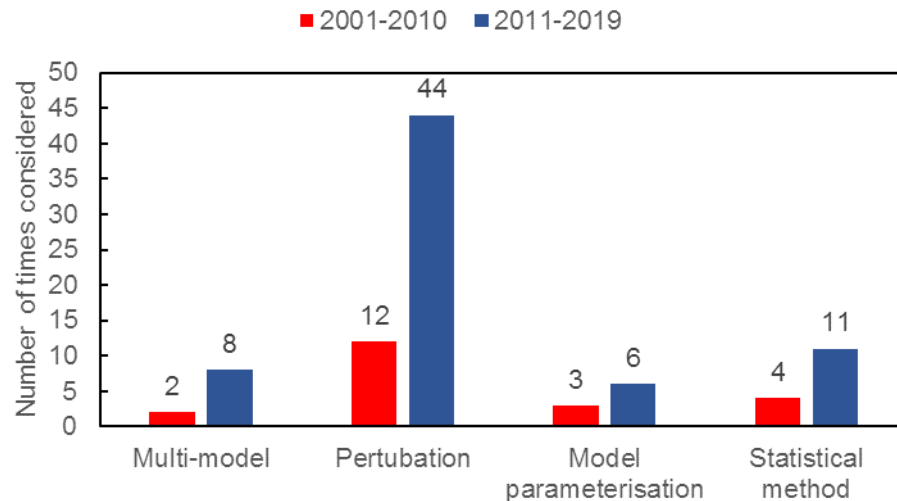

Figure 10 Ensemble generation methods used
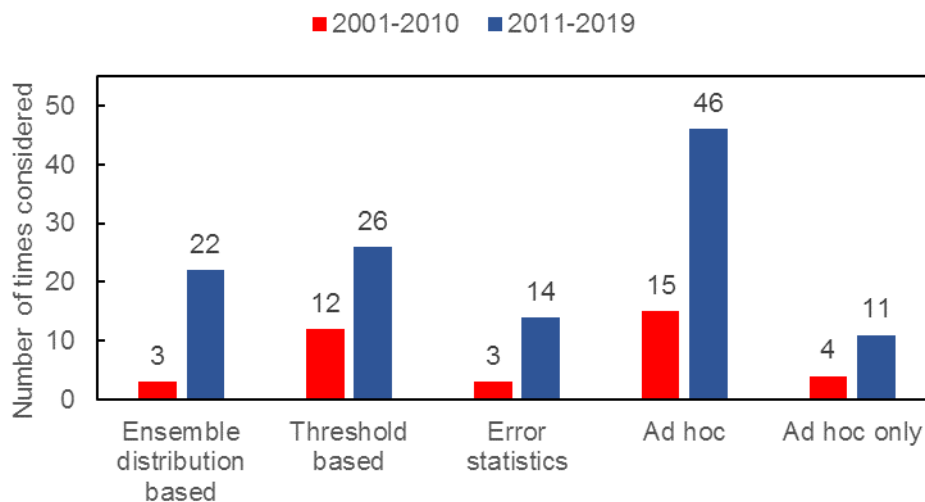
Figure 11 Ensemble evaluation methods used
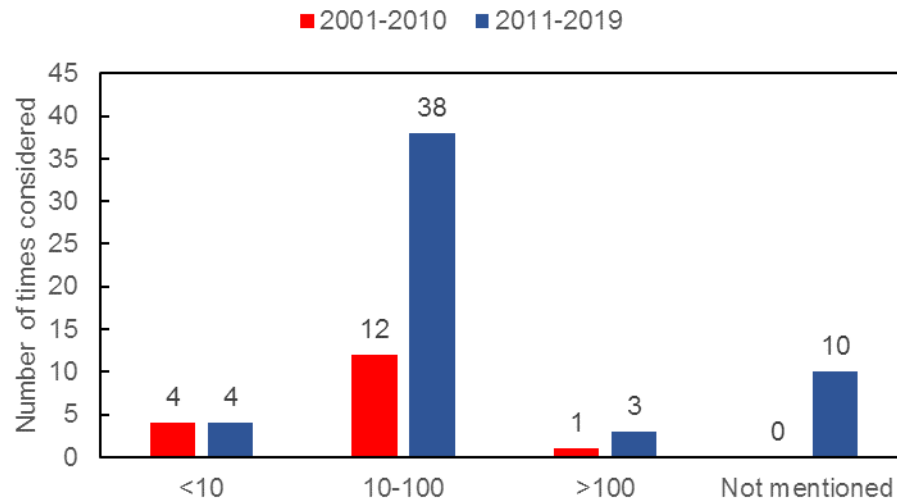
Figure 12 Ensemble member numbers considered

In summary, the review has highlighted that there are an increasing number of studies investigating ensemble flood forecasting. The majority studies have focussed on ensemble forecasting of flooding associated with fluvial processes at the catchment-scale for lead times of up to 15 days. There is considerable emphasis in the reviewed studies on forecasting river flows using hydrological models and ensemble members generated using perturbed forcing. Emerging areas of research tend to focus on extending ensemble forecasting methods to predict a wider range of variables, different types of events and to longer lead times, and the use of ensemble forecasts. These emerging research areas are identifying numerous challenges and research opportunities. Technical challenges are related to extending ensemble forecasting methods include the additional data and computational requirements, and the need for innovative modelling, ensemble generation and forecast evaluation approaches. There are also outstanding challenges related to how ensemble forecasts can be used to inform real-world management of and responses to flooding. The following section expands on these challenges and identifies future research opportunities.

CURRENT CHALLENGES AND FUTURE OPPORTUNITIES

From the review outcomes discussed, we can see that although the overall number of studies on ensemble flood forecasting has increased significantly in the last decade, this increase is not uniform across the different categories reviewed. For example, it varies with the geographic focus of the study, types of floods forecast, and ensemble modelling and evaluation methods used. This is partially due to the quality and availability of observational data. It is also because ensemble forecasting is a relatively new concept in flood-related management applications, although it has been a research focus over the past 2-3 decades in hydro-meteorological modelling. In the following section, we summarise some of the areas for future focus and consideration in the ensemble flood forecasting research and practise communities, including:

- Improving data quality and availability,
- Catering for different types of flood events,
- Extending flood forecast variables,
- Extending and bridging the gap in flood forecast lead time, and

- Extending ensemble generation and evaluation.

We suggest potential research directions in order to move forward and to encourage the uptake of ensemble forecasting in real-world flood management and response.

Improving data quality and availability

All forecasts, whether ensemble or deterministic, are dependent on observational data of high quality for all parts of the flood forecasting chain. Data acquisition for forecasts has undergone rapid and fundamental changes in recent decades with the rise of remote sensing and new types of data, but has at the same time confronted a decline of traditional observations (Dutra et al., 2014). This poses new opportunities as well as new challenges for the forecasting systems with regards to having the best possible input for hydrometeorological forecasting.

There are three main sources for observational data used in hydrometeorological forecasting around the globe; 1) observations from ground-based gauges, 2) remote sensing and 3) modelled reanalysis (Michaelides, 2009). For some areas, ground-based radar can also be used for some variables, such as wind speed and precipitation.

Conventional observational data

Conventional observational data from in situ measurement stations are still the backbone to use in model construction, calibration and validation of hydrometeorological forecasts when they are available and of sufficient quality. However, the global decrease in the number of observational measuring stations in the last decade (Lins, 2008) is especially troublesome when it comes to non-stationarity in observations due to climate change, and the ability to calibrate and validate hydrological models.

Another major challenge for observational data is the lack of metadata (i.e. information about the data), which hinders the assessments of uncertainty. Location, measuring method, type of instrumentation are all variables that contributes to the uncertainty. Further, variables such as river discharge are very often derived from another primary variable, which in effect adds to the uncertainty of the observation.

A problem that is not related to the uncertainty in the actual observation, is the representativity of the data point. Hydrological modelling most often requires spatial fields of forcing data, and these fields are interpolated to continuous fields from point observations with the assumptions of being spatially coherent. This is a problem especially for precipitation, which is highly variable in space and time, also with regards to the field correlation in space and time.

Conventional observation stations are declining in numbers (Lins, 2008). However, automation means that the data collected is available in real-time and with a lot more metadata compared to manually managed stations. An important aspect of observational data collection is to agree on standards for preserving metadata, such as European Commissions INSPIRE directive, which started in 2007 (INSPIRE, 2019). Making sure that all the available information about the data is preserved is time-consuming. Protocols for data exchange, thereby homogenizing data formats, is critical to sharing data across geographical regions and technical platforms. While updating measuring stations is expensive, and could lead to a sparser network of observations, the higher quality of data obtained could mean that the data

is more useful for forecasting. It is worth noting that the majority of these observation stations are located in the developed world; whereas the large parts of the developing world are rarely covered. How to improve observational data collection globally remains a critical challenge.

Remote sensing and reanalysis data

Remotely sensed data, mainly from satellites, has become more and more important as source of information in hydrology (McCabe, 2017). The main advantage is that it contains an abundance of information, but it needs to be calibrated, thinned and quality controlled before it can be used in applications. This often requires observational data of good quality for calibration and validation of the remotely sensed data, therefore this is not independent of ground-based observational point data. Drawbacks of satellite data are that they are sometimes blocked, the coarse spatial resolution, the blocking and the intermittent fly-over of polar-orbiting satellites. However, recent years have seen an increase in small, cheap satellites to complement the larger programmes, making more data available (McCabe, 2017). Ground-based weather radars have proven to give very good spatial coverage of precipitation and can, through calibration, be used for both near-real-time observations and nowcasting (Ross et al. 2011).

Satellite and remotely sensed data are likely to increase in quality and detail in the future (Beck et al., 2017b; McCabe et al., 2017). New missions, such as the SWOT (Surface Water and Ocean Topography) mission at NASA, will provide the community with new data sources that can be incorporated into all parts of the modelling chain (Biancamaria, 2016). However, the limitations of the data will still require observational data of very high quality in the future.

Reanalysis products use models to 'fill in the gaps' where no observations are available, providing best estimates of the state of the atmosphere, and land surface, through model simulations that assimilate observations where available. The most useful global reanalysis data typically cover the satellite era (1979 and onwards), with ECMWF's latest ERA5 reanalysis (C3S, 2017) representing the state-of-the-art. As the output is model-generated, it will suffer from model biases and drifts over time, and the water budget may not be closed (Kauffeld et al., 2015). A typical problem is the 'drizzle effect' for precipitation with too many wet days and underestimation of extreme events. These short comings can be adjusted for through post-processing (Maraun et al, 2010), however this also requires good quality observational data(C3S, 2017).

Reanalysis data is also now being used as source of forcing data for hydrological applications. Increasing model resolution and improved model physics enables the data to be used for training of models, as well as providing initial conditions for forecasting.

New data sources

There are also new products that take advantage of all available data sources, such as the Multi-Source Weighted Ensemble Precipitation (MSWEP, Beck et al., 2017a). They take the best data from all sources of precipitation, to produce an optimum dataset with global coverage. However, for it to be useful for forecasts, it needs to be available in near-real time.

Machine-learning and artificial intelligence methods are alternative ways to make better use of existing data, and are becoming more and more popular in hydrology (Shen et al., 2018). In its extreme form, machine learning methods are even substituting physical models with "black box" approaches which use

all available data with little or no prior knowledge of the physical properties. It is unlikely that these model systems will replace traditional models, it is more probable that they will be used to complement existing physical models.

Opportunities for the future lie in developing hybrid systems, mixing conventional observational data, delivering high quality, high frequency data, with remotely sensed data (Beck et al, 2017a). In a sense, this is what is being done in a reanalysis product, but the drawback of using these systems in operational forecasting has been the low resolution and model bias. However, using the uncertainty inherent in the observations and models can also open the door to ensemble methods in initial conditions of hydrometeorological forecasts.

Catering for different flood events

The majority of flood forecasting systems reviewed here predict river flows within defined river channels. Systems either display little hydrological regulation (Liechti et al., 2013) or neglect the effects of regulation times of flooding (Ushiyama et al., 2014; Yu et al., 2018). The studies generally focus on meso-scale catchments where the interactions between antecedent catchment conditions and forecast weather are both important and there are multiple different methods of producing forecasts with similar performance. At the centre of most of these systems is a model that transforms rainfall to river flow, but the characteristics of these models vary widely from physically based and conceptual hydrological models (Thiemig et al., 2015; Verbunt et al., 2007) to machine learning methods (Kimura et al., 2016; Tiwari and Chatterjee, 2010).

Reducing errors in riverine flood forecasts has been an ongoing research challenge since the advent of hydrological forecasting (O'Connell and Clarke, 1981) and remains outstanding. Errors in flood forecasts are most commonly reduced by assimilating streamflow data to update state variables (Srikanthan et al., 2008) or applying a statistical method to post-process predictions (Yu et al., 2015). Increasingly, studies are investigating the assimilation of remotely sensed observations, such as soil moisture, to update model state variables in flood forecasting models and reduce forecast errors (Li et al., 2016). The advantage of assimilating observations of variables other than streamflow is that they can provide additional information on the partitioning updates between the different state variables and potentially improve the skill of forecasts at long lead times. However, there are outstanding research challenges in how best to relate remotely sensed observations to the state variables of hydrological models and the most effective strategy for assimilation (Li et al., 2016).

More recent studies report on forecasting of flash floods. Various definitions of flash flooding exist (Hapuarachchi et al., 2011), but common to all definitions is a short time lag between intense rainfall and a flood peak. Flash floods occur when river flow reacts very rapidly to intense rainfall. They typically occur over limited spatial extents, in areas of steep slopes, and shallow or saturated soils (Liechti et al., 2013). Many of the flash flood forecasting studies reported extend riverine forecasting methods by introducing more accurate precipitation nowcasts (Marty et al., 2013) or some form of improved data assimilation (Chen et al., 2013). However, flash flooding also occurs in urban areas (Einfalt et al., 2009) and in rural landscapes away from riverine channels, but ensemble forecasting methods of these types of floods are not widely reported. This may be due to limitations on the availability of data or that

methods for generating flash flood forecasts need to be refined to reflect the dominant processes involved.

Accurately forecasting flash flooding is an intrinsically difficult problem requiring the prediction of high rainfall intensities in the precise location where it occurs. At very short lead times (0-3hrs), ensemble radar nowcasts can provide precipitation forecasts to support flash flood forecasting (Kim et al., 2009). Blending radar nowcasts with downscaled high-resolution NWP model outputs can potentially extend lead times, however beyond 6 hours precipitation forecasts from NWP models are the dominant source of any skill (Bowler et al., 2006). While the performance of rainfall forecasts from NWP models is improving, models rarely resolve the fine scale processes required to accurately predict the rainfall that causes flash flooding (Emerton et al., 2016). Improving the ability of NWP models to predict the location and intensity of rainfall remains a challenge in improving the quality of flash flood forecasts and is likely to require increases in the spatial resolution of NWP models as well as advances in observation, data assimilation and model physics.

While flash flooding predominantly is a pluvial process, the land surface hydrology can modulate the transformation of precipitation to runoff and hence the extend of flooding. To represent the effects of land surface hydrology on flash floods, hydrological modelling needs to be undertaken at fine spatial resolutions to better represent the processes that lead to flash flooding (Antonetti et al., 2019). There are many outstanding research challenges with hyper-resolution hydrological modelling, particularly with respect to the characterisation of hydrological process, availability of high-resolution observations and growth in computational requirements as model resolution increases (Bierkens et al., 2015).

Forecasting of compound flood events receives little attention in the reviewed studies. Compound floods are those that are caused by two or more events that influence flooding (Leonard et al., 2014). They may be multiple instances of the same event type, for example consecutive extreme rainfall events over a single catchment or current extreme rainfall in multiple catchments upstream of a river confluence; or the interaction of two different types of flood inducing phenomena, such as coastal flooding arising from the interaction of a riverine flood and high sea levels caused by high tides or storm surge (Leonard et al., 2014; Zscheischler et al., 2018).

Compound floods resulting from meteorological forcing are typically handled by the existing approaches to ensemble flood forecasting. Many studies illustrate forecast of multi-peak floods arising from multiple precipitation extremes (Hopson and Webster, 2010; Komma et al., 2008; Marty et al., 2013), while Verbunt et al. (2007) evaluate ensemble flood forecasts arising from the combination of intense precipitation and anomalously high temperatures accelerating snow melt.

Few studies have focussed on compound floods resulting from different forcing, with only one reviewed study forecasting flooding from the combination of intense rainfall and sea level surge during tropical storms (Saleh et al., 2017). Ensemble forecasting of compound flood events is key research challenge. Methods of ensemble forecasting that involve pre or post-processing will require new innovation to handle the joint probability of extreme events for different phenomena. The development of coupled models also will address the challenge of forecasting compound flood events. One such example is the coupled atmosphere, land surface and hydrodynamic models currently being developed for the Great Lakes of North America (Gronewold and Fortin, 2012). The ability of these coupled models to forecast compound flooding has yet to be reported in the literature, however there have been numerous

research and technical challenges in the process of coupling and parameterising models (Gaborit et al., 2017; Gronewold et al., 2018; Pietroniro et al., 2007; Vionnet et al., 2019).

Extending flood forecast variables

The majority of flood forecasting systems reviewed here predict river flow as their main output variable, yet there are an increasing number of studies describing models predicting other flood variables, such as inundation and water levels. Typically, producing forecasts of inundation and water levels requires the use of a hydraulic model in addition to a rainfall-runoff model. This adds another layer of uncertainty to the forecasting chain, due to localised variations in river channels and floodplain geometry that must be parameterised using observed data, which can be limited (Boelee et al. 2017). A key challenge and limitation in terms of providing robust forecasts of flood inundation that appropriately represent the forecast uncertainty, is the computer power required to run multiple simulations of both a hydrological and hydraulic model (Pappenberger et al. 2005), alongside the accuracy of the rainfall forecasts used as input to the models.

Teng et al. (2017) state that "despite active research in the field, rapid and accurate flood modelling at high spatial-temporal resolutions remains a significant challenge in hydrologic and hydraulic studies". While significant progress has been made, with current state-of-the-art 2D and 3D hydrodynamic models able to accurately simulate flood inundation, alongside velocity and water level, these require extensive computing resources and are typically only run for small catchments, particularly at high resolutions. Simplified, often 1D, versions are less computationally expensive, but have limitations in the simulation of flood waves and the representation of topography (Teng et al., 2017).

There exists a huge range of flood inundation models, from those using empirical methods to physically-based hydrodynamic models, and various studies (Neels and Pender, 2013; Teng et al., 2017) have indicated that the differences in the predictions from these various approaches only serves to highlight the significant uncertainty we still face in flood prediction, and the importance of employing ensemble prediction techniques.

While a key challenge in ensemble forecasts of flood inundation lies in the computer power required, which imposes limits on producing ensemble inundation forecasts in real-time, recent research (e.g. Chu et al., 2019) has begun to evaluate the use of emulation models, including those developed using machine learning techniques, for flood inundation forecasting. The application of such techniques to flood inundation forecasting is a recent and developing area of research. While challenges still remain (Chu et al.,2019), there is potential for such models to be used in real-time forecasting by significantly reducing computation time.

An important recent advance and ongoing area of research is the use of remote sensing techniques to improve the accuracy of flood models (Li et al., 2016; Li et al., 2018). Remote sensing data can be used to calibrate and validate flood inundation models (Dasgupta et al., 2019), and in future will likely be integrated into the modelling chain. One key way to incorporate and benefit from new, real-time datasets is through data assimilation, which is commonly used in meteorological ensemble forecasting, but currently less frequently used within hydrological (or hydraulic) modelling (Emerton et al., 2016).

Recent advances in hydrodynamic modelling alongside computing (e.g. parallel computing, open source code) have led to vast improvements in flood inundation modelling. With ongoing research and model

development it will become more common for variables beyond river flow, such as inundation and water level, which are directly related to flood impacts, to be included in ensemble flood forecasts (Cooten et al., 2011; Georgas et al., 2016; Gomez et al., 2019). This is particularly promising in the near future with forecasting services worldwide moving more and more towards the provision of impact-based forecasts (WMO, 2018). Ensemble forecasts of flood inundation would allow for a better representation of uncertainty when considering exposure, vulnerability and risk associated with a forecasted event.

Extending and bridging the gap in ensemble flood forecast lead time

The papers reviewed in this study have mostly focused on short- to medium-range ensemble forecasts of flood events. However, earlier indications of potential flood events, many weeks or even months in advance, could provide crucial information for flood preparedness and disaster risk reduction. A key area of ongoing and future research in hydrological ensemble prediction is that of extending the lead time of flood forecasts to sub-seasonal scales (Bennett et al., 2014). Many operational forecasting centres produce seasonal forecasts of meteorological variables, which are often averaged into "seasonal-mean" estimates of climate, such as whether the coming season will be warmer or cooler than usual (Weisheimer and Palmer, 2014). The equivalent long-range forecasts for hydrological variables can also provide crucial information for flood preparedness and mitigation (Yuan et al., 2015a). Indeed, the use of seasonal probabilistic and ensemble hydro-meteorological forecasting systems for water management has a long history. The first seasonal-scale ensemble streamflow forecasts were introduced in the mid-1970s (Wood et al, 2016; Day, 1985), and were used to support seasonal volumetric inflows and daily sequences for managing reservoirs. These forecasts are currently implemented operationally at regional to national scales in numerous countries. Improving seasonal forecasts remains an active and ongoing area of research, from local (Mendoza et al, 2017) to regional (Bell et al., 2017; Bennett et al., 2016; Crochemore et al., 2016; Meißner et al., 2017; Mo et al., 2014; Prudhomme et al., 2017; Wood et al., 2002, 2005; Yuan et al., 2013) and global (Candogan Yossef et al., 2017; Yuan et al., 2015b) scales.

In recent years, several operational continental-scale seasonal hydro-meteorological forecasting systems have been launched, using meteorological forecasts as input to hydrological models to produce seasonal forecasts of hydrological variables. Some examples of operational seasonal forecasts of hydrological variables include: the Copernicus Emergency Management Service (CEMS) European Flood Awareness System (EFAS; Arnal et al., 2018; Cloke et al., 2013) and Global Flood Awareness System (GloFAS, Emerton et al., 2018), the European Service for Water Indicators in Climate Change Adaptation (SWICCA; Copernicus, 2018), the Australian Government Bureau of Meteorology Seasonal Streamflow Forecasts (Bennett et al., 2017; BoM, 2018) and the USA's Advanced Hydrological Prediction Service (AHPS; McEnery et al, 2005) and Hydrologic Ensemble Forecast Service (HEFS; Demargne et al., 2014; Emerton et al., 2016).

Despite the chaotic nature of the atmosphere (Lorenz, 1963), which introduces a limit of weather event predictability, seasonal climate predictions are possible as they rely on components that vary on longer timescales, known as "second type predictability" (Lorenz, 1993). In the seasonal hydrological forecasting context, predictability comes from the inertia or climatological evolution of initial conditions and a lesser amount of predictability arising from large-scale modes of climate variability (Emerton et

al., 2018; Wood and Lettenmaier, 2008). Various patterns of climate variability, and their associated teleconnections, influence river flow and flooding across the globe, such as the El Niño Southern Oscillation (ENSO; Chiew and McMahon, 2002; Emerton et al., 2017; Guimarães Nobre et al., 2017; Ward et al., 2014a, 2014b, 2016), North Atlantic Oscillation (NAO), Southern Oscillation (SOI), Indian Ocean Dipole (IOD) and Pacific Decadal Oscillation (PDO) (Yuan et al., 2015b). As such, they are able to contribute to the seasonal predictability of hydrological variables. Coupled atmosphere-ocean-land or earth system models (ESMs) "are key in representing these large-scale modes of climate variability in order to produce seasonal hydro-meteorological forecasts" (Emerton et al., 2018), along with empirical forecast models relating climate system states to surface climate, and in recent years, hybrid approach combining dynamical and empirical approaches (e.g. Weisheimer and Palmer, 2014).

Another way to extend forecast lead time is to use seamless forecasting, which was first introduced by Palmer and Webster (1993), and aims to provide added predictability and lead time through integration of short- and medium-range forecasts with monthly- and seasonal-range forecasts. Seamless forecasting is based on the idea that "short-lived phenomena under certain conditions may persist and increase predictability at longer timescales" (Wetterhall and Di Guiseppe, 2018). Currently, in both meteorological and hydrological forecasting, systems producing medium-range and seasonal forecasts are typically distinct, and provide separate forecast products. In meteorological forecasting, ECMWF produces an extended-range forecast, which extends the medium-range forecasts out to monthly timescales (Vitart et al., 2008) using a seamless forecasting approach, and the UK Met Office Unified Model is designed to work across multiple timescales (Brown et al., 2012). So far, there is only limited application of this approach in operational hydrological and flood forecasting. Its application could be expanded to benefit operations by providing earlier indications of anomalous conditions. A recent study by Wetterhall and Di Giueseppe (2018) used an experimental configuration of EFAS, merging ECMWF's extended-range and seasonal forecasts before using them as input to a hydrological model, and showed that this seamless approach was able to provide more skilful forecasts -- up to 7 weeks ahead -- than forecasts using only the seasonal meteorological forecasts as input. The practice of merging short range NWP with longer range meteorological forecasts has long been an operational approach used in the US NWS, although until the advent of HEFS, the short-range forecast component was deterministic.

In future, we are likely to see more extended-range and seamless forecasts of hydrological variables, particularly as earth-system models continue to be developed and improved, allowing decision-makers to benefit from the combination of earlier indications of anomalous events, and shorter-range forecasts of specific flood events.

Extending ensemble generation and evaluation

Current practice in ensemble flood forecasting is mostly based on generating ensemble forecasts from different meteorological inputs, different initial conditions, multiple hydrological models, or multiple parameter sets, or a combination of above (Cloke and Pappenberger, 2009; Duan et al., 2019; Roundy et al., 2019). The most common approach to ensemble flood forecasting is to generate flood forecasts by perturbed initial conditions for either NWP models, the output of which are used as forcing inputs to drive a hydrological model(s), or for the hydrological models themselves (Moradkhani et al., 2019). For the former approach, raw ensemble precipitation forecasts must be post-processed via some type of

statistical techniques to remove the systematic and spread biases inherent in those forecasts (Li et al., 2017).

There are different types of post-processors, i.e., the statistical models relating raw forecasts to corresponding observations. There are numerous methods for characterising the raw forecast/observation dependence relationships. A popular method is to use various regression techniques, including multi-variate linear regression (not suited for precipitation forecasts, Glann and Lowry, 1992), nonlinear regression (Wilks and Hamill, 2007; Scheurer, 2014; Baran and Nemoda, 2016), logistic regression (Wilks, 2009), and quantile regression (Weerts et al., 2010). Another commonly used method is to construct joint probability models of raw forecasts and observations (Krzysztofowicz and Maranzano, 2006; Schaake et al., 2007; Wang et al., 2009). Other types of post-processors include analogs (van den Dool, 1994; Hamill and Whitaker, 2006; Todini, 2013), ensemble dressing (Pagano et al., 2013), multi-model averaging methods (Krishnamurti et al., 1999; Berrocal et al., 2008), among others. Reforecast data (i.e. the hindcasts of meteorological variables generated using a current meteorological model) covering a sufficiently long period of time (i.e. 20-30 years) are essential to constructing robust post-processors (Hamill et al., 2004). As ensemble precipitation forecast skill varies with lead times, separate post-processors for events of different lead-times need to be developed. For example, the statistical model relating the first day forecasts and observations would be different from the one relating the fifth day forecasts and observations as their statistical dependence structure would change. Once post-processors for different events covering the entire forecast horizon are constructed, it is then necessary to find ways to create ensemble space-time precipitation series needed to drive a hydrological model. The Schaake shuffle method has been a widely used method for creating space-time series of meteorological variables from post-processors of different events (Clark et al., 2004).

For the latter approach, where the ensemble flood forecasts area generated by directly running a hydrological model(s) using perturbed initial conditions (Moradkhani et al., 2019), numerous methods can be used. The simplest way is to assume a certain distribution of the initial conditions and employ a Monte Carlo approach to sample the distribution to initiate the forecasts. More sophisticated way is to use ensemble data assimilation methods such as Ensemble Kalman Filter (Chen, et al., 2011) or some variational data assimilation methods (Lee et al., 2011) to merge model state variables and appropriate observations. It has been argued that the skill derived from model initial conditions can exceed that from the meteorological forecasts, especially at seasonal time scale (Maurer and Lettenmaier, 2003; Wood and Lettenmaier, 2008; Shukla and Lettenmaier, 2011; Shukla et al., 2013).

As all models are based on abstraction of real-world processes and are not expected to perform well under all conditions, ensemble forecasts based on multi-model methods have gained traction over the recent years (Ziehmann, 2000; Georgakakos et al., 2004; Ajami et al., 2006; Duan et al., 2006) and have also been used in many of the studies reviewed. Multi-model ensemble methods can be as simple as the poorman's ensemble, which treats all model predictions as equally good (Ebert, 2001), to more sophisticated methods such as super-ensemble approach (Krishnamurti et al., 1999) and the Bayesian Model Averaging method (Raftery, 2005), which weigh model predictions based on how well the predictions match the observations. Multi-model predictions have shown to have more skills than single model predictions in many studies (Ajami et al., 2006; Ajami, 2007; Duan et al., 2007; Madadgar and Moradkhani, 2014) and could be a promising approach to consider uncertainty in ensemble flood forecasting for flood management.

Hydrological models generally contain a large number of parameters that vary with local conditions and are not directly observable. They must be inferred based on the fitness of model predictions with observations. Due to various uncertainties such as model initial conditions, model structure and the data used to calibrate the model, there is no unique set of parameters that are always the best for a particular application (Beven and Binley, 1992; Kuczera, 1983). Therefore, generating ensemble flood forecasts based on multi-parameter sets has been promoted by many researchers (Bates and Campbell, 2002; Ajami et al., 2007) and have also be employed in a number of the studies reviewed.

There are also attempts to consider different sources of uncertainties in an integrated manner. For example, Kavetski et al. (2006) and Ajami et al. (2007) proposed different methods to consider errors from inputs, model parameters and or model structure in a Bayesian framework. However, methods to consider all uncertainties in an integrated manner in operational hydrological forecasting are still not available yet and the development of such methods is a challenge for future research.

Ensemble flood forecasting represents a paradigm shift in hydrological forecasting, as the forecast is provided in a probabilistic form, instead of the single trace time series in traditional forecasting. Traditional verification measures of forecast skills such as root-mean square error (RMSE), mean absolute error (MAE), Nash-Sutcliffe Efficiency (NSE), or correlation coefficient are no longer adequate to verify probabilistic forecasts when used in isolation, even though they have been used to verify the ensemble means in many of the studies reviewed. Other measures suited for probabilistic forecasts should be used. Good ensemble forecasts embody three traits: (1) the probability distribution is sharp, (2) forecasts do equally well for events of different magnitudes, and (3) forecast probability is consistent with observational frequency. There are a number of measures that have been developed for verification of ensemble forecast performance (Anctil and Ramos, 2019), including Brier scores (BS), continuous ranked probability score (CRPS), ranked histograms (RH), reliability diagrams (RD), and relative operating characteristics (ROC). A summary of the verification measures can be found in Brown et al. (2010). For comparison purpose, some of those verification metrics have been scaled to certain bench marks to yield skill scores. The benchmarks used are usually based on climatology or skill score of the raw forecasts.

In addition, in the majority of the studies reviewed, between 10 and 100 ensemble members were used. This is consistent with most of the NWP products, the outputs of which are ultimate inputs for flood models. In a few studies, less than 10 ensemble members were used. Although depending on the specific cases, such a small number of ensemble members would be difficult to represent any type of uncertainty that needs to be quantified for a flood system. In 10 studies in the second review period, the number of ensemble members used was not mentioned, and as such, it is difficult for readers to understand the level of uncertainty considered and potentially undermines the credibility of the findings presented. Even with over 100 members (that are used in 4 studies), it would be difficult to represent extreme events with an occurrence probability of 1% or less. This issue could be exacerbated when the extreme flood events are caused by multiple (sometimes dependent) processes, such as such as coastal floods caused by extreme rainfall and storm surge, due to the increased risk (and rareness) of these flood events (Wu et. al, 2018). However, increasing ensemble members will increase computational effort significantly. Therefore, strategies for adequately representing uncertainty in compound floods using ensembles forecasts within a feasible computational budget remain a future research challenge.

Moving towards the adoption of ensemble flood forecasts for water management

The past 15 years have marked significant accomplishments in the development and operational adoption of strategies for ensemble flood forecasting. When HEPEX (A global community in hydrological ensemble prediction) was initiated in 2004, ensemble flood forecasting was virtually unheard of in official practice, yet initial research was underway toward understanding the ramifications of using ensemble datasets and methods in hydrologic prediction (Franz et al, 2005; Schaake et al, 2006). Since then, a coherent view of the critical scientific challenges in ensemble flood forecasting has emerged, and hundreds of studies (many of which are reviewed in this article), have been published to shed light on theoretical and practical aspects of challenges such as hydrologic ensemble data assimilation, ensemble forecast post-processing, meteorological ensemble forecast downscaling, and ensemble forecast verification. In addition, a HEPEX-led Handbook of Hydrometeorological Ensemble Forecasting (Duan et al., 2019) was written to discuss not only the relevant challenges and existing strategies, but also the details of numerous real-world case studies. Complementary advances have also been made in hydrologic model development and parameter estimation, which represent further critical challenges (Clark et al, 2017). This progress has enabled a number of countries to implement operational ensemble flood forecasting systems (Addor et al., 2011; Pappenberger et al., 2016, and references therein; Emerton et al, 2016, and references therein), even if these still generally run in parallel to their legacy deterministic flood forecasting systems. There are by now many clear operational examples of ensemble forecasts that are used to support decision-making, with many examples to choose from, such as the New York City Department of Environmental Protection's Operations Support Tool, a state-of-the-art decision support system for the city's water supply system (NASEM, 2018), in the Sonoma County Water Authority Lake Mendocino system operations (Delaney et al, 2019), for navigation scheduling operations in Germany (Hemri and Klein, 2017; Meißner and Klein, 2015; Meißner et al, 2017), and in many private sector operations, particularly those related to hydropower.

The development and adoption of ensemble flood forecasts in risk mitigation for humanitarian objectives have also expanded, with recent examples including the application of global-domain ensemble flood forecasts from systems such as GLOFAS. Though partnerships with the International Red Cross, for instance, the operational GLOFAS ensemble forecasts support strategic decisions in the staging of emergency resources ahead of and during flooding events in Africa and South America (ECMWF, 2020; GloFAS , 2020; RCCC, 2019).

The new ensemble forecasting systems are diverse but not ubiquitous, and this current landscape indicates that despite substantial theoretical progress, operational ensemble flood forecasting has yet to become the standard. Note that the history of operational ensemble NWP, which has moved along at least a decade or more relative to ensemble numerical hydrological prediction (NHP), suggests that both deterministic and ensemble NHP are likely to co-exist for the foreseeable future, with some centres or stakeholders favouring products generated via more complex systems that are computationally infeasible in ensemble contexts. To the extent that the two enterprises compete for resources, and may involve different development and operations groups, this parallel path may be counterproductive even while it may best serve a broad range of stakeholders.

The challenges currently facing the expansion of ensemble flood forecasting into operations, and further into the decision workflows of water and emergency management and related sectoral uses can be broadly grouped into several major categories: (1) scientific hurdles, (2) the operational readiness of the

forecasting community, (3) communication and outreach to users, and (4) the operational capacity of the decision-making sector. Several additional, critical difficulties affect both traditional deterministic forecasting as well as ensemble forecasting, including the sufficiency of observational networks, the need for effective approaches to parameter estimation, and the lack of reliable hydrologic data assimilation (DA) methods, all of which are exacerbated in regional or larger scale applications.

In general, such unmet scientific challenges are manifest in poor forecast quality, which impedes their acceptance and stakeholder perceptions of usability. Perhaps the most critical scientific need for ensemble forecasting is a reliable DA implementation, which particularly affects ensemble systems because they tend to be automated (versus interactive, as with many deterministic systems). Many systems, and especially regional to global ones, lack DA at present and consequently their outputs can be unusable. Despite ample research into DA in controlled or small sample (e.g., single basin, short record) studies, real-world demonstrations of DA in operational real-time systems are generally lacking. Other hurdles exist, but may not slow adoption to the same extent. For instance, because forecast spread (uncertainty) is a component of the forecast, techniques to ensure forecast reliability (essentially, accurate spread or a depiction of uncertainty) are critical. Forecast post-processing techniques are moderately successful in adjusting "total forecast uncertainty" (e.g. Bellier et al, 2018), but research is needed into whether improved ensemble initialisation or ensemble modelling techniques in hydrology (as in NWP) may also be an effective strategy. Likewise, improved pre-processing of ensemble NWP would also benefit forecasts, and particularly forecasts of extremes, but this need has not stopped ensemble forecasts from being generated with ensemble NWP, and by myriad methods, to date.

While operational flood forecasting is a long-standing practice, flood forecast production centres often lack a detailed understanding of what is involved in producing a skilful, automated, reliable ensemble forecast, including the aforementioned DA methods, as well as NWP ensemble pre-processing, streamflow ensemble post-processing, verification practices, the value of hindcasting, and even intuitive product formation. Operational centres in hydrology often lack a strong connection to the research community, and may be unaware of beneficial advances. The existing operational data and computing systems may be unready to expand to accommodate an ensemble workflow, especially if it must compete for resources and attention with a well-accepted deterministic workflow. Forecast staff also require training and support, and such resources take time to prepare, to resource, and to standardize throughout an organization. For instance, the deployment of HEFS in the US took years, which included not only data and forecast systems but also such human resources capacity building. Recently, innovative methods have also been developed to highlight the complex decision-making involved in flood management. One example is the IMPREX serious game, called "Pathways to running a flood forecasting centre: an adventure game", developed by Arnal et al. (2017; www.imprex.arctik.tech) as part of the Horizon 2020 IMPREX project. The project aims to draw attention to the role forecasts can play in managing extreme weather, with the example of protecting a city from floods using ensemble forecasts and problem-solving skills. This is just one of a range of serious games that have been developed around ensemble hydrological forecasting (Arnal et al., 2017; HEPEX, 2020).

At the last HEPEX international Workshop (Melbourne, Australia, February 2016), a poll of participants revealed that the highest ranked priorities for HEPEX involved communication related topics. Participants from forecast user groups questioned the costs versus benefits of shifting to using an ensemble forecast, as well as the practical mechanics of how such forecasts might be used for decision-

making. This value proposition is being supported by researchers (Zappa et al. 2013) and demonstrated episodically, by groups that have become early adopters (such as Germany's BfG forecasting agency, which assessed the marginal economic benefits of ensemble forecasts for navigation), but not communicated uniformly throughout the potential stakeholder community. Stakeholders with strong incentives to use forecast information (such as hydropower companies, which can translate improved or expanded information into economic gains) are relatively advanced in understanding such questions, but such stakeholders may not openly share their practices with the community.

A related challenge is that the operational capacity to ingest and incorporate ensemble forecasts into decision support systems (DSS) for water management may be lacking. DSS such as water management models may be designed to simulate a single inflow scenario, and require manual input from reservoir managers to determine release schedules, meaning that a significant upgrade and change to a more automated paradigm would be needed to shift to using ensemble forecasts. Optimal rules for making risk-based decisions may also be unknown, and require time and experience for development. Any aspects of operations that require expert input or judgement, or interaction with multiple parties, can be a hurdle to reliance on ensemble-based information. While sophisticated optimal control system approaches have long existed (e.g. the use of stochastic dynamic programming in reservoir operations) in some parts of the water management sector (and particularly in private or hydropower-related entities), these are by far not the standard, particularly for publicly owned systems.

SUMMARY AND CONCLUSIONS

We provide an overview of research activities in ensemble flood forecasting based on a detailed review of 70 papers focusing on various aspects of ensemble flood forecasting from around the world. We have identified five challenges for current research and outlined potential key future research opportunities. A summary of the review outcomes and how they related to the challenges and further lead to opportunities is provided here.

Although ensemble flood forecasting has been a research focus for 2-3 decades, the majority of the research activities focus on hydro-meteorological forecasting. This is evident from the review outcome that rainfall and flow are predicted (via climate and hydrological modelling) in over 80% of the cases investigated and only a few studies considered flood inundation levels or extent, which are directly related to flood impact. This is also reflected in the review outcome that most of the flood events considered occurred on a catchment or basin scale, with less than 20% of studies considering floods on local scales, which mostly relate to urban or flash floods. The strong focus of the studies on catchment hydrological processes is directly related to data availability. Conventional observational data are commonly collected for rural catchments and very few urban catchments are gauged. However, this situation may change as new data sources, such as remote sensing and reanalysis data, become more accessible and the quality of these data further improve. In addition, computing resources requirement is another limiting factor for the consideration of flood inundation variables. Most commonly used flood inundation models are 2D hydrodynamic models, which are computationally expensive, making it difficult to consider uncertainty using ensemble forecasting techniques. Furthermore, the disconnect between the hydro-metrological and hydraulic modelling research communities further exacerbates the situation.

In the papers reviewed, only more recent studies reported on flash floods, which have a very short lead time (i.e. less than 6 hours). This is not only due to limitations on data availability, but also the high requirement on the precision of the prediction of both rainfall intensity and location. Improving the forecast accuracy of flood location requires high resolution NWP models and hydrological models (if a pluvial process is involved), both of which are outstanding challenges. In addition, very few papers considered compound floods, which are caused by multiple (often dependent) driving mechanisms, such as coastal floods caused by extreme rainfall and sea levels. Compound extreme flooding is an emerging topic for flood management due to the growing awareness (and publicity) of such events from increased intensity of extreme weather events giving arise to them, such as Tropical Cyclone Yasi in Australia in 2011 and Hurricane Sandy in the US in 2012, both caused extensive flooding and devastating consequences in coastal areas. The reason behind the lack of consideration of uncertainty using ensemble techniques in compound flood forecasting is also multi-fold. First, it is more difficult to generate/collect data of multiple dependent flood drivers during the same time period. There are only limited locations where such observational data may exist. This situation cannot be improved easily with new emerging data sources, such as satellite observations, due to the extreme rareness of compound floods. Different compound flood drivers, such as rainfall and sea levels, are commonly studied by different research communities, which creates another barrier for data exchange. In addition, integrated climate-ocean-land modelling (especially on a local scale) is still a challenging task.

Finally, most of the papers reviewed focus on answering research questions, either related to modelling techniques (e.g. ensemble model studies) or how to use ensemble modelling techniques to simulate past flood events or assess flood risk in specific locations (e.g. ensemble utilisation studies). The application of ensemble flood forecasting to solve a real-world problem was only considered in less than 30% of the studies, most of which are from the more recent review period. The increasing trend in application related studies is a promising sign that more flood practitioners are taking up ensemble forecast as a useful technique. However, the current scope of applications is quite limited, with a focus on flood warning, risk analysis and reservoir operation. The papers reviewed did not include studies on the use of ensemble forecasting in other emerging application areas, such as evacuation planning in emergency management and risk mitigation for humanitarian actions.

Challenges lie ahead to make ensemble flood forecasting a common tool used for flood management or water management in general. First, forecast flood variables need to be extended to include flood inundation depth and extent which are directly related to flood impact and commonly used in the water industry for flood risk/hazard evaluation, as discussed above. Second, forecast lead time needs to be extended to cater for various types of water management tasks. For example, environmental flood allocation in the Murray Darlings Basin in Australia often requires information up to 3 months ahead, necessitating the extension of flood forecasts lead time to sub-seasonal and seasonal scales (i.e. very long range). Long-range forecasts will also improve flood preparedness and mitigation. In addition, effective communication of ensemble forecasting to both scientific and non-scientific communities who are not familiar with the concept is key to encourage the uptake of ensemble flood forecasting in water management, through improved exchange of data, techniques and ideas.

We conclude that in order to encourage the uptake of ensemble flood forecasting in flood management or water management in general, there is a need to not only improve technical aspects of flood forecasting, but also to bridge the gap between different scientific research communities, as well as

between the scientific community and flood forecasters and managers. We hope this paper has contributed towards achieving this goal.

REFERENCES

Addor, N., Jaun, S., Fundel, F., Zappa, M., (2011). An operational hydrological ensemble prediction system for the city of Zurich (Switzerland): skill, case studies and scenarios. Hydrol. Earth Syst. Sci., 15(7): 2327-2347. DOI:10.5194/hess-15-2327-2011

Ajami NK. (2007), "An integrated hydrologic Bayesian multimodel combination framework: confronting input, parameter, and model structural uncertainty in hydrologic prediction". Water Resour Res, 43:1–19. https://doi.org/10.1029/2005WR004745

Ajami, N., Q. Duan, X. Gao, S. Sorooshian, (2006), "Multi-Model Combination techniques for Hydrological Forecasting: Application to Distributed Model Intercomparison Project Results", Journal of Hydrometeorology, 7, 755-768. https://doi.org/10.1175/JHM519.1

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., and Pappenberger, F. (2013). "GloFAS - global ensemble streamflow forecasting and flood early warning." Hydrology and Earth System Sciences, 17(3), 1161-1175. https://doi.org/10.5194/hess-17-1161-2013

Arnal, L., H. L. Cloke, E. Stephens, F. Wetterhall, C. Prudhomme, J. Neumann, B. Krzeminski and F. Pappenberger, (2018) "Skilful seasonal forecasts of streamflow over Europe?", Hydrol. Earth Syst. Sci, **22**, 2057–2072, doi:10.5194/hess-22-2057-2018

Arnal, L., Wetterhall, F. and Pappenberger, F.: Flood forecast decision-making games, ECMWF Newsletter, 152, available online at: https://www.ecmwf.int/en/newsletter/152/news/flood-forecast-decision-making-games, 2017.

Amengual, A., Carrio, D. S., Ravazzani, G., and Homar, V. (2017). "A Comparison of Ensemble Strategies for Flash Flood Forecasting: The 12 October 2007 Case Study in Valencia, Spain." Journal of Hydrometeorology, 18(4), 1143-+. DOI: 10.1175/JHM-D-16-0281.1

Anctil, F., M.-H. Ramos, (2019) "Verification metrics for hydrological ensemble forecasting", in Duan et al. (eds): Handbook of Hydrometeorological Ensemble Forecasting, Springer, https://doi.org/10.1007/978-3-642-39925-1_3

Antonetti, M., C. Horat, I. V. Sideris and M. Zappa (2019). "Ensemble flood forecasting considering dominant runoff processes - Part 1: Set-up and application to nested basins (Emme, Switzerland)." Natural Hazards and Earth System Sciences 19(1): 19-40.

Arnal, L., Anspoks, L., Manson, S., Neumann, J., Norton, T., Stephens, E., Wolfenden, L., and Cloke, H. L. (2019) "Are we talking just a bit of water out of bank? Or is it Armageddon? Front line perspectives on transitioning to probabilistic fluvial flood forecasts in England.", Geosci. Commun. Discuss., https://doi.org/10.5194/gc-2019-18.

Bao, H. J., and Zhao, L. N. (2012). "Development and Application of an Atmospheric-Hydrologic-Hydraulic Flood Forecasting Model Driven by TIGGE Ensemble Forecasts." Acta Meteorologica Sinica, 26(1), 93-102. https://doi.org/10.1007/s13351-012-0109-0

Baran S. and Nemoda D. (2016) "Censored and shifted gamma distribution based EMOS model for probabilistic quantitative precipitation forecasting." Environmetrics, 27:280–292. https://doi.org/10.1002/env.2391

Barthelemy, S., Ricci, S., Le Pape, E., Rochoux, M., Thual, O., Goutal, N., Habert, J., Piacentini, A., Jonville, G., Zaoui, F., and Gouin, P. (2016). "Ensemble-based algorithm for error reduction in hydraulics in the context of flood forecasting." 3rd European Conference on Flood Risk Management, M. Lang, F. Klijn, and P. Samuels, eds. https://doi.org/10.1051/e3sconf/20160718022

Barthelemy, S., Ricci, S., Rochoux, M. C., Le Pape, E., and Thual, O. (2017). "Ensemble-based data assimilation for operational flood forecasting - On the merits of state estimation for 1D hydrodynamic forecasting through the example of the "Adour Maritime" river." Journal of Hydrology, 552, 210-224. https://doi.org/10.1016/j.jhydrol.2017.06.017

Bates, B. C., and E. P. Campbell (2001), "A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall–runoff modelling." Water Resour. Res., 37, 937– 947, doi:10.1029/2000WR900363.

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A. (2017a) "MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data", Hydrol. Earth Syst. Sci., 21, 589–615, https://doi.org/10.5194/hess-21-589-2017.

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., Brocca, L., Pappenberger, F., Huffman, G. J., and Wood, E. F. (2017b): Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling, Hydrol. Earth Syst. Sci., 21, 6201–6217, https://doi.org/10.5194/hess-21-6201-2017, 2017.

Bell, V. A., Davies, H. N., Kay, A. L., Brookshaw A. and Scaife A. A. (2017) "A national-scale seasonal hydrological forecast system: development and evaluation over Britain", Hydrol. Earth Syst. Sci., **21**(9), 4681–4691, doi:10.5194/hess-21-4681-2017

Bellier, J., Zin, I., Siblot, S., and Bontron, G. (2016). "Probabilistic flood forecasting on the Rhone River: evaluation with ensemble and analogue-based precipitation forecasts." 3rd European Conference on Flood Risk Management, M. Lang, F. Klijn, and P. Samuels, eds. https://doi.org/10.1051/e3sconf/20160718011

Bellier, J., Zin, I., and Bontron, G. (2018). "Generating coherent ensemble forecasts after hydrological postprocessing: Adaptations of ECC-based methods." Water Resources Research, 54, 5741– 5762. https://doi.org/10.1029/2018WR022601.

Bennett, J. C., Q. J. Wang, M. Li, D. E. Robertson and A. Schepen (2016) "Reliable long-range ensemble streamflow forecasts: Combining calibrated climate forecasts with a conceptual runoff model and a staged error model", Water Resour. Res., **52**(10), 8238–8259, doi:10.1002/2016WR019193

Bennett JC, Wang QJ, Pokhrel P, Robertson DE. (2014) "The challenge of forecasting high streamflows 1-3 months in advance with lagged climate indices in southeast Australia." Natural Hazards and Earth System Sciences 14:219-233. https://doi.org/10.5194/nhess-14-219-2014

Bennett, J. C., Wang, Q. J., Robertson, D. E., Schepen, A., Li M. and Michael K. (2017) "Assessment of an ensemble seasonal streamflow forecasting system for Australia", Hydrol. Earth Syst. Sci., **21**, 6007–6030, doi:10.5194/hess-21-6007-2017

Berrocal VJ, Raftery AE, Gneiting T. (2007) "Combining spatial statistical and ensemble information in probabilistic weather forecasts." Mon Weather Rev, 135:1386–1402. https://doi.org/10.1175/MWR3341.1

Biancamaria, S., Lettenmaier, D.P. and Pavelsky, T.M. (2016) "The SWOT Mission and Its Capabilities for Land Hydrology" Surv Geophys 37: 307. https://doi.org/10.1007/s10712-015-9346-y

Bierkens, M.F.P. et al. (2015) "Hyper-resolution global hydrological modelling: what is next?" Hydrological Processes, 29(2): 310-320. DOI:10.1002/hyp.10391

Boelee, L., D. Lumbroso, P. Samuels, E. Stephens and H. Cloke (2017) "A review of the understanding of uncertainty in a flood forecasting system and the available methods of dealing with it", E-proceedings of the 37th IAHR World Congress, August 13-18 2017, Kuala Lumpur, Malaysia, available online at: http://eprints.hrwallingford.co.uk/1409/

Bowler, N.E., Pierce, C.E., Seed, A.W. (2006) "STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP." Q. J. R. Meteorol. Soc., 132(620): 2127-2155. DOI:10.1256/qj.04.100

Braud, I., Borga, M., Gourley, J., Hürlimann, M., Zappa, M., and Gallart, F. (2016). Flash floods, hydro-geomorphic response and risk management. Journal of Hydrology, 541, 1-5. https://doi.org/10.1016/j.jhydrol.2016.08.005

Brown, A., J. Demargne, D.-J. Seo, and Y. Liu (2010) The Ensemble Verification System (EVS): A software tool for verifying ensemble forecasts of hydrometeoro-logical and hydrologic variables at discrete locations. Environ. Modell. Software,25, 854–872.

Brown, A., Milton, S., Cullen, M., Golding, B., Mitchell, J., and Shelly, A. (2012) "Unified modeling and prediction of weather and cli- mate: A 25-year journey", B. Am. Meteorol. Soc., 93, 1865–1877. https://doi.org/10.1175/BAMS-D-12-00018.1

Bruen, M., Krahe, P., Zappa, M., Olsson, J., Vehvilainen, B., Kok, K., and Daamen, K. (2010). Visualizing flood forecasting uncertainty: some current European EPS platforms - COST731 working group 3. Atmospheric Science Letters, 11(2), 92-99. https://doi.org/10.1002/asl.258

Bureau of Meteorology (BoM), (2018) Seasonal Streamflow Forecasts: Water Information. (Accessed 24 April 2018), http://www.bom.gov.au/water/ssf/about.shtml

Buizza, R., M. Miller and T. N. Palmer (1999) "Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System", Q. J. R. Meteorol. Soc., 125, 2887-2908. https://doi.org/10.1002/qj.49712556006

Candogan Yossef, N., R. van Beek, A. Weerts, H. Winsemius and M. F. Bierkens (2017) "Skill of a global forecasting system in seasonal ensemble streamflow prediction", *Hydrol. Earth Syst. Sci*, **21**, 4103–4114, doi:10.5194/hess-21-4103-2017

Cane, D., Ghigo, S., Rabuffetti, D., and Milelli, M. (2013). "Real-time flood forecasting coupling different postprocessing techniques of precipitation forecast ensembles with a distributed hydrological model. The case study of May 2008 flood in western Piemonte, Italy." Natural Hazards and Earth System Sciences, 13(2), 211-220. DOI: 10.5194/nhess-13-211-2013

Cao, Q., A. Mehran, F.M. Ralph, and D.P. Lettenmaier (2019) "The role of hydrological initial conditions on Atmospheric River floods in the Russian River basin", Journal of Hydrometeorology, 20, DOI:10.1175/JHM-D-19-0030.1.

Chen, F., W.T. Crow, P. J. Starks, D.N Moriasi (2011) "Improving hydrologic predictions of a catchment model via assimilationof surface soil moisture", Adv. in Water Res., 34: 526–536, https://doi.org/10.1016/j.advwatres.2011.01.011

Chen, H., Yang, D. W., Hong, Y., Gourley, J. J., and Zhang, Y. (2013). "Hydrological data assimilation with the Ensemble Square-Root-Filter: Use of streamflow observations to update model states for real-time flash flood forecasting." Advances in Water Resources, 59, 209-220. https://doi.org/10.1016/j.advwatres.2013.06.010

Chiew, F. H. S. and T. A. McMahon (2002) "Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability", *Hydrol. Sci. J.*, **47**(3), 505–522, doi:10.1080/02626660209492950

Clark MP and Hay LE. (2004) "Use of medium-range numerical weather prediction model output to produce forecastsof streamflow." J Hydrometeorol, 5:15–32.doi:10.1175/1525-7541(2004)005h0015:UOMNWPi2.0.CO;2.37.

Clark, MP, MFP Bierkens, L Samaniego, R Woods, R Uijlenhoet, K Bennett, V Pauwels, X Cai, AW Wood, and C Peters-Lidard. (2017). "The evolution of process-based hydrologic models: Historical challenges

and the collective quest for physical realism." Hydrology and Earth System Sciences. 21. 3427-3440. https://doi.org/10.5194/hess-21-3427-2017

Cloke H. and Pappenberger F. (2009) "Ensemble flood forecasting: a review." J Hydrol, 375:613–626. https://doi.org/10.1016/j.jhydrol.2009.06.005

Cloke, H., F. Pappenberger, J. Thielen and V. Thiemig (2013) "Operational European Flood Forecasting." *Environmental Modelling*, J. Wainwright. and M. Mulligan, Eds., John Wiley and Sons Ltd, Chichester, UK, 415–434. https://doi.org/10.1002/9781118351475.ch25

V. Cooten, S., K.E. Kelleher, K. Howard, J. Zhang, J.J. Gourley, J.S. Kain, K. Nemunaitis-Monroe, Z. Flamig, H. Moser, A. Arthur, C. Langston, R. Kolar, Y. Hong, K. Dresback, E. Tromble, H. Vergara, R.A. Luettich, B. Blanton, H. Lander, K. Galluppi, J.P. Losego, C.A. Blain, J. Thigpen, K. Mosher, D. Figurskey, M. Moneypenny, J. Blaes, J. Orrock, R. Bandy, C. Goodall, J.G. Kelley, J. Greenlaw, M. Wengren, D. Eslinger, J. Payne, G. Olmi, J. Feldt, J. Schmidt, T. Hamill, R. Bacon, R. Stickney, and L. Spence, (2011) The CI-FLOW Project: A System for Total Water Level Prediction from the Summit to the Sea. Bull. Amer. Meteor. Soc., 92, 1427–1442, https://doi.org/10.1175/2011BAMS3150.1

Copernicus Climate Change Service (C3S) (2017) "ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate." Copernicus Climate Change Service Climate Data Store (CDS), date of access 01 Nov. 2019. https://cds.climate.copernicus.eu/cdsapp#!/home

Cloke, H.L, Pappenberger, F., Jan van Andel, S., Thielen, J., Ramos, M.-H.(eds) (2013) "Hydrological ensemble prediction systems." Hydrological processes special issue, Hydrol Process 2013, 27:1–4. doi:10.1002/hyp.9679.

Chu, H., W. Wu, Q. J. Wang, R. Nathan and J. Wei (2019) "An ANN-based emulation modelling framework for flood inundation modelling: application, challenges and future directions", Environmental Modelling and Software, 124, https://doi.org/10.1016/j.envsoft.2019.104587

Copernicus (2018) SWICCA | Service for Water Indicators in Climate Change Adaptation. Accessed 12 January 2018, http://swicca.climate.copernicus.eu/

Crochemore, L., M.-H. Ramos and F. Pappenberger (2016) "Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts", *Hydrol. Earth Syst. Sci*, **20**, 3601–3618, doi:10.5194/hess-20-3601-2016

Dahl, N., and Xue, M. (2016). "Prediction of the 14 June 2010 Oklahoma City Extreme Precipitation and Flooding Event in a Multiphysics Multi-Initial-Conditions Storm-Scale Ensemble Forecasting System." Weather and Forecasting, 31(4), 1215-1246. https://doi.org/10.1175/WAF-D-15-0116.1

Das, A. K., and Kaur, S. (2013). "Performance of IMD multi-model ensemble and WRF (ARW) based sub-basin wise rainfall forecast for Mahanadi basin during flood season 2009 and 2010." Mausam, 64(4), 625-644.

Dasgupta, A., Grimaldi, S., Ramsankaran, R., Pauwels, V.R.N., Walker, J.P., Chini, M., Hostache, R. and Matgen, P. (2019). Flood Mapping Using Synthetic Aperture Radar Sensors From Local to Global Scales.

In Global Flood Hazard (eds G.J.-P. Schumann, P.D. Bates, H. Apel and G.T. Aronica). doi:10.1002/9781119217886.ch4.

Day, Gerald N. (1985) Extended Streamflow Forecasting Using NWSRFS. *Journal of Water Resources Planning and Management-asce*, 111(2), https://doi.org/10.1061/(ASCE)0733-9496(1985)111:2(157)

Delaney, C, Mendoza, J. R., Hartman, R. K. and Hecht, C. W. (in review), Forecast Informed Reservoir Operations Using Ensemble Streamflow Predictions for a Multi-Purpose Reservoir in Northern California, Water Resources Research

Demargne, J., L. Wu, S. K. Regonda, J. D. Brown, H. Lee, M. He, D.-J. Seo, R. Hartman, H. D. Herr, M. Fresch, J. Schaake and Y. Zhu (2014) "The Science of NOAA's Operational Hydrologic Ensemble Forecast Service", *Bull. Am. Meteorol. Soc.*, **95**(1), 79–98, doi:10.1175/BAMS-D-12-00081.1

Dietrich, J., Denhard, M., and Schumann, A. H. (2009). "Can ensemble forecasts improve the reliability of flood alerts?" Journal of Flood Risk Management, 2(4), 232-242.   DOI: 10.1111/j.1753-318X.2009.01039.x

Dietrich, J., Schumann, A. H., Redetzky, M., Walther, J., Denhard, M., Wang, Y., Pfutzner, B., and Buttner, U. (2009). "Assessing uncertainties in flood forecasts for decision making: prototype of an operational flood management system integrating ensemble predictions." Natural Hazards and Earth System Sciences, 9(4), 1529-1540.   DOI: 10.5194/nhess-9-1529-2009

Dietrich, J., Trepte, S., Wang, Y., Schumann, A. H., Voss, F., Hesser, F. B., and Denhard, M. (2008). "Combination of different types of ensembles for the adaptive simulation of probabilistic flood forecasts: hindcasts for the Mulde 2002 extreme event." Nonlinear Processes in Geophysics, 15(2), 275-286.   DOI: 10.5194/npg-15-275-2008

Doycheva, K., Horn, G., Koch, C., Schumann, A., and Konig, M. (2017). "Assessment and weighting of meteorological ensemble forecast members based on supervised machine learning with application to runoff simulations and flood warning." Advanced Engineering Informatics, 33, 427-439. https://doi.org/10.1016/j.aei.2016.11.001

Duan, Q., N. Ajami, X. Gao, S. Sorooshian, (2007), "Multi-model Hydrologic Ensemble Predictions Using Bayesian Model Averaging", Advances in Water Resources, doi:10.1016/j.advwatres.2006.11.014

Duan, Q., F. Pappenber, A. Wood, H. Cloke, J. Schaake (Eds), (2019), Handbook of Hydrometeorological Ensemble Forecasting, Springer, 1528p., https://doi.org/10.1007/978-3-642-39925-1

Dutra, E., Wetterhall, F., Di Giuseppe, F., Naumann, G., Barbosa, P., Vogt, J., Pozzi, W., and Pappenberger, F. (2014) "Global meteorological drought – Part 1: Probabilistic monitoring", Hydrol. Earth Syst. Sci., 18, 2657–2667, https://doi.org/10.5194/hess-18-2657-2014.

Ebert EE. (2001) "Ability of a poor man's ensemble to predict the probability and distribution of precipitation." Mon Weather Rev, 129:2461–2480. https://doi.org/10.1175/1520-0493(2001)129<2461:AOAPMS>2.0.CO;2

ECMWF, (2020) ECMWF works with universities to support response to tropical cyclone Idai, https://www.ecmwf.int/en/newsletter/160/news/ecmwf-works-universities-support-response-tropical-cyclone-idai, Last access 10 February 2020

Einfalt, T., Hatzfeld, F., Wagner, A., Seltmann, J., Castro, D. and Frerichs, S. (2009) "URBAS: forecasting and management of flash floods in urban areas". Urban Water Journal, 6(5): 369-374. DOI:10.1080/15730620902934819

Emerton, R., H. L. Cloke, E. M. Stephens, E. Zsoter, S. J. Woolnough and F. Pappenberger (2017) "Complex picture for likelihood of ENSO-driven flood hazard", *Nat. Commun*., **8**, 14796, doi:10.1038/ncomms14796

Emerton, R. E., E. M. Stephens, F. Pappenberger, T. C. Pagano, A. H. Weerts, A. W. Wood, P. Salamon, J. D. Brown, N. Hjerdt, C. Donnelly, C. A. Baugh and H. L. Cloke (2016) "Continental and Global Scale Flood Forecasting Systems", *WIREs Water*, 3 (3), 391-418, doi:10.1002/wat2.1137

Emerton, R., E. Zsoter, L. Arnal, H. L. Cloke, D. Muraro, C. Prudhomme, E. M. Stephens, P. Salamon and F. Pappenberger (2018) "Developing a global operational seasonal hydro- meteorological forecasting system: GloFAS-Seasonal v1.0", *Geosci. Model Dev*., **11**(8), 3327– 3346, doi:10.5194/gmd-11-3327-2018

Fan, F. M., Collischonn, W., Quiroz, K. J., Sorribas, M. V., Buarque, D. C., and Siqueira, V. A. (2016). "Flood forecasting on the Tocantins River using ensemble rainfall forecasts and real-time satellite rainfall estimates." Journal of Flood Risk Management, 9(3), 278-288. https://doi.org/10.1111/jfr3.12177

Fleming, S. W., Bourdin, D. R., Campbell, D., Stull, R. B., and Gardner, T. (2015). "DEVELOPMENT AND OPERATIONAL TESTING OF A SUPER-ENSEMBLE ARTIFICIAL INTELLIGENCE FLOOD-FORECAST MODEL FOR A PACIFIC NORTHWEST RIVER." Journal of the American Water Resources Association, 51(2), 502-512. https://doi.org/10.1111/jawr.12259

Franz, K., Ajami, N., Schaake, J., and Buizza, R. (2005), "Hydrologic ensemble prediction experiment focuses on reliable forecasts", Eos Trans. AGU, 86(25), 239– 239, doi:10.1029/2005EO250004.

Étienne Gaborit, Vincent Fortin, Xiaoyong Xu, Frank Seglenieks, Bryan Tolson, Lauren M. Fry, Tim Hunter, François Anctil6, and Andrew D. Gronewold (2017) "A hydrological prediction system based on the SVS land-surface scheme: efficient calibration of GEM-Hydro for streamflow simulation over the Lake Ontario basin." Hydrol. Earth Syst. Sci., 21(9): 4825-4839. DOI:10.5194/hess-21-4825-2017

Glahn, H.R., D.A. Lowry, (1973) "The use of model output statistics (MOS) in objective weather forecasting." J. Appl. Meteorol., 11:1203–1211. https://doi.org/10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2

GloFAS, (2020) Global Flood Awareness System Partners, Copernicus Emergency Management Service: https://www.globalfloods.eu/get-involved/partners/, Last access 10 February 2020

Golding, B., Roberts, N., Leoncini, G., Mylne, K., and Swinbank, R. (2016). "MOGREPS-UK Convection-Permitting Ensemble Products for Surface Water Flood Forecasting: Rationale and First Results." Journal of Hydrometeorology, 17(5), 1383-1406. https://doi.org/10.1175/JHM-D-15-0083.1

Gronewold, A.D., Fortin, V. (2012) "Advancing Great Lakes Hydrological Science Through Targeted Binational Collaborative Research." Bulletin of the American Meteorological Society, 93(12): 1921-1925. DOI:10.1175/bams-d-12-00006.1

Gronewold, A.D., Fortin, V., Caldwell, R., Noel, J. (2018) "Resolving Hydrometeorological Data Discontinuities along an International Border." Bulletin of the American Meteorological Society, 99(5): 899-910. DOI:10.1175/bams-d-16-0060.1

Guimaraes Nobre, G., B. Jongman, J. Aerts and P. J. Ward (2017) "The role of climate variability in extreme floods in Europe", *Environ. Res. Lett.*, **12**(8), 084012, doi:10.1088/1748- 9326/aa7c22

Hall, J. and Solomatine, D. (2008) A framework for uncertainty analysis in flood risk management decisions, International Journal of River Basin Management, 6:2, 85-98, DOI: 10.1080/15715124.2008.9635339

Hamill TM, Whitaker JS, Wei X. (2004) "Ensemble re-forecasting: improving medium-range forecast skill using retrospective forecasts." Bull Am Meteorol Soc. , 132:3825–3830. https://doi.org/10.1175/1520-0493(2004)132<1434:ERIMFS>2.0.CO;2

Hapuarachchi, H., Wang, Q., Pagano, T. (2011) "A review of advances in flash flood forecasting." Hydrological Processes, 25(18): 2771-2784. https://doi.org/10.1002/hyp.8040

He, Y., Wetterhall, F., Cloke, H.L., Pappenberger, F., Wilson, M., Freer,, J., McGregor, G. (2009) "Tracking the uncertainty in flood alerts driven by grand ensemble weather predictions", Meteorological Applications, 16, 91-101. https://doi.org/10.1002/met.132

He, Y., Wetterhall, F., Bao, H. J., Cloke, H., Li, Z. J., Pappenberger, F., Hu, Y. Z., Manful, D., and Huang, Y. C. (2010). "Ensemble forecasting using TIGGE for the July-September 2008 floods in the Upper Huai catchment: a case study." Atmospheric Science Letters, 11(2), 132-138. https://doi.org/10.1002/asl.270

Hemri, S., and Klein, B. (2017). "Analog-based postprocessing of navigation-related hydrological ensemble forecasts." Water Resources Research, 53, 9059– 9077. https://doi.org/10.1002/2017WR020684.

HEPEX: HEPEX Games and Resources, 2020, Available online at: https://hepex.inrae.fr/resources/hepex-games/.

Hopson, T. M., and Webster, P. J. (2010). "A 1-10-Day Ensemble Forecasting Scheme for the Major River Basins of Bangladesh: Forecasting Severe Floods of 2003-07." Journal of Hydrometeorology, 11(3), 618-641. https://doi.org/10.1175/2009JHM1006.1

Hsiao, L. F., Yang, M. J., Lee, C. S., Kuo, H. C., Shih, D. S., Tsai, C. C., Wang, C. J., Chang, L. Y., Chen, D. Y. C., Feng, L., Hong, J. S., Fong, C. T., Chen, D. S., Yeh, T. C., Huang, C. Y., Guo, W. D., and Lin, G. F. (2013). "Ensemble forecasting of typhoon rainfall and floods over a mountainous watershed in Taiwan." Journal of Hydrology, 506, 55-68. https://doi.org/10.1016/j.jhydrol.2013.08.046

INSPIRE, (accessed) 1 December 2019, https://inspire.ec.europa.eu/

Kavetski, D., G. Kuczera, and S. W. Franks (2006), "Bayesian analysis of input uncertainty in hydrological modeling:2. Application", Water Resour. Res.,42, W03408, doi:10.1029/2005WR004376

Kauffeldt A. (2014) Disinformative and Uncertain Data in Global Hydrology: Challenges for Modelling and Regionalisation [dissertation]. Uppsala: Uppsala Universitet; 2014, 79 p. 30.

Kauffeldt, A., Halldin, S., Pappenberger, F., Wetterhall, F., Xu, C.-Y., and Cloke, H. L. (2015), "Imbalanced land surface water budgets in a numerical weather prediction system." *Geophys. Res. Lett.*, 42, 4411–4417. doi: 10.1002/2015GL064230.

Kim, S., Tachikawa, Y., Sayama, T., and Takara, K. (2009). "Ensemble flood forecasting with stochastic radar image extrapolation and a distributed hydrologic model." Hydrological Processes, 23(4), 597-611. https://doi.org/10.1002/hyp.7188

Kimura, N., Hsu, M. H., Tsai, M. Y., Tsao, M. C., Yu, S. L., and Tai, A. (2016). "A river flash flood forecasting model coupled with ensemble Kalman filter." Journal of Flood Risk Management, 9(2), 178-192. https://doi.org/10.1111/jfr3.12128

Krishnamurti TN, Kishtawal CM, LaRow TE, Bachiochi DR, Zhang Z, Williford CE, Gadgil S, Surendran S. (1999) "Improved weather and seasonal climate forecasts from multimodel superensemble." Science, 285:1548–1550. DOI: 10.1126/science.285.5433.1548

Komma, J., Bloschl, G., and Reszler, C. (2008). "Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting." Journal of Hydrology, 357(3-4), 228-242. DOI:https://doi.org/10.1016/j.jhydrol.2008.05.020

Komma, J., Reszler, C., Bloschl, G., and Haiden, T. (2007). "Ensemble prediction of floods - catchment non-linearity and forecast probabilities." Natural Hazards and Earth System Sciences, 7(4), 431-444.

Kovalchuk, S. V., Krikunov, A. V., Knyazkov, K. V., and Boukhanovsky, A. V. (2017). "Classification issues within ensemble-based simulation: application to surge floods forecasting." Stochastic Environmental Research and Risk Assessment, 31(5), 1183-1197. https://doi.org/10.5194/nhess-7-431-2007

Kuchment, L. S., and Gelfan, A. N. (2009). "A study of effectiveness of the ensemble long-term forecasts of spring floods issued with physically based models of the river runoff formation." Russian Meteorology and Hydrology, 34(2), 100-109. DOI:10.3103/S106837390902006X

Kuczera, G. (1983), "Improved parameter inference in catchment models. 1. Evaluating parameter uncertainty", Water Resour. Res. 19(5), 1151– 1172. https://doi.org/10.1029/WR019i005p01151

Lee, C. S., Ho, H. Y., Lee, K. T., Wang, Y. C., Guo, W. D., Chen, D. Y. C., Hsiao, L. F., Chen, C. H., Chiang, C. C., Yang, M. J., and Kuo, H. C. (2013). "Assessment of sewer flooding model based on ensemble quantitative precipitation forecast." Journal of Hydrology, 506, 101-113. https://doi.org/10.1016/j.jhydrol.2012.09.053

Lee, H., D.-J. Seo, V. Koren (2011) "Assimilation of streamflow and in situ soil moisture data into operational distributed hydrologic models: Effects of uncertainties in the data and initial model soil moisture states", Adv. in Water Res., 34: 1597–1615, https://doi.org/10.1016/j.advwatres.2011.08.012

Leonard, M., Seth Westra, Aloke Phatak, Martin Lambert, Bart van den Hurk, Kathleen McInnes, James Risbey, Sandra Schuster, Doerte Jakob and Mark Stafford-Smith (2014) "A compound event framework for understanding extreme impacts." WIREs Climate Change, 5(1): 113-128. DOI:10.1002/wcc.252

Leoncini, G., Plant, R. S., Gray, S. L., and Clark, P. A. (2013). "Ensemble forecasts of a flood-producing storm: comparison of the influence of model-state perturbations and parameter modifications." Quarterly Journal of the Royal Meteorological Society, 139(670), 198-211. https://doi.org/10.1002/qj.1951

Li, H., L. Luo, E. F. Wood, and J. Schaake, (2009) "The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting", J. Geophys. Res., 114, D04114, doi:10.1029/2008JD010969.

Li, W., Duan, Q., Miao, C., Ye, A., Gong, W., Di, Z. (2017). "A review on statistical postprocessing methods for hydrometeorological ensemble forecasting", WIREs Water, 4:e1246, DOI: 10.1002/wat2.1246

Li, Y., Grimaldi, S., Walker, J., Pauwels, V. (2016) "Application of Remote Sensing Data to Constrain Operational Rainfall-Driven Flood Forecasting: A Review." Remote Sensing, 8(6): 456. https://doi.org/10.3390/rs8060456

Yuan Li, Stefania Grimaldi, Valentijn R.N. Pauwels, Jeffrey P. Walker (2018) "Hydrologic model calibration using remotely sensed soil moisture and discharge measurements: The impact on predictions at gauged and ungauged locations", Journal of Hydrology, Volume 557, Pages 897-909, ISSN 0022-1694, https://doi.org/10.1016/j.jhydrol.2018.01.013.

Liechti, K., Panziera, L., Germann, U., and Zappa, M. (2013) "The potential of radar-based ensemble forecasts for flash-flood early warning in the southern Swiss Alps." Hydrology and Earth System Sciences, 17(10), 3853-3869. https://doi.org/10.5194/hess-17-3853-2013

Lins, Harry F., (2008) "Challenges to Hydrological Observations", World Meteorological Organisation Bulletin, Vol 57 (1), https://public.wmo.int/en/bulletin/challenges-hydrological-observations

Liu, L., Gao, C., Xuan, W. D., and Xu, Y. P. (2017). "Evaluation of medium-range ensemble flood forecasting based on calibration strategies and ensemble methods in Lanjiang Basin, Southeast China." Journal of Hydrology, 554, 233-250. https://doi.org/10.1016/j.jhydrol.2017.08.032

Liu, P., Lin, K. R., and Wei, X. J. (2015). "A two-stage method of quantitative flood risk analysis for reservoir real-time operation using ensemble-based hydrologic forecasts." Stochastic Environmental Research and Risk Assessment, 29(3), 803-813. https://doi.org/10.1007/s00477-014-0986-0

Lorenz, E. N. (1963) "Deterministic Nonperiodic Flow", *J. Atmos. Sci.*, 20(2), 130–141, doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2

Lorenz, E. N. (1993) *The essence of chaos*, University of Washington Press, 240pp.

Mackey, B. P., and Krishnamurti, T. N. (2001). "Ensemble forecast of a typhoon flood event." Weather and Forecasting, 16(4), 399-415. https://doi.org/10.1175/1520-0434(2001)016<0399:EFOATF>2.0.CO;2

Maraun, D., et al. (2010), "Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user", Rev. Geophys., 48, RG3003, doi:10.1029/2009RG000314.

Marty, R., Zin, I., and Obled, C. (2013). "Sensitivity of hydrological ensemble forecasts to different sources and temporal resolutions of probabilistic quantitative precipitation forecasts: flash flood case studies in the Cevennes-Vivarais region (Southern France)." Hydrological Processes, 27(1), 33-44. DOI:10.1002/hyp.9543

Maurer, A. Kumar and Lettenmaier, D. P. (2002) "Long-range experimental hydrologic forecasting for the eastern United States", *J. Geophys. Res.*, **107**(D20), 4429, doi:10.1029/2001JD000659

Maurer, E. P., and Lettenmaier, D. P. (2003) "Predictability of seasonal runoff in the Mississippi River basin", J. Geophys. Res., 108(D16), 8607, doi:10.1029/2002JD002555.

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N. E. C., Franz, T. E., Shi, J., Gao, H., and Wood, E. F (2017) "The future of Earth observation in hydrology", Hydrol. Earth Syst. Sci., 21, 3879–3914, https://doi.org/10.5194/hess-21-3879-2017.

McEnery, J., J. Ingram, Q. Duan, T. Adams, and L. Anderson (2005) "NOAA'S ADVANCED HYDROLOGIC PREDICTION SERVICE." Bull. Amer. Meteor. Soc., 86, 375–386, https://doi.org/10.1175/BAMS-86-3-375

McSharry, P. E., Little, M. A., Rodda, H. J. E., and Rodda, J. (2013). "Quantifying flood risk of extreme events using density forecasts based on a new digital archive and weather ensemble predictions." Quarterly Journal of the Royal Meteorological Society, 139(671), 328-333. https://doi.org/10.1002/qj.2136

Meißner, D., Klein, B. and Ionita, M. (2017) "Development of a monthly to seasonal forecast framework tailored to inland waterway transport in central Europe", *Hydrol. Earth Syst. Sci.*, **21**(12), 6401–6423, doi:10.5194/hess-21-6401-2017.

Meißner D. and Klein, B. (2015) "Probabilistic Shipping Forecast". In: Duan Q., Pappenberger F., Thielen J., Wood A., Cloke H., Schaake J. (eds) Handbook of Hydrometeorological Ensemble Forecasting. Springer, Berlin, Heidelberg.   DOI: 10.1007/978-3-642-40457-3_58-1

Mendoza, PA, AW Wood, EA Clark, E Rothwell, MP Clark, B Nijssen, LD Brekke, and JR Arnold (2017) "An intercomparison of approaches for improving predictability in operational seasonal streamflow forecasts", *Hydrol. Earth Syst. Sci.,* 21, 3915–3935, 2017. https://doi.org/10.5194/hess-21-3915-2017

Michaelides, S., V. Levizzani, E. Anagnostou, P. Bauer, T. Kasparis, J.E. Lane, (2009) "Precipitation: Measurement, remote sensing, climatology and modelling", Atmospheric Research, Volume 94, Issue 4, 512-533, https://doi.org/10.1016/j.atmosres.2009.08.017.

Mo, K. C., D. P. Lettenmaier, K. C. Mo and D. P. Lettenmaier (2014) "Hydrologic Prediction over the Conterminous United States Using the National Multi-Model Ensemble", *J. Hydrometeorol.*, **15**(4), 1457–1472, doi:10.1175/JHM-D-13-0197.1

Montani, A., Marsigli, C., Nerozzi, F., Paccagnella, T., Tibaldi, S., and Buizza, R. (2003). "The Soverato flood in Southern Italy: performance of global and limited-area ensemble forecasts." Nonlinear Processes in Geophysics, 10(3), 261-274.   DOI: 10.5194/npg-10-261-2003

Moradkhani, H., G.S. Nearing, P. Abbaszadeh, and S. Pathiraja (2019) "Fundamentals of data assimilation and theoretical advances", in Duan et al. (eds): Handbook of Hydrometeorological Ensemble Forecasting, Springer, https://doi.org/10.1007/978-3-642-39925-1_30.

NASEM (National Academies of Sciences, Engineering, and Medicine) (2018) Review of the New York City Department of Environmental Protection Operations Support Tool for Water Supply. Washington, DC: The National Academies Press. https://doi.org/10.17226/25218.Nayak, M. A., Herman, J. D., and Steinschneider, S. (2018). "Balancing Flood Risk and Water Supply in California: Policy Search Integrating Short-Term Forecast Ensembles With Conjunctive Use." Water Resources Research, 54(10), 7557-7576. https://doi.org/10.1029/2018WR023177

Neelz, S., Pender, G. (2013) "Benchmarking the Latest Generation of 2D Hydraulic Modelling Packages." DEFRA/Environment Agency, UK. http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/SC120002_Benchmarking_2D_hydraulic_models_Report.sflb.ashx. Accessed online 01 Nov. 2019.

Nester, T., Komma, J., Viglione, A., and Bloschl, G. (2012). "Flood forecast errors and ensemble spread-A case study." Water Resources Research, 48. https://doi.org/10.1029/2011WR011649

O'Connell, P.E., Clarke, R.T. (1981) "Adaptive hydrological forecasting—a review / Revue des méthodes de prévision hydrologique ajustables." Hydrological Sciences Bulletin, 26(2): 179-205. DOI:10.1080/02626668109490875

Pagano TC, Shrestha DL, Wang QJ, Robertson D, Hapuarachchi P. (2013) "Ensemble dressing for hydrological applications." Hydrol Process, 27:106–116. doi:10.1002/hyp.9313.

Palmer, T. and Webster, P. (1993) "Towards a unified approach to climate and weather prediction", in: Proceedings of 1st Demetra Conference on Climate Change, 1993.

Pappenberger, F., K. J. Beven, N. M. Hunter, P. D. Bates, B. T. Gouweleeuw, J. Thielen and A. P. J. de Roo, (2005) "Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS)", Hydrology and Earth System Sciences, 9 (4), 381-393. https://doi.org/10.5194/hess-9-381-2005

Pappenberger, F., Cloke, H. L., Parker, D. J., Wetterhall, F., Richardson, D. S., and Thielen, J. (2015) "The Monetary Benefit of Early Flood Warnings in Europe", Environmental Science and Policy, 51, 278–91, https://doi.org/10.1016/j.envsci.2015.04.016.

Pappenberger, F., Pagano, T. C., Brown, J. D., Alfieri, L., Lavers, D. A., Berthet, L., et al. (2016) Hydrological ensemble prediction systems around the globe. In Q. Duan, F. Pappenberger, J. Thielen, A. Wood, H. L. Cloke, and J. C. Schaake (Eds.), Handbook of hydrometeorological ensemble forecasting (p. (35 pp.). https://doi.org/10.1007/978-3-642-40457-3_47-1

Prudhomme, C., J. Hannaford, S. Harrigan, D. Boorman, J. Knight, V. Bell, C. Jackson, C. Svensson, S. Parry, N. Bachiller-Jareno, H. Davies, R. Davis, J. Mackay, A. McKenzie, A. Rudd, K. Smith, J. Bloomfield, R. Ward and A. Jenkins (2017) "Hydrological Outlook UK: an operational streamflow and groundwater level forecasting system at monthly to seasonal time scales", *Hydrol. Sci. J.*, **62**(16), 2753–2768, doi:10.1080/02626667.2017.1395032

Pietroniro, A., V. Fortin, N. Kouwen, C. Neal, R. Turcotte, B. Davison, D. Verseghy, E. D. Soulis, R. Caldwell, N. Evora, and P. Pellerin (2007) "Development of the MESH modelling system for hydrological ensemble forecasting of the Laurentian Great Lakes at the regional scale." Hydrol. Earth Syst. Sci., 11(4): 1279-1294. DOI:10.5194/hess-11-1279-2007

Raftery AE, Gneiting T, Balabdaoui F, Polakowski M. (2005) "Using Bayesian model averaging to calibrate forecast ensembles." Mon Weather Rev,133:1155–1174. https://doi.org/10.1175/MWR2906.1

Ramos, M. H., Bartholmes, J., and Thielen-del Pozo, J. (2007). "Development of decision support products based on ensemble forecasts in the European flood alert system." Atmospheric Science Letters, 8(4), 113-119. https://doi.org/10.1002/asl.161

Rossa, A., Liechti, K., Zappa, M., Bruen, M., Germann, U., Haase, G., et al. (2011). The COST 731 Action: a review on uncertainty propagation in advanced hydro-meteorological forecast systems. Atmospheric Research, 100(2-3), 150-167. https://doi.org/10.1016/j.atmosres.2010.11.016

Ravazzani, G., Amengual, A., Ceppi, A., Homar, V., Romero, R., Lombardi, G., and Mancini, M. (2016). "Potentialities of ensemble strategies for flood forecasting over the Milano urban area." Journal of Hydrology, 539, 237-253. https://doi.org/10.1016/j.jhydrol.2016.05.023

Roundy, J.K., Q. Duan, J.C. Schaake (2019) "Hydrological predictability, scales, and uncertainty issues", in Duan et al. (eds): Handbook of Hydrometeorological Ensemble Forecasting, Springer, https://doi.org/10.1007/978-3-642-39925-1_8

Saleh, F., Ramaswamy, V., Wang, Y., Georgas, N., Blumberg, A., and Pullen, J. (2017). "A multi-scale ensemble-based framework for forecasting compound coastal-riverine flooding: The Hackensack-Passaic watershed and Newark Bay." Advances in Water Resources, 110, 371-386. https://doi.org/10.1016/j.advwatres.2017.10.026

Scheuerer M. (2014) "Probabilistic quantitative precipitation forecasting using ensemble model output statistics." Q J Roy Meteor Soc, 140:1086–1096. https://doi.org/10.1002/qj.2183

Schulte, J. A. (2017). "Sub-Ensemble Coastal Flood Forecasting: A Case Study of Hurricane Sandy." Journal of Marine Science and Engineering, 5(4). https://doi.org/10.3390/jmse5040059

Schaake, J., Franz, K., Bradley, A., and Buizza, R. (2006) "The Hydrologic Ensemble Prediction EXperiment (HEPEX)", Hydrol. Earth Syst. Sci. Discuss., 3, 3321–3332, https://doi.org/10.5194/hessd-3-3321-2006.

Shamir, E., Lee, B. J., Bae, D. H., and Georgakakos, K. P. (2010). "Flood Forecasting in Regulated Basins Using the Ensemble Extended Kalman Filter with the Storage Function Method." Journal of Hydrologic Engineering, 15(12), 1030-1044. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000282

Shaw E, Beven K, Chappell NA, Lamb R. (2011) Hydrology in Practice. 4th ed. Oxfordshire, England: Spon Press; 543 p. 31.

Shen, C., Laloy, E., Elshorbagy, A., Albert, A., Bales, J., Chang, F.-J., Ganguly, S., Hsu, K.-L., Kifer, D., Fang, Z., Fang, K., Li, D., Li, X., and Tsai, W.-P. (2018) "HESS Opinions: Incubating deep-learning-powered hydrologic science advances as a community", Hydrol. Earth Syst. Sci., 22, 5639–5656, https://doi.org/10.5194/hess-22-5639-2018, 2018. https://doi.org/10.5194/hess-22-5639-2018

Shi, H. Y., Li, T. J., Liu, R. H., Chen, J., Li, J. Y., Zhang, A., and Wang, G. Q. (2015). "A service-oriented architecture for ensemble flood forecast from numerical weather prediction." Journal of Hydrology, 527, 933-942. https://doi.org/10.1016/j.jhydrol.2015.05.056

Shukla, S. and D. P. Lettenmaier (2011) "Seasonal hydrologic prediction in the United States: understanding the role of initial hydrologic conditions and seasonal climate forecast skill", Hydrol. Earth Syst. Sci., 15, 3529-3538, doi:10.5194/hess-15-3529-2011

Shukla, S., J. Sheffield, E. F. Wood and D. P. Lettenmaier (2013) "On the sources of global land surface hydrologic predictability", *Hydrol. Earth Syst. Sci.*, **17**, 2781-2796, http://dx.doi.org/10.5194/hess-17-2781-2013.

Siddique, R. and A. Mejia (2017) Ensemble Streamflow Forecasting across the U.S. Mid-Atlantic Region with a Distributed Hydrological Model Forced by GEFS Reforecasts. J. Hydrometeor., 18, 1905–1928, https://doi.org/10.1175/JHM-D-16-0243.1

Siqueira, V. A., Collischonn, W., Fan, F. M., and Chou, S. C. (2016). "Ensemble flood forecasting based on operational forecasts of the regional Eta EPS in the Taquari-Antas basin." Rbrh-Revista Brasileira De Recursos Hidricos, 21(3). http://dx.doi.org/10.1590/2318-0331.011616004

Srikanthan, R., Amirthanathan, G., Kuczera, G. (2008) "Application of Ensemble Kalman Filter for flood forecasting in Australian rivers." Australasian Journal of Water Resources, 12(3): 245-255. DOI:10.1080/13241583.2008.11465351

Teng, J., A. J. Jakeman, J. Vaze, B. F. W. Croke, D. Dutta and S. Kim (2017) "Flood inundation modelling: A review of methods, recent advances and uncertainty analysis", Environmental Modelling and Software, 90, 201-216, https://doi.org/10.1016/j.envsoft.2017.01.006

Thiemig, V., Bisselink, B., Pappenberger, F., and Thielen, J. (2015). "A pan-African medium-range ensemble flood forecast system." Hydrology and Earth System Sciences, 19(8), 3365-3385. https://doi.org/10.5194/hess-19-3365-2015

Thiemig, V., Pappenberger, F., Thielen, J., Gadain, H., de Roo, A., Bodis, K., Del Medico, M., and Muthusi, F. (2010). "Ensemble flood forecasting in Africa: a feasibility study in the Juba-Shabelle river basin." Atmospheric Science Letters, 11(2), 123-131. https://doi.org/10.1002/asl.266

Tiwari, M.K., Chatterjee, C.(2010) "Uncertainty assessment and ensemble flood forecasting using bootstrap based artificial neural networks (BANNs)." Journal of Hydrology, 382(1): 20-33. DOI:https://doi.org/10.1016/j.jhydrol.2009.12.013

Todini E. (2013) "From HUP to MCP: analogies and extended performances." J Hydrol, 477:33–42. https://doi.org/10.1016/j.jhydrol.2012.10.037

Ushiyama, T., Sayama, T., Tatebe, Y., Fujioka, S., and Fukami, K. (2014). "Numerical Simulation of 2010 Pakistan Flood in the Kabul River Basin by Using Lagged Ensemble Rainfall Forecasting." Journal of Hydrometeorology, 15(1), 193-211. DOI:10.1175/jhm-d-13-011.1

Verbunt, M., Walser, A., Gurtz, J., Montani, A., and Schar, C. (2007). "Probabilistic flood forecasting with a limited-area ensemble prediction system: Selected case studies." Journal of Hydrometeorology, 8(4), 897-909. DOI:10.1175/JHM594.1

Verkade, J. S. and Werner, M. G. F. (2011) "Estimating the benefits of single value and probability forecasting for flood warning", Hydrol. Earth Syst. Sci., 15, 3751-3765, https://doi.org/10.5194/hess-15-3751-2011.

Vincendon, B., Ducrocq, V., Nuissier, O., and Vie, B. (2011). "Perturbation of convection-permitting NWP forecasts for flash-flood ensemble forecasting." Natural Hazards and Earth System Sciences, 11(5), 1529-1544. https://doi.org/10.5194/nhess-11-1529-2011

Vitart, F., Buizza, R., Alonso Balmaseda, M., Balsamo, G., Bidlot, J. R., Bonet, A., Fuentes, M., Hofstadler, A., Molteni, F., and Palmer, T. N. (2008) "The new VarEPS-monthly forecasting system: A first step towards seamless prediction", Q. J. Roy. Meteor. Soc., 134, 1789–1799. https://doi.org/10.1002/qj.322

Vionnet, V., Fortin, V., Gaborit, E., Roy, G., Abrahamowicz, M., Gasset, N., and Pomeroy, J. W. (2019) "High-resolution hydrometeorological modelling of the June 2013 flood in southern Alberta, Canada", Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-152

Wang QJ, Robertson DE, Chiew FHS. (2009) A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites. Water Resources Research45: W05407. DOI:10.1029/2008WR007355.

Ward, P. J., S. Eisner, M. Flörke, M. D. Dettinger and M. Kummu, (2014a) "Annual flood sensitivities to El Niño–Southern Oscillation at the global scale", *Hydrol. Earth Syst. Sci.*, **18**(1), 47–66, doi:10.5194/hess-18-47-2014

Ward, P. J., B. Jongman, M. Kummu, M. D. Dettinger, F. C. Sperna Weiland and H.C. Winsemius (2014b) "Strong influence of El Niño Southern Oscillation on flood risk around the world." *Proc. Natl. Acad. Sci.*, **111**(44), 15659–64, doi:10.1073/pnas.1409822111

Ward, P. J., Kummu, M. and Lall, U. (2016) "Flood frequencies and durations and their response to El Niño Southern Oscillation: Global analysis", *J. Hydrol.*, **539**, 358–378, doi:10.1016/J.JHYDROL.2016.05.045

Weisheimer, A. and Palmer, T. N. (2014) "On the reliability of seasonal climate forecasts., *J. R. Soc. Interface*, **11**(96), 20131162, doi:10.1098/rsif.2013.1162

Wetterhall, F. and Di Giuseppe, F. (2018) "The benefit of seamless forecasts for hydrological predictions over Europe", Hydrol. Earth Syst. Sci., 22, 3409–3420, https://doi.org/10.5194/hess-22-3409-2018.

Wilks DS. (2009) "Extending logistic regression to provide full-probability-distribution MOS forecasts." Meteorol Appl, 16:361–368. https://doi.org/10.1002/met.134

Wilks DS and Hamill TM. (2007) "Comparison of ensemble-MOS methods using GFS reforecasts." Mon Weather Rev, 135:2379–2390. https://doi.org/10.1175/MWR3402.1

Weerts AH, Winsemius HC, Verkade JS. (2011) "Estimation of predictive hydrological uncertainty using quantile regression: examples from the National Flood Forecasting System (England and Wales)." Hydrol Earth Syst Sci, 15:255–265. https://doi.org/10.5194/hess-15-255-2011

World Meteorological Organisation (WMO). (2015) Disaster Risk Reduction (DRR) Programme, Available at: www.wmo.int/pages/prog/drr/projects/Thematic/ MHEWS/MHEWS_en.html. (Accessed July 17, 2015).

World Meteorological Organisation (WMO), (2018) Impact-based forecasting and warning: weather-ready nations, WMO Bulletin, 67 (2). https://public.wmo.int/en/resources/bulletin/impact-based-forecasting-and-warning-weather-ready-nations (Accessed 01 Nov., 2019)

Wood, A. W., and D. P. Lettenmaier (2008) "An ensemble approach for attribution of hydrologic prediction uncertainty",Geophys.Res.Lett.,35, L14401,doi:10.1029/2008GL034648.Weisheimer A, Palmer TN. (2014) "On the reliability of seasonal climate forecasts." J R Soc Interface. 2014;11(96):20131162. Published 2014 Apr 30. doi:10.1098/rsif.2013.1162.

Wood, A. W., A. Kumar and D. P. Lettenmaier (2005) "A retrospective assessment of National Centers for Environmental Prediction climate model–based ensemble hydrologic forecasting in the western United States", *J. Geophys. Res.*, **110**(D4), D04105, doi:10.1029/2004JD004508.

Wood, AW, T Pagano, and M Roos (2016) Tracing the Origins of ESP, HEPEX Historical Hydrology Series, Edition 1 (online at: https://hepex.irstea.fr/tracing-the-origins-of-esp/).

Wu, J., Lu, G. H., and Wu, Z. Y. (2014). "Flood forecasts based on multi-model ensemble precipitation forecasting using a coupled atmospheric-hydrological modeling system." Natural Hazards, 74(2), 325-340. 10.1007/s11069-014-1204-6

Wu, W., McInnes, K., O'Grady, J., Hoeke, R., Leonard, M., Westra, S. (2018) "Mapping Dependence Between Extreme Rainfall and Storm Surge." Journal of Geophysical Research: Oceans 123(4) 2461-2474. https://doi.org/10.1002/2017JC013472

Yang, T. H., Yang, S. C., Ho, J. Y., Lin, G. F., Hwang, G. D., and Lee, C. S. (2015). "Flash flood warnings using the ensemble precipitation forecasting technique: A case study on forecasting floods in Taiwan caused by typhoons." Journal of Hydrology, 520, 367-378. https://doi.org/10.1016/j.jhydrol.2014.11.028

Ye, J. Y., Shao, Y. H., and Li, Z. J. (2016). "Flood Forecasting Based on TIGGE Precipitation Ensemble Forecast." Advances in Meteorology. http://dx.doi.org/10.1155/2016/9129734

Yu, W., Nakakita, E., and Jung, K. (2016). "Flood Forecast and Early Warning with High-Resolution Ensemble Rainfall from Numerical Weather Prediction Model." Procedia Engineering, 154, pp 498-503. https://doi.org/10.1016/j.proeng.2016.07.544

Yu, W., Nakakita, E., Kim, S., and Yamaguchi, K. (2015). "Improvement of rainfall and flood forecasts by blending ensemble NWP rainfall with radar prediction considering orographic rainfall." Journal of Hydrology, 531, 494-507. DOI:https://doi.org/10.1016/j.jhydrol.2015.04.055

Yu, W., Nakakita, E., Kim, S., and Yamaguchi, K. (2016). "Impact Assessment of Uncertainty Propagation of Ensemble NWP Rainfall to Flood Forecasting with Catchment Scale." Advances in Meteorology. http://dx.doi.org/10.1155/2016/1384302

Yu, W., Nakakita, E., Kim, S., and Yamaguchi, K. (2016). "Improving the accuracy of flood forecasting with transpositions of ensemble NWP rainfall fields considering orographic effects." Journal of Hydrology, 539, 345-357. https://doi.org/10.1016/j.jhydrol.2016.05.047

Yu, W., Nakakita, E., Kim, S., and Yamaguchi, K. (2018). "Assessment of Ensemble Flood Forecasting with Numerical Weather Prediction by considering Spatial Shift of Rainfall Fields." KSCE Journal of Civil Engineering, 22(9), 3686-3696. DOI:10.1007/s12205-018-0407-x

Yuan, X., E. F. Wood, N. W. Chaney, J. Sheffield, J. Kam, M. Liang and K. Guan (2013) "Probabilistic Seasonal Forecasting of African Drought by Dynamical Models", *J. Hydrometeorol.*, **14**(6), 1706–1720, doi:10.1175/JHM-D-13-054.1

Yuan, X., Wood, E. F. and Ma, Z. (2015a) "A review on climate-model-based seasonal hydrologic forecasting: physical understanding and system development", *Wiley Interdiscip. Rev. Water*, **2**(5), 523–536, https://doi.org/10.1002/wat2.1088

Yuan, X., J. K. Roundy, E. F. Wood and J. Sheffield (2015b) "Seasonal Forecasting of Global Hydrologic Extremes: System Development and Evaluation over GEWEX Basins", *Bull. Am. Meteorol. Soc.*, **96**(11), 1895–1912, doi:10.1175/BAMS-D-14-00003.1

Yussouf, N., Kain, J. S., and Clark, A. J. (2016). "Short-Term Probabilistic Forecasts of the 31 May 2013 Oklahoma Tornado and Flash Flood Event Using a Continuous-Update-Cycle Storm-Scale Ensemble System." Weather and Forecasting, 31(3), 957-983. https://doi.org/10.1175/WAF-D-15-0160.1

Zappa, M., Fundel, F., and Jaun, S. (2013). A 'Peak-Box' approach for supporting interpretation and verification of operational ensemble peak-flow forecasts. Hydrological Processes, 27(1), 117-131. https://doi.org/10.1002/hyp.9521

Zappa, M., Jaun, S., Germann, U., Walser, A., and Fundel, F. (2011). Superposition of three sources of uncertainties in operational flood forecasting chains. Atmospheric Research, 100(2-3), 246-262. https://doi.org/10.1016/j.atmosres.2010.12.005

Zscheischler, J., Westra, S., Hurk, B.J.J.M. *et al.* (2018) "Future climate risk from compound events." *Nature Clim Change* **8,** 469–477 doi:10.1038/s41558-018-0156-3

To identify relevant papers to review, a keyword search of the ISI Web of Science was conducted in February 2020 with the keyword "ensemble flood forecast*" in the field of "title" for the period of 1900-2019 for papers in English. After removing reprints, review papers, book chapters and conference papers that are only available in abstracts, the search resulted in a total of 70 papers, including 66 journal papers from 28 journals and 4 conference papers from 3 conference proceedings. A summary of the 70 papers, including the case study system(s) and the flood events or time periods studied is provided in Appendix B.

In this study, 70 journal and conference papers were reviewed in order to understand the current status of ensemble flood forecasting, identify aspects needing improvement and guide future research directions. To understand the development of ensemble flood forecasting over time, a comparison between papers published during the early and later times of the review period was conducted. The early period is defined between 2001 when the first selected paper was published and 2010; and the later period is from 2011 to 2019. The dividing point of 2010 was selected considering the time when most of the EPS products became available and the general uptake time of EPS in research, which is approximately 2 to 3 years.

The authors acknowledge that it is not always possible to include all details of a study in the final paper published due to length limit requirement of many journals. However, the authors believe that there is essential information that needs to be included in every study for the readers to make an objective assessment of the studies carried out, understand the limitations and interpret the outcomes so future studies can be built on the knowledge developed in previous studies. Bearing this in mind, four areas of the selected papers were reviewed, including the focus or major contribution of the studies, the case studies and datasets analysed, flood modelling conducted, and finally ensemble modelling and evaluation techniques used. The review criteria used and review outcomes for each of the four areas are summarised in the four sections below.

A.1: Review criteria for study focus

The focus of the studies was investigated to understand the emphasis of current studies and the gaps for future research. The studies reviewed are categorised into 1) ensemble modelling focused studies, where the major contribution of the study is to develop or investigate ensemble modelling techniques; 2) ensemble utilisation focused studies, where the major contribution is to apply existing ensemble forecast products to simulate or understand specific flood events; and 3) application focused studies, where the major contribution is to translate ensemble forecasts into real-world applications.

Although there are many technical aspects ensemble modelling focused studies could investigate, the reviewed papers mainly focused on three aspects of ensemble modelling, including data assimilation, ensemble generation and post-processing. Ensemble utilisation focused studies can be categorised into three categories, including the development of methods for ensemble flood forecasting, the comparison or evaluation of different ensemble forecasting products, and feasibility studies examining if certain ensemble products (e.g. NWP) can be used for flood forecasts in specific areas. The application focused studies mainly considered application of ensemble flood forecasts in flood warning, reservoir operation (for flood control or navigation) and flood risk analysis.

A.2: Review criteria for case studies and datasets

The case studies investigated in the papers were analysed based on the geographic locations of the case study systems, the type(s) of flood considered, the spatial scale and temporal resolution of the case studies, including forecast lead time considered. The geographic locations are based on the contents the case study systems are located, and global studies are also identified. The types of flood considered include floods from rural or urban catchments, flash floods that have a typical onset time of less than six hours (e.g. within a rainfall event) (Hapuarachchi et al., 2011) and compound floods that are caused by multiple flood driving processes (Zscheischler et al., 2018). The spatial scale of the floods can be divided into local floods, which impact a relatively small local area (e.g. < 100 km^2), basin or catchment scale (e.g. between 100 km^2 and 2,000 km^2) and regional scale (e.g. > 2,000 km^2) that affect a relatively large area that could be across multiple countries or catchments. The temporal resolution investigated includes the frequency of the flood data predicted (e.g. sub-daily or daily). In addition, the lead time of flood forecasted is also examined, which is divided into very short range (i.e. less than 24 hours), short range (i.e. one to two days), medium range (i.e. three to 15 days), long range (i.e. 15 to 30 days) and very long range (i.e. longer than 30 days). Since it is possible for more than one system to be investigated in one paper, the review was conducted based on the number of times each case study falls into each defined category.

A.3: Review criteria for flood modelling

The flood modelling aspect of the papers was investigated to understand the types of flood models used, flood variables predicted and flood driving processes considered. The types of models used in the papers reviewed include numerical weather prediction models (predicting weather variables, such as rainfall), hydrological models (simulating rainfall-runoff processes), hydraulic models (1D, 2D, 3D or conceptual models simulating flood levels or inundation) and data driven models, ranging from simple regression models to machine-learning techniques (simulating various flood variables). Flood variables predicted are directly related to flood models used and include rainfall intensity (as surrogate of intensity of floods), stream flow, flood inundation levels and area. Flood driving processes considered include precipitation (i.e. through a pluvial process), riverine floods (i.e. through a fluvial process), ocean processes such as high astronomical tide or storm surge, other meteorological conditions such as temperature, pressure, wind and humidity, as well as catastrophic events, such as dam failure or glacial lake outburst floods. It should be noted that more than one type of flood model, flood variable or driving processes may be considered in a study and therefore the total number in each category in the summary could exceed the total number of papers reviewed, as mentioned above.

A.4: Review criteria for ensemble modelling

The ensemble modelling aspect of the papers was investigated based on how the ensembles were generated, how the ensemble forecasts were evaluated, and the number of ensemble members used. The methods used to generate ensembles are divided into four categories, including Multi-model-based ensemble, initial/boundary condition perturbation, parameterisation of models, and statistical approach

to account for errors (e.g. Bayesian approach or residual error) in an integrated manner. Ensemble forecasting evaluation methods includes 1) methods considering the whole distribution of ensemble forecast, such as Brier Score (BS), Continuous Ranked Probability Score (CRPS), CRPS, Reliability diagram, Relative Operating Characteristic (ROC) curve, Fractions Skill Score (FSS),and Rank histogram, among others; 2) traditional threshold values based methods, such as contingency table showing hits and misses, Critical Success Index (CSI), Gilbert Skill Score (GSS) and probability of threshold exceedance/detection, among others; 3) traditional error statistics, such as Root Mean Square Error (RMSE) and Taylor diagram; and 4) ad hoc methods, which are mainly based on visual inspection (e.g. the spaghetti plot, postage stamp maps or quantile maps) or evaluation of ensemble means only. In addition, the number of ensemble members used in each of the studies was also investigated to understand the extent of uncertainties considered and the applicability of ensemble forecasts in representing different flood producing conditions, especially extreme or component flood events.

.

APPENDIX B LIST OF SELECTED PAPERS, AND THE CASE STUDY INFORMATION

Author, year	Case study system	Events/periods analysed
Goodarzi et al. (2019a)	Kan basin, Tehran, Iran	14 individual flood events.
Ziliani et al. (2019)	An experimental model of a river located in the Occidental Alps (Italy)	A dam failure flood event.
Liu et al. (2019)	The Yarlung Zangbo River basin, China.	Years 1998 to 2015
Lee et al. (2019)	Ten headwater basins in Texas, US	Flood events over a 10-year period with a peak flow exceeding 100m^3/s.
Tian et al. (2019)	Two mountainous catchments in Jin-Jing-Ji Region of Northern China	Four storm events.
Han and Coulibaly (2019)	Humber River watershed, Canada	January 2011 to December 2015
Goodarzi et al. (2019b)	Kan Basin, Iran	14 historical storms.
Antonetti et al. (2019)	The Emme catchment, Switzerland.	Four flood events.
Gomez et al. (2019)	The tidal Delaware River near Philadelphia, Pennsylvania, US	Years 2008 to 2013
Nayak et al. (2018)	Folsom Reservoir in California, US	100 year synthetic data generated based on 12 year observations
Yu et al. (2018)	The Futatsuno and Nanairo dam catchments in the Shingu river basin, Japan	The typhoon Talas, 2011
Amengual et al. (2017)	Serpis River basin, Spain	The 12 October 2007 flash flood event
Barthelemy et al. (2017)	Adour river, southwest France	Eight flood events 2009 - 2014
Doycheva et al. (2017)	Mulde river basin, Germany	June 2013 flood event
Kovalchuk et al. (2017)	Baltic Sea, St. Petersburg, Russia	Hypothetical flood events
Liu et al. (2017)	Lianjiang Basin, southeast China.	Years 1987 to 2012
Saleh et al. (2017)	Confluence of Passaic and Hackensack rivers and Newark Bay, New York, USA	Hurricane Irene (2011) and Hurricane Sandy (2012)
Schulte (2017)	13 stations in the New York harbour, US	Hurricane Sandy (2012)
Dahl and Xue (2016)	Oklahoma City, US	The 14 June 2010 flood event
Fan et al. (2016)	Tocantins River, Brazil	2011/2012 rainy season
Golding et al. (2016)	Edinburgh, Scotland	July 2011 and 7 events July - August 2012
Kasiviswanathan et al. (2016)	Bow River, Alberta, Canada	5 June –7 August 2005, 9 May –25 August 2007, and 24 May–7 August 2010
Kimura et al. (2016)	Tanshui River watershed, northern Taiwan	Seven typhoon events
Ravazzani et al. (2016)	Milano urban area, Italy	17–18 September 2010 and 7–8 July 2014

Siqueira et al. (2016)	Taquari-Antas basin, southern Brazil.	Two flood events on 6 June 2014 and 21 July 2011
Barthelemy et al. (2016)	Adour and Marne catchments, France	One 2011 flood event
Bellier et al. (2016)	5 basins in the French Upper Rhone river region	18-26 year periods
Ye et al. (2016)	Upstream catchment of Yishu River, China	12 flood events 2001 - 2004, 2007 - 2008 and 2010
Yu et al. (2016)	The Shingu river basin, Japan	The 2013 Typhoon Talas event
Yu et al. (2016)	33 catchments in Shingu river basin, Japan	Typhoon Talas (2011)
Yu et al. (2016)	Shingu river basin, Japan	Typhoon Talas (2011)
Yussouf et al. (2016)	Oklahoma City, USA	The 31 May 2013 flash flood event
Liu et al. (2015)	Three Gorges Reservoir, China	Flood season 2003 - 2010
Thiemig et al. (2015)	36 catchments in Africa	39 medium to large flood events in 2003
Yang et al. (2015)	Urban areas in Pingtung County, Taiwan	Six typhoon events 2010 - 2014
Fleming et al. (2015)	Englishman River, British Columbia, Canada	2013-2014 storm season
Shi et al. (2015)	Juma River basin and upper Baishui River basin, China	Storm on 21 July 2012 for first basin and storm on 17 July 2010 for second basin
Yu et al. (2015)	Futatsuno and Nanairo dam catchments in Shingu river basin, Japan	Typhoon Talas (2011)
Ushiyama et al. (2014)	Kabul River Basin, Pakistan	27 - 29 July 2010
Wu et al. (2014)	Wangjiaba sub-region in Huaihe River Basin, China	5-8 July 2007 and 13-17 August 2008
Alfieri et al. (2013)	Large world river basins	January 2009 - December 2010 and 2010 Pakistan flood
Cane et al. (2013)	Western Piemonte, Italy	May 2008 flood
Chen et al. (2013)	Cobb Creek watershed, US and Chuzhou watershed, China	12 events between 1 June 2006 and 30 September 2009 for Cobb Creek watershed. 16 events 1981 - 2002 for Chuzhou watershed
Das and Kaur (2013)	Mahanadi basin, Indian	Flood seasons 2009 - 2010
Hsiao et al. (2013)	Lanyang basin, Taiwan	Typhoon Nanmadol (2011) and (deterministic forecast of) 21 typhoons in 2011
Lee et al. (2013)	314 urban areas, Taiwan	Typhoon Megi and Typhoon Nanmadol (2011)
Leoncini et al. (2013)	Boscastle, UK	Boscastle flash flood event on 16 August 2004
Liechti et al. (2013)	Calancasca Catchment, Verzasca Catchment and Ticino catchment, southern Switzerland	Between June 2007 and December 2010
Marty et al. (2013)	Cevennes-Vivarais Catchments, Southern France	4-10 September 2005, 19-25 November 2007, 20-23

		October 2008, 31 October - 4 November 2008
McSharry et al. (2013)	Britain, UK	257 rainfall events 1866–1968
Bao and Zhao (2012)	Huaihe River, China	July 2007
Nester et al. (2012)	43 catchments in Austria and Germany	Four years of data
Vincendon et al. (2011)	Gardons river, C`eze river and Ard`eche river, France	Two flash flood events: 21–22 October 2008 and 1–2 November 2008
He et al. (2010)	Huai catchment, China	Three flood events July-September 2008
Hopson and Webster (2010)	Ganges Rivers, Bangladesh	Severe floods 2003 - 2007
Shamir et al. (2010)	Nakdong River, South Korea	Jan 2006 - Aug 2008
Thiemig et al. (2010)	Juba–Shabelle river basin, Africa	1 January 1961 - 31 December 1981
Tiwari and Chatterjee (2010)	Mahanadi River basin, India	2001- 2005
Dietrich et al. (2009)	Upper Mulde river basin, Germany and Czech Republic	Seven events 2002 - 2008
Dietrich et al. (2009)	Mulde river basin, Germany	May 2007 - April 2008 and August 2002 extreme flood
Kim et al. (2009)	Yodo-River basin, Japan	Two events from August 1992 and June 1993
Kuchment and Gelfan (2009)	Vyatka Rive, Russia	41 spring floods (1940–1980)
Dietrich et al. (2008)	Mulde river basin, Germany	August 2002 extreme event
Komma et al. (2008)	Kamp catchment, Austria	Six largest flood events on record: August 2002 (two), July 2005, August 2005 (two), March 2006
Komma et al. (2007)	Kamp catchment, Austria	Five flood events on 8 and 13 August. 2002, 11 July 2005, 16 and 22 August 2005
Srikanthan et al. (2008)	Georges River in Sydney, Australia	The May 2003 flood event
Ramos et al. (2007)	Danube river basin, France	Flood events in 2005
Verbunt et al. (2007)	Alpine tributaries of Rhine River basin, central Europe	Spring 1999 and November 2002 floods
Montani et al. (2003)	Soverato, southern Italy	September 2000 flood
Mackey and Krishnamurti (2001)	Southeast China	Typhoon Winnie (1997)