Deliverable 4.3

Guidelines for uncertainty propagation analysis

This project has received funding from the European Union´s Seventh Framework Programme for Research, Technological Development and Demonstration under Grant Agreement Nº 603663 for the research project PEARL (Preparing for Extreme And Rare events in coastal regions).

pearl

Preparing for Extreme And Rare events in coastaL regions

Deliverable 4.3

Guidelines for uncertainty propagation analysis

Document Information

Project Number	***603663***	Acronym	**PEARL**
Full Title	Preparing for Extreme and Rare events in coastaL regions		
Project URL	http://www.pearl-fp7.eu/		
Document URL			
EU Project Officer	Denis Peter		

Deliverable	Number	D.4.3	Title	Guidelines for uncertainty propagation analysis
Work Package	Number	WP4	Title	Flood forecasting and early warning systems for coastal regions
Authors	Marios Anagnostou, George Karavokiros, Panagiotis Kossieris (NTUA) Henrik Madsen, Nina Sto Domingo (DHI) Jelena Batica (UNSA) Xavi Llort (HYDS)			

Date of Delivery	Contractual	31/12/2016	Actual	30/11/2017
Status	version 1.0		final Y	
Nature	prototype Y report Y dissemination □			
Dissemination level	public Y consortium □			

Abstract (for dissemination, 100 words)	This report presents the results from PEARL Task 4.3 on *Advanced uncertainty methods and tools for early warning*. Several aspects of uncertainty have been explored associated with the measurement of precipitation and various components of an Early Warning System (EWS). The main sources of uncertainty in EWS along with the most important methodologies and tools for their analysis have been discussed. Methodologies have been categorized by source of uncertainty while their strengths and shortcomings have been addressed and references were given for further research. A methodological framework for the quantification of uncertainty in precipitation forecasts, combining numerical weather predictions and satellite estimates has been presented and a methodology for addressing radar precipitation uncertainty in nowcasting has been introduced and uncertainty propagation in model chains has been

	analysed identifying critical uncertainty propagation issues and bottlenecks. Finally, the above methodologies have been applied in two case studies in the cities of Marbella, Spain and Greve, Denmark.
Keywords	flood forecasting, early warning system, EWS, uncertainty, radar precipitation, nowcasting, model chain, Marbella, Greve

Version Log				
Issue Date	Rev. No.	Author	Change	Approved by
29/7/2015	0.1	George Karavokiros	Preliminary version, TOC	
20/9/2016	0.2	Marios Anagnostou	Forecast and nowcast uncertainty in Early Warning Systems	
2/12/2016	0.3	Henrik Madsen, Nina Sto Domingo	Model chain uncertainty and Greve case study	
23/11/2016	0.4	Jelena Batica	Uncertainty analysis using Monte Carlo based methods	
23/12/2016	0.5	Panagiotis Kossieris	Literature review on EWS	
18/1/2017	0.6	Jelena Batica	Updates on uncertainty analysis using Monte Carlo based methods	
31/10/2017	0.7	Xavi Llort	Nowcast and Radar precipitation uncertainty. Marbella case study.	
30/11/2017	1.0	George Karavokiros, Henrik Madsen, Xavi Llort	Final version including all corrections	

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant agreement n° 603663 for the research project PEARL (Preparing for Extreme And Rare events in coastaL regions).

Abstract

This report presents the results from PEARL Task 4.3 on *Advanced uncertainty methods and tools for early warning*, elaborated jointly by the project partners NTUA, DHI, UNSA and HYDS. Several aspects of uncertainty have been explored associated with the measurement of precipitation and various components of an Early Warning System (EWS). The main sources of uncertainty in EWS along with the most important methodologies and tools for their analysis have been discussed. Methodologies have been categorized by source of uncertainty while their strengths and shortcomings have been addressed and references were given for further research (Section 2). After that, a methodological framework for the quantification of uncertainty in precipitation forecasts, combining numerical weather predictions and satellite estimates has been presented in Section 3. In the same section a methodology for addressing radar precipitation uncertainty in nowcasting has been introduced and uncertainty propagation in model chains has been analysed identifying critical uncertainty propagation issues and bottlenecks. Finally, in Section 4, the above methodologies have been applied in two case studies in the cities of Marbella, Spain and Greve, Denmark. The EWS of Marbella demonstrates the ability of the system for uncertainty reduction by continuous adjustment of forecasts utilizing updated real-time information. In the Greve case study, uncertainty is quantified in probabilistic terms and visualized using innovative uncertainty visualization tools.

Contents

List of figures

List of tables

1 Introduction

At a time of global changes, the world is striving to adapt to inevitable and profound alterations. The increasing numbers of forest fires, severe and more frequent floods are already taking place. Climate change may cause additional environmental stresses and societal crises especially in regions already vulnerable to natural hazards and prone to social conflicts. Between 2000 and 2016, severe events in Europe affected more than 1 million people, causing around 1500 human deaths and resulting in 20 billion euro direct economic losses (CRED, 2016). To mitigate and reduce these impacts, various risk management frameworks and plans have been issued, both in European and regional levels (e.g., European Commission, 2007, 2004), aiming to improve the prevention, protection, preparedness and emergency response to such events. In this framework, multi-hazard early warning systems have a key role to play in the prevention of damages and fatalities by informing local authorities about imminent events and hence allowing for advance warning of affected populations.

According to International Strategy for Disaster Reduction of United Nations (UNISDR, 2009), an Early Warning System (EWS) is defined as "The set of capacities needed to generate and disseminate timely and meaningful warning information to enable individuals, communities and organizations threatened by a hazard to prepare and to act appropriately and in sufficient time to reduce the possibility of harm or loss". An effective and people-centered EWS is compiled by the integration of four key elements (WMO, 2015): i) *Risks Knowledge*, ii) *Monitoring and Warning service*, iii) *Dissemination* and iv) *Emergency response capacity*. The role, objectives, and functionalities of each element are presented in Figure 1. The key idea behind early warning is that the earlier and more accurately we are able to predict short- and long-term potential risks associated with natural and human induced hazards, the more likely we will be able to manage and mitigate a disaster's impact on society, economies, and environment.

Figure 1: The main elements of a people-centered EWS (source: WMO, 2015)

The efficient implementation of any EWS requires the synergy of the above four components, based on an appropriate legislative and legal framework, along with the close collaboration of many agencies and actors (e.g. International bodies, National and Local governments, Regional institutions and organizations, Communities, NGOs, Science community, Media, Private Sector). Failure of any part of the system will imply failure of the whole system. For example, accurate warnings will have no impact if they are not issued in sufficient lead time, if the population is not prepared or if the alerts are received but not disseminated by the agencies receiving the messages.

The core of any EWS is the forecast service that collects, processes and analyses the various data and information to produce forecasts on the characteristics of a forthcoming flood event. During the last decades, many forecast services of different complexity and sophistication have been developed, tailored to the special characteristics of the hazard and the focal area (Alfieri et al., 2012), as well as the interplay between different hazards such as floods and fires (Kochilakis et al., 2016). The spatial and temporal extent of the different water-related hazards is summarised in the top panel of Figure 2. The bottom panel presents the corresponding meteorological inputs that are used in the forecasting process along with a qualitative assessment of their skill towards the forecast lead time.

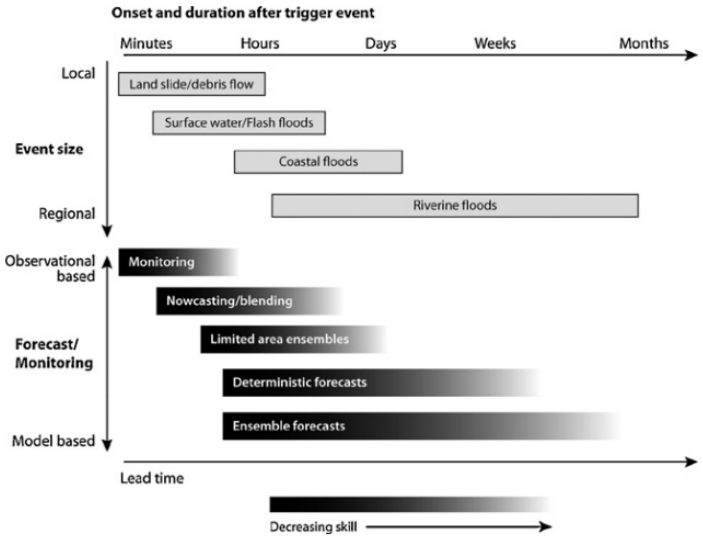


Figure 2: Spatio-temporal extent of water-related hazards and the corresponding meteorological input for forecasting (source: Alfieri *et al.*, 2012)

According to Borga *et al.* (2011) and Collier (2007), the forecast service of a Flood EWS is typically composed by: i) a precipitation detection system, typically consisted of remote sensing (radar, satellites) and ground rainfall and water-level gauges, ii) a regional and global weather prediction model, able to provide short-range meteorological forecasts and iii) hydrological and/or hydraulic models to simulate the response of the catchment, at a wide range of time scales, over predictions (Figure 3).

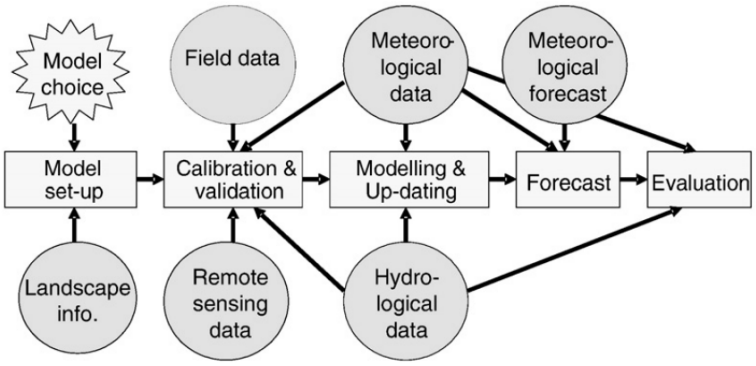


Figure 3: Main components, data and tasks of a complete Flood EWS (Arheimer et al., 2011)

However, each of the above EWS components is subjected to considerable uncertainties and errors that *cascades* from one process to the next, i.e., from atmospheric conditions and observations to rainfall forecasts, from rainfall to runoff forecasts, and from runoff to flood inundation forecasts (Pappenberger et al., 2005b). The quantification of uncertainty at each step of the process along with the study of cascade effect in model chain is of high importance towards the accurate and timely detection of severe flood events.

In the next section (Section 2), the main sources of uncertainty in EWS along with the most important methodologies and tools for their analysis are discussed. In the same section, methodologies are categorized by source of uncertainty while their strengths and shortcomings are addressed and references are given for further research. In Section 3 a methodological framework for the quantification of uncertainty in precipitation forecasts, combining numerical weather predictions and satellites estimates, is presented. The methodology has been applied for the Greve region producing ensemble forecasts and in Marbella (Spain) to build an EWS. Finally, the application of the methodologies in two case studies in the cities of Marbella, Spain and Greve, Denmark is discussed in Section 4.

2 Uncertainty in early warning systems

2.1 Uncertainty in precipitation forecasts

Precipitation is the key factor that drives the hydrometeorological circulation and it is the most important element of the global water and energy cycle, and the driving force for one of the most devastating natural hazards, that is, floods (Nikolopoulos et al., 2010). Three-fourths of the energy absorbed by the earth's atmosphere is derived from the latent heat released from condensation of water vapour (Meischner, 2004). However, the simulation and forecast of precipitation events is rather a hard task due to the high spatio-temporal variability of the physical process (Alemohammad, 2015). This is in principal the main source of uncertainty in the forecast chain and consequently the lack of efficiency of any EWS which depends on the lead time where accurate weather forecasts are available. The importance of accurate representation of the spatial variability of rainfall on simulating floods/ flash floods have been demonstrated by several studies (Zoccatelli et al 2010; Lobligeois et al., 2014). Moreover, recent studies have shown the importance of accurate rainfall estimates at small spatial scales for the prediction of debris flow/landslide initiation (Nikolopoulos et al. 2014).

Towards more accurate predictions, modern flood and flash flood EWS demand for improved quantitative measurements of precipitation, at temporal scales of minutes and spatial scales of a few square kilometres. This is of high importance especially in urban catchments and small watersheds, which are characterized by particularly fast response, and hence require even higher resolution of the rainfall products (Dabbert et al. 2000). During last decades, different techniques have been developed to provide precipitation forecasts at different spatial extents, temporal analysis and forecast range. These techniques are based mainly on Numerical Weather Predictions (NWP), weather radar and satellite estimates, as well as on their blending.

Nowcasts refer to short-range (up to 6 hours ahead) precipitation forecasts (Golding, 2000). Until recent years, such predictions have been based on *extrapolation techniques* that use radar and satellite observations to capture the spatio-temporal dynamic of precipitation field and forecast the motion of a discrete set of rain objects (Alfieri *et al.*, 2012; Borga *et al.*, 2011; Collier, 2007). The error in estimates is typically adjusted using real-time measurements from ground rain-gauges. Forecasting extrapolation techniques can be grouped into those that use correlation techniques, those that track the centroid of an object and those that use NWP wind advection techniques (Bowler *et al.*, 2004). Other approaches concern the use of artificial neural network methods and Bayesian statistical techniques (Hapuarachchi *et al.*, 2011).

Although this type of methods successfully simulate and forecast the precipitation field at coarser spatial and temporal scales, they show some weaknesses in developing convection of rainfall fields in new areas and capturing cell splitting and decay, that usually control heavy and rapid rain dynamics and hence, flash floods (Borga *et al.*, 2011). Until recently, NWP provided forecasts over a coarse spatial grid (12 km or larger) which it was not possible to capture very localized processes (Collier, 2007). As has already mentioned above rainfall products at finer temporal and spatial scale is vital in forecasting of flood events in urban catchments and small watersheds where fast and localised phenomena are evolved. Predictions on finer temporal resolution (lower than 1 hour) and spatial resolution (a few kilometers) achieved with the new generation of NWP models, which offer the prospect of producing useful forecasts of convective storms on scales applicable for flood prediction (Lean and Clark, 2003). The accuracy of predictions was further improved and the uncertainty bounds were narrowed down by combining NWP forecasts with radar estimations. A characteristic operational forecast system that follows this data assimilation approach is the Nimrod precipitation nowcasting system (Golding, 1998) of the UK Met Office that combines radar advection with NWP to provide precipitation forecasts from 1 up 6 hours lead time (Collier, 2007).

Although, assimilation of radar-data on NWP improves the accuracy of predictions, the high stochastic and complex nature of the atmosphere does not allow the exact prediction of its state (Lorenz, 1969). To take into account these uncertainties, ensembles of many NWPs, known as ensemble prediction systems (EPS), rather than single deterministic forecasts, are used. At a second step, the EPS feed into the hydrological model to produce a set of different flood prediction scenarios. The use of ensembles enables the improvement of the accuracy of general weather forecasts at medium-range (Tracton and Kalnay, 1993) and at short-range lead times (Du *et al.*, 1997) and (Collier, 2007). EPS are typically derived using several perturbations (from 10 to 50) of the initial conditions of the model according to a Monte Carlo sampling approach. According to (Buizza *et al.*, 2005), typical ensemble perturbation approaches are : i) the error breeding method, ii) the perturbed observations method, iii) the singular vectors method and iv) the ensemble Kalman Filter method. A typical operational EPS system is the Short-Term Ensemble Prediction System (STEPS), which merges an extrapolated nowcast with different downscaled NWP model forecasts in order to provide the probability of precipitation (Bowler *et al.*, 2004).

2.1.1 Forecast and satellite precipitation uncertainty

In this subsection our focus is on the investigation of precipitation uncertainty that results from Numerical Weather Prediction (NWP) models and provide as reliable as possible surface precipitation estimates to subsequent models (e.g. runoff, flood forecasting) based on NWP and satellite data.

Operational NWP systems consist of global atmospheric models which run typically at horizontal resolution ranging from 20 to 50 km and of mesoscale limited area models with horizontal resolution ranging from 5 to 10 km. As the trend for increasingly higher resolutions continues, following the improvement of the dramatic increase in computational power, state-of-the-art non-hydrostatic limited area models running at operational basis may achieve today very high horizontal resolutions of 1-3 km. Based on the NWPs, meteorological agencies issue periodically rainfall forecasts for specific areas with a very dense horizontal grid spacing. The forecasting period may vary from hours up to a few days, typically with a time step of one hour.

When examining single rainfall events, one may observe the very high spatio-temporal variability of the process. This is typical especially for major storm events, which are of particular interest in this project as they are more likely to cause floods, depending on the characteristics of the drainage area and the event. While the relevant characteristics of a hydrosystem are usually sufficiently known (main hydrographic network, drainage system) and change very slowly over time, rainfall events cannot be reliably predicted in the sub-grid temporal and spatial scale needed to feed the rainfall-runoff models. Forecasts in the highest resolution provided today by weather agencies are typically given in 1hr/2km, which leaves open the question of the distribution of the actual event within a given temporal and spatial frame. This information however is substantial for the assessment of flood risk in many hydrosystems and can hardly be addressed in a deterministic way.

In addition, accuracy and reliability of hydrological modelling studies heavily depend on quality and availability of precipitation estimates. Difficulties in representation of high rainfall variability over regions with poorly covered or without any ground based sensors make satellite remote sensing techniques attractive for meteorological and hydrological research. Over the past decade, precipitation retrievals from earth-orbiting satellites have increased in availability and improved in quality. Current satellite retrieval techniques are capable of producing precipitation retrievals on high spatial resolutions; however, high uncertainties are associated with such high-resolution products. Moreover, satellite-based rainfall measurements are now almost globally available at high spatial resolution. These products have uncertainties that necessitate the use of error characterization and correction procedures based on more accurate in situ rainfall measurements.

Many satellites missions have been launched by different organizations associated with earth observing missions. These include three main groups of on-board instruments used for precipitation estimation:

- Infrared (IR) and Visible (VIS) bandwidth instruments that are mostly installed on the Geostationary Orbit (GEO) satellites. These instruments are not used to directly estimate precipitation intensities and rates. However, they are able to estimate precipitation with a proper post-processing algorithm for cloud top or near cloud top emission observation.
- Based on satellite cloud top emission observations (i.e. temperature), it is possible to use the Geostationary Operational Environmental Satellite (GOES) Precipitation Index (GPI) algorithm for the precipitation estimation in the tropics. This method uses the Infrared (IR) observations of cloud top temperature that is far removed from the physics of the precipitation generation process.
- Finally, there are also instruments in IR and VIS bandwidths mounted on Low Earth Orbit (LEO) platforms. These include the recently launched Moderate-Resolution Imaging Spectroradiometer (MODIS) on-board of Aqua and Terra satellites and the Visible and InfraRed Scanner (VIRS) on-board of Tropical Rainfall Measuring Mission (TRMM) satellites.

For hydrometeorological applications, there is a second group of satellite-based instruments, namely are the Passive Microwave (PMW) or radiometers. Microwave frequencies are the most widely used for precipitation retrieval on-board of LEO platforms. Two different approaches are applied for the estimation of precipitation in relation with PMW instruments: a) The scattering of the background emission (used as a reference from the Earth's surface) by the precipitation ice particles and b) the latent heat emission from precipitation droplets (the ground decreases the background microwave radiation which droplets tend to increase). Nowadays, there are several PMW instruments in orbit equipped with different features: Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E), Special Sensor Microwave Imager (SSM/I) and its new version Special Sensor Microwave Imager/Sounder (SSMIS), Advanced Microwave Sounding Unit-B (AMSU-B), TRMM Microwave Imager (TMI), and the recent launched Global Precipitation Measurement (GPM) Microwave Imager (GMI). Each of these instruments has several channels that are capable of detecting different types of precipitation. GMI is the most recent and advanced PMW instrument that has 13 channels from 10 GHz to 183 GHz. PMW measurements of rainfall have their own unique advantages as well as shortcomings. However, in order to overcome those common problems and to obtain reasonably accurate rainfall estimation with a good spatial and temporal sampling resolution, useful for studies of climate change and large-scale hydrological processes, it is necessary to merge different types of measurements (Alemohammad, 2015; Nikolopoulos et al. 2013).

In order to produce better estimations of rainfall, researchers have tried to combine different types of rainfall observations. Several models have been developed that combine satellite measurements with ground base measurements, and provide rainfall estimations for both missing pixels and for time periods for which no satellite measurements exist. The current limitations on accurate estimation of extremes prevent us from designing short-term warning systems based on satellite data. Future advances on precipitation estimates in detection of extremes may lead to a significant improvement in early warning systems and hazard mitigation. Additional research efforts on extremes evaluation using ground reference measurements (e.g., ground radars and gauges) is necessary to understand to what extend one can observe extremes reliably (Sorooshian and AghaKouchak, 2010). Traditionally, rain gauges have been used to measure surface rainfall rates. Gauges are considered as the most accurate sensors for measurements over a limited area (nearly a point), but their small coverage (especially over complex terrain and tropical regions) limits the adequacy in representing the spatial structure of highly variable rainfall fields over large spatial scales (Nikolopoulos et al. 2010). Weather radars, on the other hand, have the capability to monitor precipitation at high spatial and temporal scales and they have stimulated great interest and support within the hydrologic and meteorological community (Anagnostou et al. 2004). However, satellite

sensors offer unique advantages in regions where other capabilities are inexistent or sparse. They are able to detect extreme events more reliably in large scales than in small scales, since observation noise has minor impact on large scale events. There is indeed a trade-off between resolution and observation skill. Currently, the most common used resolution in satellite observations is 0.25o/3hr.

2.2 Uncertainty in hydrological modelling

Hydrological and hydraulic models support the transformation of meteorological predictions (e.g., precipitation, air temperature, wind speed, humidity etc.) into runoff forecasts and, as Serban and Askew (1991) point out, are "*the heart of any flow forecasting system*". In the EWS framework, models of different complexity have been developed and employed during the last decades (see reviews in Collier, 2007; Hapuarachchi et al., 2011). Data-driven models generate runoff forecasts through simplified statistical cause-effect relationships (i.e., without any physical interpretation), which are established from historic rainfall and runoff data. The main representatives of this category are the Transfer Function models (Sene and Tilford, 2004) and the neural network-based approaches (Chiang et al., 2007; Kim and Barros, 2001; Piotrowski et al., 2006; Sahoo et al., 2006; Sahoo and Ray, 2006). The data-driven models are easy to set-up and find applicability especially in cases where the physical processes of the under study catchment are unknown. However, long series of observed data are required in model calibration, while the derived relationships are case specific providing a coarse representation of the phenomenon.

In contrast to the black-box approach of data-driven techniques, physically-based (i.e., use of physical laws) and conceptual models (i.e., use of semi-empirical equations) aim to represent the main components and processes of hydrologic cycle of a catchment. Lumped hydrological models (i.e., the whole basin is treated as homogeneous) found wide applicability in flood forecasting thanks to their parsimonious structure which represents the phenomenon with only a few parameters, the lower computational burden and the low requirements for data (Hapuarachchi *et al.*, 2011). Characteristic examples of such models are the Sacramento soil moisture accounting model (SAC-SMA model; Burnash et al., 1973), used by the US National Weather Services to produce flash flood guidance, and the Hydrologiska Byrans Vattenavdelning (HBV) model (Bergstrom, 1976; Kobold and Brilly, 2006). During the last years, the increased computer capacity along with the greater availability of data pave the way for the implementation of spatially semi-distributed and distributed models in the EWS: the HL-RDHM model (Koren *et al.*, 2004), the TOP-MODEL model (Berenguer *et al.*, 2005; Beven and Binley, 1992), the TOPKAPI model (Bartholmes and Todini, 2005; Ciarapica and Todini, 2002) the LISFLOOD model (De Roo et al., 2000; Pappenberger et al., 2008, 2005b; Thielen et al., 2009), the MARINE model (Estupina-Borrell et al., 2006), the TREX model (England et al., 2007). Distributed modeling enables simulation at finer temporal and spatial scales and hence it can be considered as more suitable in flash flood applications. Moore *et al.* (2006) suggests that for extreme events, distributed models may give more plausible results than lumped models. However, the high computational burden along with the high demands for observed data may restrict their operational applicability.

Since hydrological modeling is never accurate, all models are subject to varying degrees of uncertainty. In order to enhance the reliability and credibility of the model output, estimations of these uncertainties should be incorporated in the deterministic forecasts (Montanari and Brath, 2004; Tolson and Shoemaker, 2008). Further to the inherent complexity of natural systems, the total uncertainty in the output of hydrological models can be generally associated with the following factors (Efstratiadis and Koutsoyiannis, 2010): (a) measurement errors; (b) use of over-parameterized model structures, whose complexity is inconsistent with the available information about the system behaviour; (c) inappropriate representation of the temporal and spatial variability of model inputs, which are obtained either from processed data (e.g. discharge records based on stage information) or point observations (e.g. precipitation, temperature); (d) poor identification of initial and boundary conditions; (e) non-informativeness of calibration data with regard to the entire system regime; (f) use of statistically inconsistent fitting criteria (e.g. error metrics not accounting for heteroscedasticity);

(g) weaknesses of nonlinear optimization algorithms on rough and high-dimensional response surfaces; and (h) inconsistent assumption of parameters constant in time whilst the environment is changing, e.g. due to urbanization, deforestation, stream lining and other human interventions.

The total error in model predictions is often attributed into three main sources of uncertainty (Montanari and Koutsoyiannis, 2012; Pianosi and Raso, 2012; Tolson and Shoemaker, 2008):

1. **Measurement error:** it derives from the imperfect or infrequent measurement of data (i.e., flow or rainfall data) used in model calibration and validation, and as initial or boundary conditions.
2. **Parameter uncertainty:** it is defined as the uncertainty in model parameter vector and is related with the accuracy and availability of observations, the calibration procedure and the structure of the model.
3. **Model structural uncertainty:** also known as 'model error', it derives from the inability of the model to perfectly represent the real physical mechanism.

Although a lot of work has been conducted during the last decades on the estimation and assessment of model uncertainty, this scientific field still remains active and challenging. Many of the developed techniques, adopting the assumption of decomposition of total error into individual components, try to identify and assess a specific source of uncertainty. For example, see McMillan et al. (2011, 2010), Montanari and Di Baldassarre (2013), Di Baldassarre and Montanari (2009), Sikorska et al. (2013) for uncertainty in input and output data; Ebtehaj et al. (2010), Shrestha et al. (2014), Tolson and Shoemaker (2008), Vrugt and Robinson (2007b) for parameter uncertainty assesment; and Efstratiadis et al. (2015), Montanari and Brath (2004), Montanari and Grossi (2008), Sikorska et al. (2015) and Montanari and Koutsoyiannis (2012) for structural uncertainty. However, the distinction between different sources of uncertainty is rather a hard task and requires assumptions on the statistical properties of the individual errors (Sikorska et al., 2015). As an alternative to 'decomposition' approach, the estimation of global predictive uncertainty of deterministic model can be based directly on 'model residuals', i.e., studying the stochastic and statistical properties of the errors between observed and predicted values (Pianosi and Raso, 2012). A comprehensive review on uncertainty assessment techniques and their underlying assumptions can be found in Montanari (2011) and Beven (2009).

The complexity and non-linearity of hydrological models does not allow for the analysis, estimation and propagation of uncertainties via analytical methods. This process typically involves probability theory and Monte-Carlo (MC) sampling techniques, Bayesian statistical methods, machine learning techniques and fuzzy set theory. The most popular, MC-based, uncertainty assessment method is the 'Generalized Likelihood Uncertainty Estimator' (GLUE), developed by Beven and Binley (1992). This approach rejects the idea of an "*optimum model*" or parameter set in favor of the "*equifinality concept*" that assumes that many parameter set or model structures may have an equal likelihood to produce acceptable or "*behavioural*" simulations. The general GLUE methodology is evolved as follows:

> Initially, different candidate model structures are identified along with the parameters that most affect the output. Then, prior distributions of models' parameters are used to generate a large number of parameter sets via a MC-type sampling process. In the case that prior knowledge on parameters' distribution is not available, simple uniform random sampling is employed. The models are run for each generated parameter set and the model outputs are compared with the available calibration data. A quantitative measure is then used to assess the performance of each model, based on the model residuals. Different likelihood measures (e.g., Nash-Sutcliffe coefficient) or combinations of them can be used (see Beven and Freer, 2001 for a list of potential measures), while the only requirement is that the measure should increase monotonically with increasing goodness of fit and that "*non-behavioural*" models should have a likelihood of zero. The models and their corresponding parameter sets that provide a likelihood measure that reaches a minimum threshold are retained and allowed to

contribute to the distribution of predictions. The predictions of each simulation are weighted by the likelihood measure associated with that simulation, according to a standard procedure: Initially, the calculated likelihoods are rescaled to produce a cumulative sum of 1.0. At each time step *t* the cumulative distribution function of simulated discharges is then constructed, using the rescaled weights, with the highest discharge associated to a probability of not exceedance equal to 1. The uncertainty bounds (prediction limits) at each time step can be derived from the cumulative distribution function. The implementation of the above general methodology requires from the modeler to make subjective decisions at each step of the process (Beven, 2012):

- the model or models to include in the analysis
- the feasible range for each parameter value
- the sampling strategy for the parameter sets
- an appropriate likelihood measure or set of measures
- the conditions for rejecting models

The main advantage of the GLUE methodology is the flexibility to account for different sources of uncertainty (Montanari, 2005). For instance, by considering different parameter sets or model structures, one can explicitly assess parameter uncertainty or structural uncertainty, respectively. By using different input and output data, the effect of measurement errors can be estimated. However, this flexibility and subjectivity in the choice of various components have raised a lot of criticism and discussion around GLUE method (e.g., see Montanari, 2005; Montanari and Koutsoyiannis, 2012; Tolson and Shoemaker, 2008 and the reviews therein). Further to the high computational burden imposed by the great number of required simulations (Shrestha et al., 2014), the method has been mainly criticized for the use of informal likelihood measures for statistical inference and for the choice of the threshold of some likelihood measure upon which models are considered as acceptable. One of the main issues is the fact that prediction limits generated by GLUE are not equivalent to statistical confidence limits. For example, Montanari (2005) showed, via a series of hydrological modelling experiments, that the 95% of GLUE prediction limits include only 62% of the observed data. The main argument against these objections relies on the fact that the use of formal likelihoods requires assumptions about the nature of the errors, which can be rarely justified in hydrological modeling (Liu et al., 2009). To overcome these issues Beven (2006) introduced the "*limits of acceptability*" approach as an alternative framework to assess models as behavioural or non-behavioural.

In real-time forecasting, the general GLUE approach has been also applied to produce uncertainty bounds in predictions of hydraulic models (Aronica et al., 2002, 1998, Pappenberger et al., 2007, 2005a, Romanowicz and Beven, 2003, 1998). In this case, the uncertainty is attributed in the error in upstream and lateral inflows, the coarse representation of floodplain geometry along with infrastructures, the simplistic representation of flow processes, the rough estimation of roughness coefficient and conveyance coefficients, as well as the errors in observed inundation maps with which the model is compared (i.e., air photographs, ground mapping, radar mapping etc.). Romanowicz and Beven (1998) applied the method to obtain the spatially and temporally varying probabilities of inundation of all points on a floodplain, taking into consideration the structural error of hydraulic model along with the uncertainties in the calibration of the parameter sets. Their method allows the updating of the posterior distributions of the predictions, using Bayesian conditioning upon the new available observations. It is worth noting that this process updates the likelihood of predictions, avoiding the need to update neither parameter values nor state variables of the model. The same authors proposed more recently a new framework to constrain the uncertainty in the prediction of flood risk by conditioning conveyance parameters upon observations of maximum inundation extend (Romanowicz and Beven, 2003). Their methodology has the following steps:

1. The number of hydraulic model parameters to be evaluated is chosen and, given the floodplain geometry and boundary conditions for an event, many simulations are made using parameter sets chosen randomly from prior distributions. For the first event, a uniform

sampling via Latin Hypercube method is conducted, given that no information on parameter distributions is available.

2. The predictions of maximum inundation for each realization are compared with the observed inundation maps. The comparison is conducted on the basis of a fuzzy-type measure that takes into account the uncertainty in the conversion of model predicted depths to inundation widths. Any non-behavioural simulations are rejected and the posterior probability distributions for the parameters are derived.
3. Posterior probabilities of inundation are calculated by forming cumulative distributions of predicted inundation over all the retained behavioural realizations for all points of interest on the floodplain.
4. Resampled marginal posterior distributions for the effective parameter values may then be used as prior distributions for the new event.
5. Likelihood values from different events may be combined using Bayes equation.

A complete framework to cascade uncertainty from weather forecasts down to rainfall-runoff model and in next step to flood inundation model was proposed by Pappenberger et al. (2005b). In their work, 52 rainfall forecasts (i.e., 50 ensembles, one high resolution deterministic forecast and a single control forecast at the coarser ensemble resolution) are used to account for meteorological uncertainty. The GLUE concept was employed to identify the behavioural rainfall-runoff (LisFlood-WB) and flood models (LisFlood-FP). Total uncertainty can then be estimated by combining all rainfall inputs with runoff predictions and flood inundation predictions. However, such a process would require significant computational resources due to the high number of combination and the distributed nature of the models. To alleviate this high computational burden, the behavioural rainfall-runoff simulations were classified into functional types (6 classes for the presented case study) and only representative parameter sets for each type were combined with rainfall forecasts to provide upstream discharge inputs (Pappenberger and Beven, 2004). These 52*6 runoff realizations form the input to the flood inundation model and ten different sets of roughness coefficients are utilized to account parameter uncertainty.

The GLUE procedure is usually employed and referred as an informal Bayesian approach (Smith et al., 2008a). The rigorous Bayesian inference is based on the Bayes formula that combines the prior distributions, which comprise prior knowledge, with a likelihood measure, with reflects how well the model predicts the available observations, to form the posterior distributions of parameters and model errors which are used to estimate the probability of predicting the next observation conditional on the model (Beven, 2012). As new data becomes available, Bayes formula can be applied sequentially by setting the posterior distribution at the end at one step as prior of the next step. The main difference to the GLUE method lies on the specification of a formal likelihood measure which implies explicit beliefs on the nature of the errors and hence assumptions about their statistical descriptions (Montanari and Koutsoyiannis, 2012). Typically, a Gaussian error model is assumed, implying that the residuals are normal-distributed, mutually independent and homoscedastic variables. However, the structure of model errors often does not conform to these assumptions, making the formulation of the likelihood measure a very challenging task around which a lot of debate has been raised (e.g., see Beven, 2012, 2006; Montanari, 2007; Montanari and Koutsoyiannis, 2012 and the references therein). Typical examples of Bayesian-based methods for assessing and estimating uncertainty are the Bayesian Total Error Analysis (BATEA) method by Kavetski et al. (2006), the Differential Evolution Adaptive Metropolis (DREAM) approach by Vrugt et al. (2009), the Maximum Likelihood Bayesian Model Averaging (MLBMA) by Neuman (2003) and the recently developed Dynamic Uncertainty Model by Regression on Absolute Error (DUMBRAE) by Pianosi and Raso (2012).

A general framework, named Bayesian Forecasting System (BFS), that takes into account both precipitation and hydrological uncertainties in real-time forecasting was studied by Krzysztofowicz (2002, 1999). The system produces short-term probabilistic river discharge forecasts combining probabilistic quantitative precipitation forecasts with a deterministic hydrologic model to simulate the

response of a river basin to precipitation. In the BFS, the total uncertainty is decomposed into precipitation uncertainty, which is assumed dominant, and hydrologic uncertainty, which encompasses all the other sources of uncertainty. The overall system is composed by three structural components: i) a *precipitation uncertainty processor* that maps precipitation uncertainty into output uncertainty under the hypothesis that there is no hydrologic uncertainty (Kelly and Krzysztofowicz, 2000), ii) a *hydrologic model uncertainty processor that* accounts for hydrologic uncertainty under the hypothesis that there is no precipitation uncertainty (Krzysztofowicz and Kelly, 2000), and iii) an *integrator* that blends the two sources of uncertainty into a Bayesian schema (Krzysztofowicz, 2001). The hydrologic uncertainty processor evaluates the probability of occurrence of the predicted flow rate via a Bayesian revision of prior information. The revision is based on the assumption that model performance, in terms of residual quantiles, for future events will be similar to that for similar past events. In this framework, model residuals for different forecast lead times over a number of past events are used off-line to train the processor. The processor is updated after each new event is occurred. The efficiency of the revision process depends significantly on the appropriate selection of the prior distribution. In the initial schema, the latter is obtained via a linear model that expresses future variables as a function of past ones. The posterior distribution derives by incorporating the information provided by the prediction of the hydrological model. The processor uses mixed distributions that are transformed into the multivariate Gaussian space through the normal quantile transformation.

2.2.1 Data assimilation techniques for dynamic error correction

The dynamic correction of model errors to constrain uncertainty in predictions is of high importance especially in real-time forecasting applications. In this framework, *data assimilation techniques* are used to update model predictions at each time step taking into consideration the deviation of forecasts from the observed values at previous steps along with real-time information on the current state of the system. Several of such techniques have been developed, aiming to update different variables, i.e., input variables, model states, model parameters and output variables (Madsen and Skotner, 2005). In all cases, the ultimate target of any data assimilation technique is the minimization of the forecast error (i.e., deviation of forecasts from observations) and the prediction variance (uncertainty) at “N step ahead” lead time for which a decision is required (Beven, 2009). This is achieved by updating either the error component directly or the parameters of the forecast model itself, as new data becomes available.

In real-time flood forecasting applications, data assimilation is usually conducted via the Kalman filter (KF) method and its extensions (the extended KF, the Ensemble KF, the Ensemble Kalman smoother) and the Particle filter. The Kalman filter methodologies, originally proposed by Kalman (1960) for the optimal control of linear systems, has been applied to update model states and model parameters in hydrological models. A comprehensive description of KF application for data assimilation on stochastic Transfer Function models is provided by Young (2002). As presented in Figure 4, the KF is a recursive two-step procedure that updates a model as new data becomes available, taking into account both model and data uncertainties. In the “predictor” step, the *time update* equations are used to project forward in time the current state $\hat{x}_{k-1}$ and the model error covariance estimates P_{k-1}, so prior estimates for the next time step are obtained ($\hat{x}_k^-, P_k^-$). Then, in the “correction” step, the posterior estimate $\hat{x}_k$ is obtained as linear combination of the prior estimate and a weighted difference between the actual measurement at that time z_k and the measurement prediction $H\hat{x}_k$. The posterior estimates are used as input to predict the new a priori estimates at the next iteration/step. The difference is weighted according to the Kalman gain K_k that is updated each time a new observation is occurred. The values of Kalman gain depends on the relative magnitudes of observation and model error.

In the simplest case, the KF assimilation methodology can be applied to update directly the model of the residuals between deterministic forecast and the observations. In this case the input data and the parameters of the model are assumed correct, and the residuals represent the lack of knowledge

on the true value of the system. If these assumptions do not hold then all sources of error should be corrected in the updating. Alternatively, one can allow the gain coefficient to evolve as random variable with values that will reflect whether the model is over- or under-predicting the observed variable (Beven, 2009).

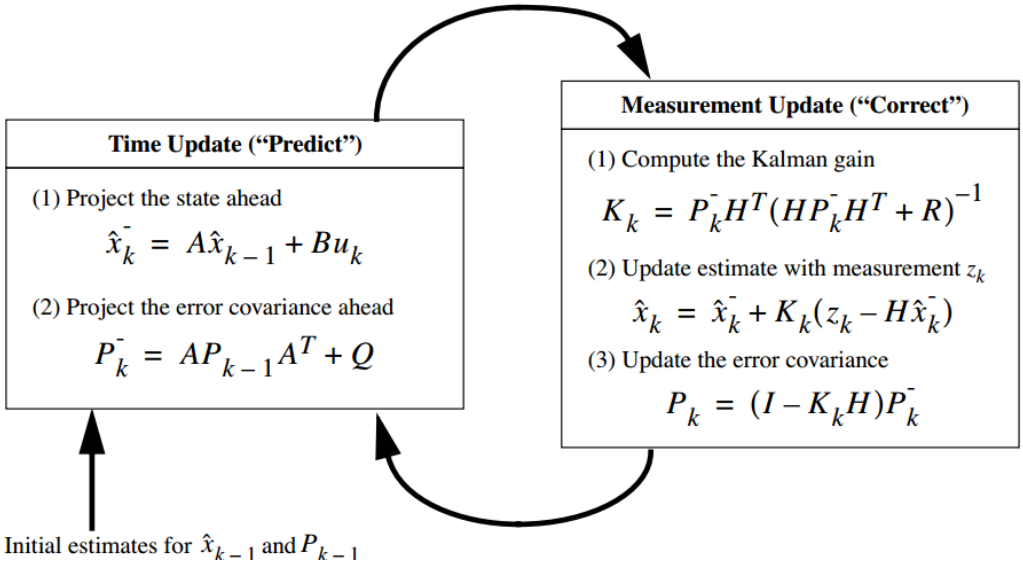


Figure 4: The Kalman filter recursive procedure

The above methodology, originally developed for linear models, was further extended to update and correct parameter and states of non-linear models. The *extended KF* is based on the linearisation of the time update equations around the non-linear model predictions between times *t* and *t*-1, using the partial derivatives of the process and measurement functions. Although the linearisation approach has been followed in some cases (Walker et al., 2001; Yeh and Huang, 2005), it is characterized by some shortcomings that restrict its wider practical application. These concern on the one hand the high computational burden in the case of large number of states and parameters, and on the other hand considerable numerical instabilities in the case of highly non-linear models (Beven, 2009).

A further extension of the KF procedure that overcomes the problems of the extended KF is the Ensemble Kalman Filter (EnKF), first introduced by Evensen (1994). In general, the ensemble approaches are evolved in a Bayesian framework where the posterior density of model states at previous step *t*-1, $p(\hat{x}_{t-1}|\hat{y}_{t-1})$, becomes the prior density at next step *t*, $p(\hat{x}_t|\hat{x}_{t-1}, \hat{y}_{t-1})$. This density is then updated via the likelihood function of new observations, $p(\hat{y}_t|\hat{x}_t)$, and the posterior density, $p(\hat{x}_t|\hat{y}_t)$,of model states is obtained according to Bayes formula (Hutton et al., 2014). In the EnKF, an ensemble of model states is obtained by randomly perturbating the input states according to a Monte Carlo sampling procedure that reflects the current error covariance matrix at the start of the period. At each updating step, the prediction-correction KF equations are applied to propagate each ensemble member forward in time, updating the distribution estimates (i.e., typically the multivariate normal is used) for model states and/or parameters (Beven, 2009). For the next forecasting step, the sampling is based on the updated distributions. This process allows the calculation of model error covariance from the ensemble sample directly, avoiding the high computational burden required to propagate the error covariance matrix in non-linear cases. Further to weather forecasting, the EnKF has been also applied to reduce predictive uncertainty in hydrological modelling applications (Clark et al., 2008; Moradkhani et al., 2005b; Vrugt et al., 2006; Vrugt and Robinson, 2007a; Xie and Zhang, 2010) as well as hydraulic models (Madsen and Canizares, 1999; Sørensen et al., 2004; Weerts and El Serafy, 2006). Critical factors for the successful implementation of the

EnKF methodology are the efficiency of sampling methodology and sufficiency of the number of members in the ensemble. This is of high importance especially in real-time forecasting applications where time constraints, along with the computational burden of model simulations, do not allow for an exhaustive sampling. Advances in sampling strategies aiming to improve the performance and efficiency of the original method were investigated by Evensen (2004, 2003). The Ensemble Kalman smoother, developed by Evensen and van Leeuwen (2000), extends the original EnKF by allowing the update of not only the current states in the model but also states at past time steps every time a new dataset becomes available.

Further to the family of Kalman-type filters, another well-known method for the dynamic correction and propagation of errors is the *Sequential Monte Carlo* approach, also known as *Particle Filtering* or *Bootstrap filtering*. This process implements a recursive Bayesian filter by MC simulations to update the state of the system (Arulampalam et al., 2001). The basic idea of the method is that the required posterior density can be represented by a set of random samples (particles) with associated weights and all estimates are based only on the samples and weights without requiring any assumption on the nature of the distributions (in contrast with the KF-type procedures). As the number of particles becomes very large, this MC process approximates the true posterior distribution adequately. At the first time step all samples have equal weights. The weights, typically expressed as a function of the likelihood measure, are updated as new observations are made available at the next forecasting realisation to reflect how well each sample was modelling each observation. The models are then propagated to the next forecast step and the weights are re-evaluated again (Beven, 2009). Various extensions and modifications of the original schema have been developed to cope with the high computational burden imposed by the sampling procedure (see reviews by Arulampalam et al. (2001) and van Leeuwen (2009)). In the framework of forecasting, application of this process can be found in Moradkhani et al. (2005a), Salamon and Feyen (2010), Weerts and El Serafy (2006) and Smith et al. (2008b).

The following Table 1 summarizes the types of uncertainty in each phase of an Early Warning System (EWS) and maps it to state-of-art tools and methods as well as indicative references which attempt to address the problem.

Table 1: Tools and methods relevant to address uncertainty in EWS

Relevant EWS phase	Source of Uncertainty	Methods and Indicative References
Meteorological Forecast	Error in NWP (stochastic physics and initial conditions)	**Ensemble Methods** (see reviews in Tracton and Kalnay (1993), Bowler *et al.* (2004), Buizza *et al.* (2005), Cloke and Pappenberger (2009), Hapuarachchi *et al.* (2011))**:** • Error breeding • Perturbed observations • Singular vectors • Ensemble Kalmar Filter
	Error in Nowcasts – Satellite and radar precipitation uncertainty	**Extrapolation methods:** Alfieri et al. (2012); Borga et al. (2011); Collier (2007) **Ensemble Methods:** Bowler *et al.* (2004) **Bayesian Methods:** Hapuarachchi *et al.* (2011) **Assimilation methods** (see Beven, 2009)**:** • Kalman Filter • Extended KF • Ensemble KF • Ensemble Kalman smoother • Particle filter
Hydrological/Hydraulic Modelling	Measurement Error (See Section 2.2)	McMillan et al. (2011, 2010); Montanari and Di Baldassarre (2013); Di Baldassarre and Montanari (2009); Sikorska et al. (2013); **GLUE**: Beven and Binley (1992), Pappenberger et al. (2007, 2005a) **Bayesian Methods**: DREAM, Vrugt et al. (2009); Krzysztofowicz (2002, 1999)
	Parameter Uncertainty (See Section 2.2)	Efstratiadis et al. (2015); Montanari and Brath (2004); Montanari and Grossi (2008); Sikorska et al. (2015); **GLUE**: Beven and Binley (1992), Romanowicz and Beven (2003, 1998), Aronica et al. (2002, 1998), Pappenberger et al., 2007, 2005a; **Bayesian Methods**: DREAM, Vrugt et al. (2009); Krzysztofowicz (2002, 1999)
	Structural Uncertainty (See Section 2.2)	Montanari and Koutsoyiannis (2012); **GLUE**: Beven and Binley (1992), Romanowicz and Beven (2003, 1998), Aronica et al. (2002, 1998), Pappenberger et al., 2007, 2005a; **Bayesian Methods**: DREAM, Vrugt et al. (2009); Krzysztofowicz (2002, 1999)

	Total Uncertainty (See Section 2.2)	Montanari and Koutsoyiannis, 2012; **GLUE**: Beven and Binley (1992); **Bayesian Methods**: BATEA, Kavetski et al. (2006); MLBMA, Neuman (2003); DUMBRAE, Pianosi and Raso (2012); Krzysztofowicz (2002, 1999)

3 A methodological framework for uncertainty analysis

The proposed in this study spatial and temporal resolution (1km/1hr) may be ideal for monitoring storms that lead to flash floods. An open question remains how reliable and accurate such data can be, particularly for sensitive applications (e.g., flash flood warning, decision making, etc.). Due to the importance of high resolution and reliable precipitation data to current and future research and applications, there is a need to investigate more multi-spectral and multi-satellite precipitation retrieval techniques in order to obtain the best possible approximation of precipitation with the highest resolution possible. As we now stand at the doorstep of a global-scale precipitation mission, named Global Precipitation Measurement (GPM) we are expecting improvements in the accuracy of precipitation estimation (available online at http://gpm.gsfc.nasa.gov/; Smith et al. 2007). Even though GPM does not directly contribute in producing high resolution precipitation estimations, it may be used for downscaling to finer resolutions or training of the high resolution IR-based precipitation estimates. In this case, Geostationary Operational Environmental Satellite-R (GOES-R) can provide the spectral information required to produce precipitation data with spatial (2km) and temporal (15 min) resolution (Sorooshian and AghaKouchak, 2010).

The first step in this study has been the selection of the National Oceanic and Atmospheric Administration (NOAA) Climate prediction center (CPC) MORPHing method (CMORPH) 8-km/1/2-hr gauge-adjusted (Joyce et al. 2004) satellite data (i.e., in our study between the 2002 and 2011). This method exhibits the best consistency among 10 different satellite rainfall products and as the proposed methodology of this study it could be also applied to the future satellite high-resolution precipitation products (i.e., GPM). Besides the precipitation satellite observation methods, quantitative rainfall forecasts from numerical weather prediction (NWP) models can be employed by a hydrological model for generating flood predictions (Nikolopoulos et al. 2013; Nikolopoulos et al. 2010). Nevertheless, the accuracy of precipitation fields derived from NWP forecasts suffers from spatiotemporal and amplitude errors depending on the model physics, dynamics, and model configuration (Schwartz et al. 2010; Bray et al. 2011). Generally, NWP initialized with analysis data can provide a reliable estimation of the synoptically forced rainfall, while satellite observations can better represent the convective rainfall temporal variability (Ebert et al. 2007). Zhang et al. (2013) used the Weather Research and Forecasting Model (WRF) with data assimilation to downscale satellite rainfall estimates for hydrological applications. Their results show that the WRF Ensemble Data Assimilation System can lead to improved high-resolution precipitation estimates when adding satellite observations into the system.

Finally, this study uses the non-hydrostatic high-resolution WRF NWP modelling system (version 3.7.0), with the dynamical core of Advanced Research WRF (WRF-ARW) (Matsangouras et al. 2015) addressing different storm types (low, medium and deep convection precipitation) and ground stations (i.e. rain gauges) used for the statistical evaluation (both satellite and NWP estimates). The simulation runs have been initialized using operational data with 0.25^{o} spatial resolution and 6-h temporal resolution fields from the European Centre for Medium-Range Weather Forecasts (ECMWF) analysis. The WRF simulations have been performed in a two-way interactive mode with 35 vertical levels and a three domain configuration, in which the coarsest spatial resolution is 6-km and two nested domains have the resolution of 3-km and 1-km, respectively. One would intuitively argue that the reliability and accuracy of forecasts depends not only on the quality and resolution of weather prediction models but also to a high degree on the density and quality of input like ground observations, soundings at smaller scales and satellite observations on larger scales.

3.1 Methodological framework and study area

a) Uncertainty analysis methodological framework

In this section we present the applied methodology addressing the aspects mentioned above.

Step 1: Production of additional realizations (i.e. ensembles) using statistical methods

The first step in this procedure is to match the estimates (in our case these are the deterministic forecasts) with the reference (either radar or satellite) precipitation data and bring them to a common spatial and temporal resolution fitting grid (Figure 7) in order to determine the error statistics.

Given a precipitation forecast of an event at a dense or very dense horizontal grid resolution (e.g. 1h/10km), a methodology has been developed providing a set of additional realizations with respect to the probability distribution of a number of parameters influencing the event. More specifically, for each storm event the following actions have been taken:

a) Disaggregation of all forecasted data to match the spatial and temporal resolution of the reference rainfall observations i.e. radar, measuring stations, based on the weighting factor data matching methodology from Kirstetter et al. 2012
b) Comparison of the forecasted with the reference rainfall data
c) Independent evaluation and ranking of each time step of the forecasted precipitation data with respect to their predictive accuracy, and
d) Statistical comparison of satellite data with NWP by plotting their corresponding quantiles (Q-Q).

A fixed cumulative distribution function (CDF) bin with cumulative probability values ranging from 0.05 up to 0.95 has been defined to determine the corresponding NWP and Satellite hourly rain rate quantiles. An adjustment function based on power-law has been introduced for best fit of the NWP versus the satellite hourly rain rate quantiles:

$$Y = \mathrm{a}\ X^{k} \qquad \text{[Eq. 1]}$$

The least squares approach has been used as fitting method.

This is similar to what in Zhang et al. 2013 has shown, where the Y and X represent the NWP forecast and Satellite hourly rain rates (in mmh^{-1}) respectively. The estimated parameters (a and k) of the function have unique and distinct values for each deterministic forecast. This way a number of relationships (i.e. adjustment functions) has been produced, based on the available storm events. The range of these relationships represents the forecasting uncertainty (in space and intensity) as each relationship gives a different forecasted field.

Step 2: Error analysis and evaluation of the precipitation forecasts

The evaluations of the high-resolution deterministic and ensemble precipitation prognoses (based on the CMORPH-rainfall observation retrievals) are based on rainfall observations from the ground stations. The proposed error assessment approach based on metrics such as the mean field bias adjustment and the pdf matching has been applied on the satellite datasets (see section 3.3). Additional events, not used in the first phase, have been selected for error analysis. The adjustment functions elaborated in the step 1 have been used to adjust the respective NWP deterministic forecasts and create several forecast realizations (ensembles). Finally, a number of error metrics have been defined including two common verification scores, the bias score (BS) and the Heidke skill score (HSS). Based on these indicators, the precipitation ensembles have been evaluated against time series collected from rain gauge stations.

b) Study area and data

The proposed methodology will be applied in the coastal study area of Greve in Denmark (Figure 1). This region often suffers from heavy precipitations, floods and storm surges. Eighteen different precipitation events ranging from low, moderate to severe storms have been selected to run NWP models with precipitation forecasts. Table 2 shows the NWP integration periods of the selected events which vary from 6 to 72 hrs. The model output files were analysed at hourly time intervals.

Table 2. Time periods of WRF simulation

WRF Simulation	Starting time	Ending time	Duration (h)
1	01 Aug 2002 1800 UTC	02 Aug 2002 0000 UTC	32
2	09 Jan 2007 0000UTC	11 Jan 2007 0000 UTC	48
3	05 Jul 2007 0000 UTC	06 Jul 2007 0000UTC	24
4	26 Jul 2007 1800 UTC	28 Jul 2007 0000 UTC	48
5	28 Jul 2007 0000 UTC	30 Jul 2007 0000 UTC	48
6	10 Jun 2009 0600 UTC	11 Jun 2009 0000 UTC	18
7	11 Jun 2009 0000 UTC	14 Jun 2009 0000 UTC	72
8	17 Aug 2010 0000 UTC	18 Aug 2010 0000 UTC	24
9	02 Jul 2011 1800 UTC	04 Jul 2011 0000 UTC	30
10	29 Jun 2012 0600 UTC	01 Jul 2012 0000 UTC	18
11	24 Sep 2012 0600 UTC	24 Sep 2012 2300 UTC	18
12	25 Sep 2012 0000UTC	27 Sep 2012 0000 UTC	48
13	15 Jun 2013 0600 UTC	17 Jun 2013 0000 UTC	18
14	10 Sep 2013 1200 UTC	11 Sep 2013 0000 UTC	12
15	14 Jul 2014 0000 UTC	15 Jul 2014 0000 UTC	24
16	31 Aug 2014 0000 UTC	01 Sep 2014 0000 UTC	24
17	13 Sep 2014 1200 UTC	14 Sep 2014 1200 UTC	24
18	19 Sep 2014 1200 UTC	19 Sep 2014 1800 UTC	6

For the verification of the precipitation estimates, the Danish Meteorological Institute (DMI) has provided in-situ precipitation observations from rain gauges. They are located, as shown in Figure 5, within or close to the NWP simulation domain. The number next to some of the rain gauge symbols indicates the station code. Only stations with bold codes have been used in this case study as only for these stations time series for the selected period, 1976 until 2012, were available. The black shading box with the dotted line delineates the selected domain of NWP simulations

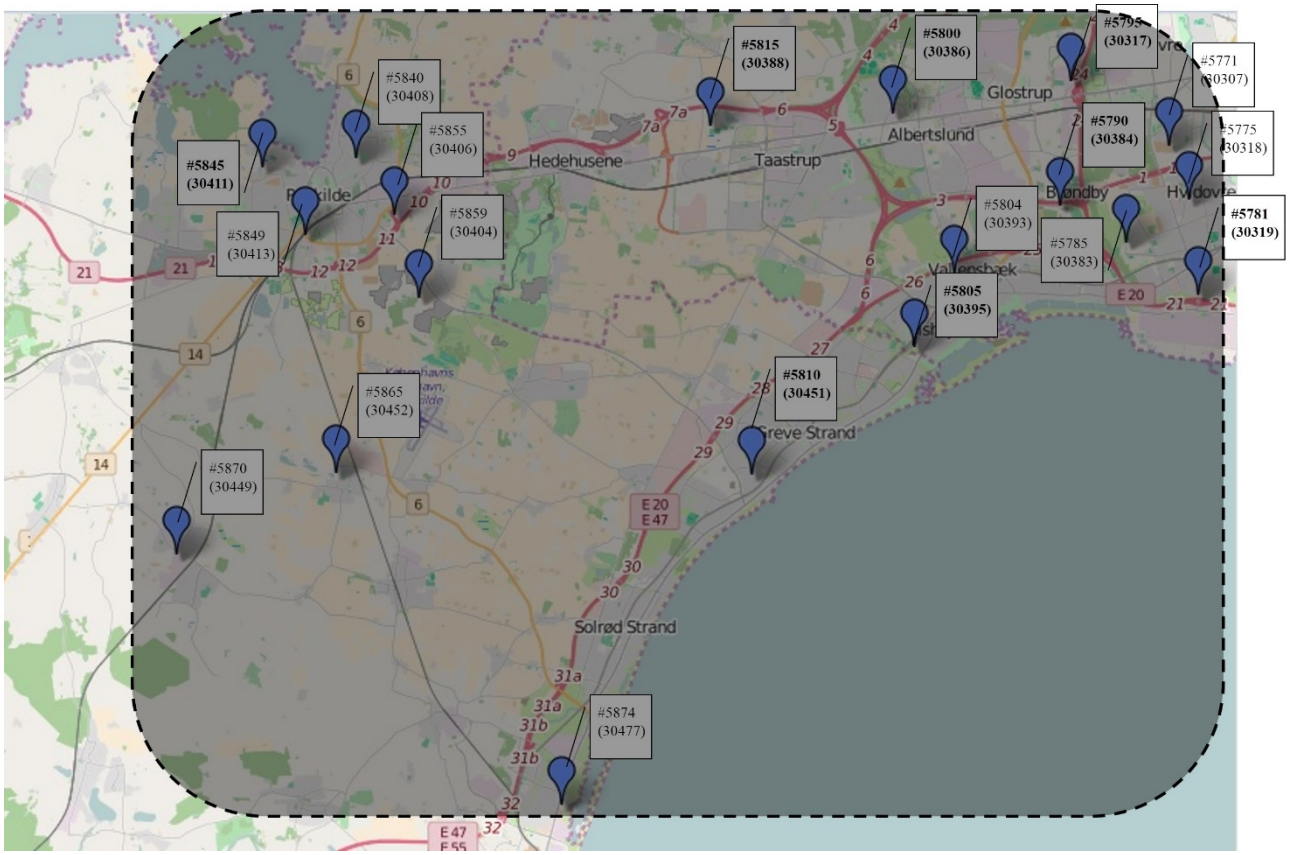


Figure 5: Map showing the study area with the location rain gauge stations operated by the Danish Meteorological Institute (DMI).

Figure 6 shows time series of the rainfall rate ($mmhr^{-1}$) for the selected rain gauges, used for the evaluation of the error in the rainfall estimates, resulting from the high-quality, rain gauge-adjusted CMORPH satellite retrievals. CMORPH is a satellite rainfall algorithm that uses motion vectors derived from half-hourly interval geostationary satellite IR imagery to propagate the relatively high-quality rainfall estimates obtained from earth orbiting satellite-based passive microwave (PMW) sensors (Joyce et al. 2004). The dynamic morphological characteristics (such as shape and intensity) of the precipitation features are morphed at consecutive times between PMW sensor samples by performing a time weighted linear interpolation. This process yields spatially and temporally continuous areal precipitation that has been guided by IR imagery and yet is independent of any IR temperature–based inversion to rainfall rate (Kidd et al. 2012).

However, there are certain issues with this technique as highlighted by Joyce et al. (2004): (i) the algorithm may miss precipitation estimates that are developed over an area between satellite (PMW) overpasses and (ii) the current snow-screening process gives nonzero rainfall estimates to the snow or ice areas, thus causing inauthentic observations over areas that are usually associated with high elevations and mountainous terrain. These deficiencies introduce uncertainties to the CMORPH precipitation estimates, which propagate in flood prediction when these estimates are used as input to a hydrologic model. Nevertheless, as indicated by several satellite error studies (Stampoulis and Anagnostou 2012; Kidd et al. 2012; Bitew et al. 2012; Dinku et al. 2007), the CMORPH methodology results in better precipitation estimates than other satellite products.

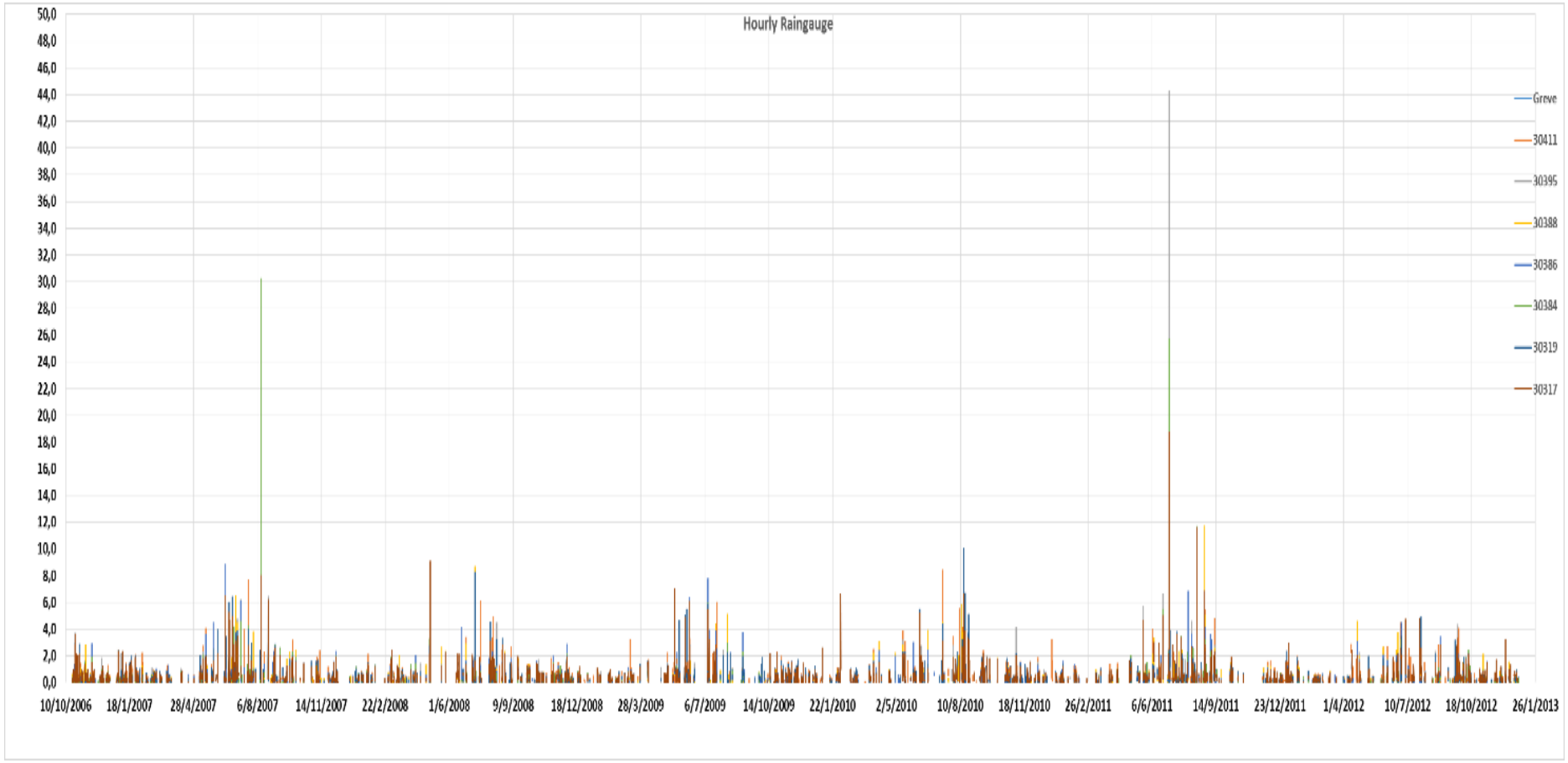


Figure 6: Hourly rainfall (mm) measurements from 2007 until the end of 2012 taken from the eight selected different rain gauge stations (shown in Figure 5) that has been used for data validation and error assessment

3.2 Results

The selected domain that has been defined in this study follows two criteria: (i) it is small enough to focus on the area of interest (i.e. covering the whole watershed in Greve) and therefore to represent the distinctive precipitation features associated with the satellite CMORPH-retrievals, and (ii) at the same time, it is large enough to follow intense rainfall events in the study area. This is significant in order for the WRF-CMORPH relationships to be specified based on the characteristics of the selected storm events.

The eighteen storm events investigated in this study are distinct in terms of the rainfall intensities and the spatiotemporal rain structures, providing a good representation of different storm types occurring in the Greve coastal area. Figure 7 shows the accumulated rainfall maps for eight selected storm events. In each pair, the left map is the one resulting from the CMORPH methodology, and the right one from WRF prognoses.

The fitting relations between the WRF versus the CMORPH rainfall estimation are shown in the Q-Q plots. They exhibit a non-linear relationship. Thus, as proposed in Zhang et al. 2013, a power-law equation [Eq. 1] has been selected to relate the deterministic WRF rainfall estimates with the CMORPH satellite estimates and produce for each rain storm additional relations (i.e. realizations) of areal rainfall estimations. Table 3 shows the parameters a and k of [Eq 1] by event, which give the best fit. It is important to note that the adjustment method does not correct the rainfall detection error, which may be a significant factor in satellite rainfall underestimation over mountainous areas and near the north and south boundaries of the satellite scan.

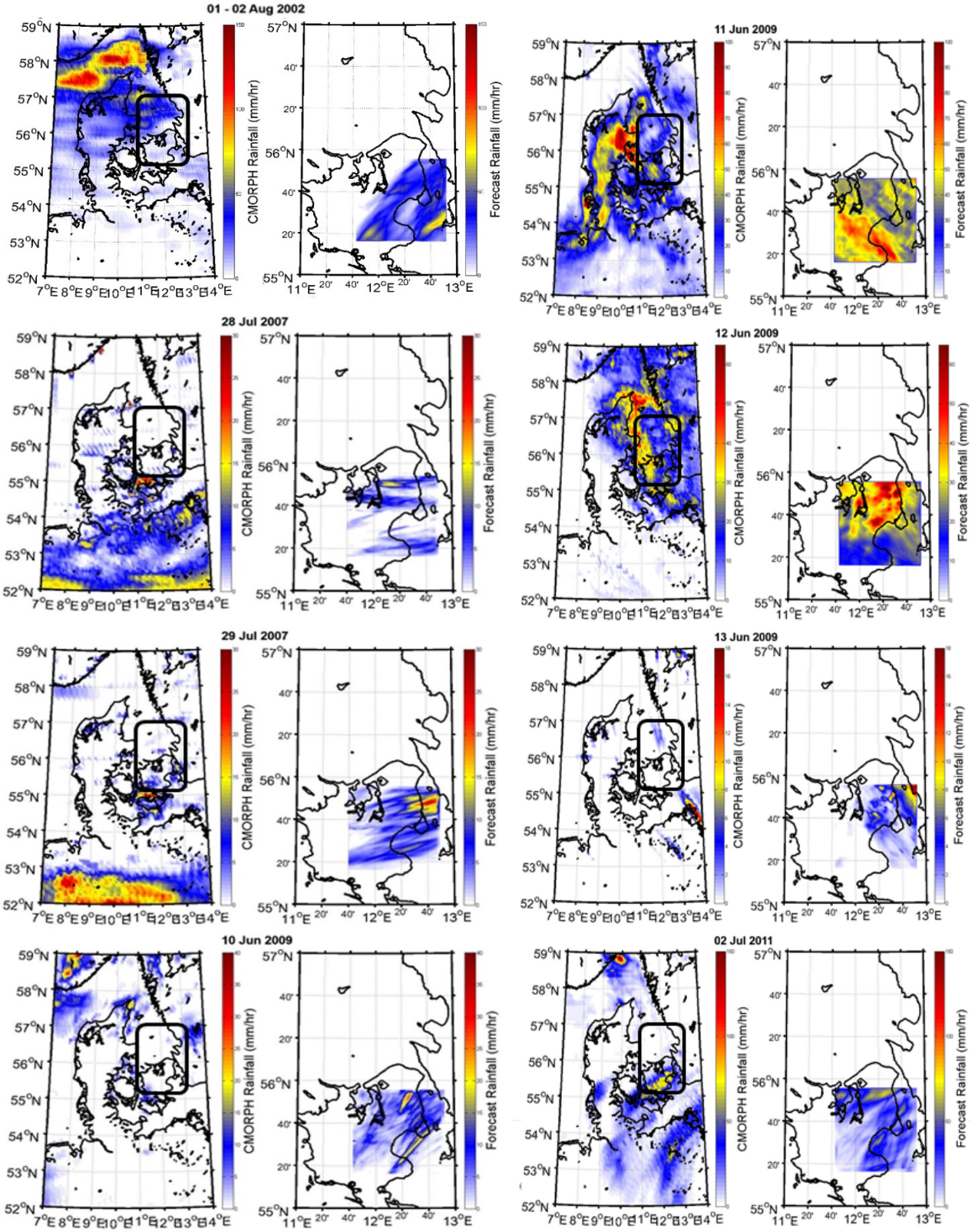


Figure 7: Accumulated rainfall maps of some selected storm events with high (>20 mm) rainfall

Table 3: Fitting parameter values

Storm Date/Parameter	*a*	*k*
02 Aug. 2002	9.07	0.42
05 Jul. 2007	4.65	0.63
26 Jul. 2007	0.91	1.31
27 Jul. 2007	2.02	0.64
28 Jul. 2007	2.41	0.46
29 Jul. 2007	3.30	0.40
10 Jun. 2009	2.51	0.40
11 Jun. 2009	16.6	0.41
12 Jun. 2009	30.8	0.31
13 Jun. 2009	0.91	0.78
17 Aug. 2010	1.23	0.71
03 Jul. 2011	13.0	0.21
29 Jun. 2012	0.52	1.19
24 Sep. 2012	10.6	0.27
25 Sep. 2012	4.17	0.35
26 Sep. 2012	6.62	0.27
15 Jun. 2013	8.26	0.20
10 Sep. 2013	7.70	0.40
14 Jul. 2014	9.42	0.30
31 Aug. 2014	9.72	0.33
13 Sep. 2014	7.03	0.41
13 Oct. 2014	8.50	0.22
19 Oct. 2014	7.42	0.22

Overall, the 1-km resolution of WRF and the 8-km resolution of CMORPH rainfall accumulations show similar patterns for the different storm events. Considering the fitting parameter values (as shown in Table 3) and the respective Q-Q plots on Figure 8, three storm events (i.e. 05/07/07, 11/06/09 and 29/06/12) may indicate a linear relationship approximation between the datasets. They show not only a strong linear but also an almost unbiased relationship (the fitting is close to the "1-1" line) between the WRF deterministic prognoses and the CMORPH rainfall estimation. In the contrary, the rest of the events show a strong non-linear relationship (i.e. power-law relationship).

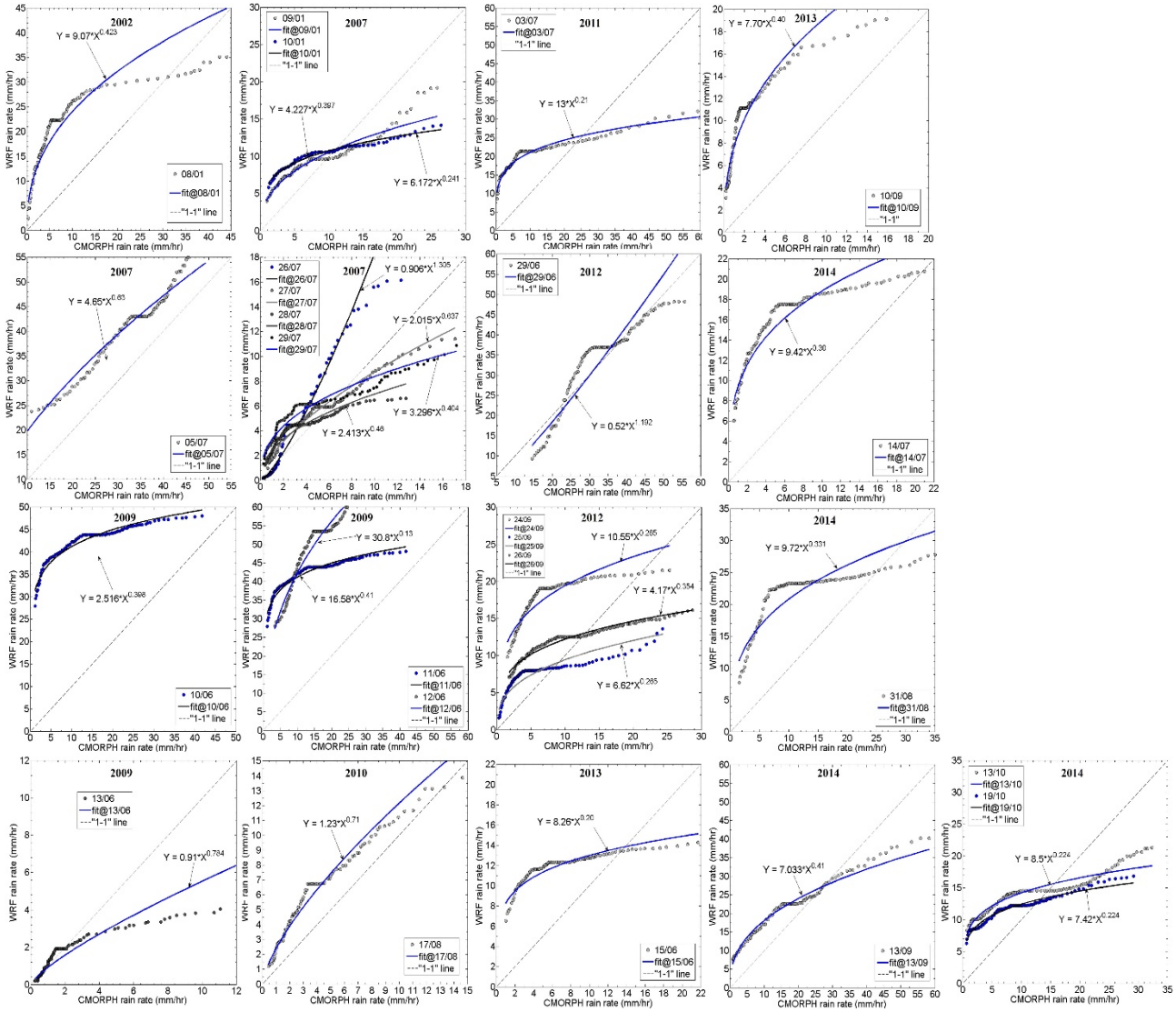


Figure 8: Q–Q plot of WRF rain rates vs CMORPH rain rates

In order to account for uncertainty of satellite precipitation data, characteristics of error should be investigated first. Understanding the spatial and temporal dependencies of errors plays a key role in developing uncertainty models. Possible metrics used to understand the multivariate characteristics of errors are the *Probability distribution function (pdf)* and the *variance of errors*. By providing reliable information on products' uncertainties, users can make adjustments to their modelling approach or apply a decision making method "on-the-fly". Therefore, the methodology presented in this study needs to be statistically evaluated with new deterministic forecasts and compared with the corresponding satellite and ground rainfall observations.

Ignoring the distribution information in bias adjustment procedure could potentially result in loss of valuable information, especially with regard to the tails of the distribution. In hydrologic applications, watershed response significantly depends on the tails of rainfall intensities. Therefore, PDF-based adjustment techniques deserve more in-depth future research.

3.3 Validation and statistical error assessment of forecast realizations

As mentioned in the previous sections, the time and space resolution of the satellite rainfall (i.e. CMORPH, TRMM, etc.) estimates should be carefully selected to match with the deterministic

forecast resolution. Similarly, the time and space resolution of the reference rainfall (i.e. ground stations precipitation observations) should be carefully matched to the resolution of the deterministic and ensemble forecasts. The validation analysis in this study is provided over the model output domain. The ground station measurements are hereafter named “reference” data, while the rainfall forecasts and ensembles will be referred as “estimates”.

At first, a number of error metrics for the statistical assessment of the methodology have been selected including Bias, Mean Bias (MB), the relative Root Mean Square Error (rRMSE) of the estimated versus the reference data. These statistical metrics are defined as follows:

$$Bias = \left(\sum_{i=1}^{N} ref. - \sum_{i=1}^{N} est.\right) \quad \text{[Eq. 2]}$$

$$MB = \frac{1}{N}\left(\sum_{i=1}^{N} ref. - \sum_{i=1}^{N} est.\right) \quad \text{[Eq. 3]}$$

$$relative\ RMSE = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(ref.-est.)^2}}{\frac{1}{N}\sum_{i=1}^{N} ref.} \quad \text{[Eq. 4]}$$

In addition, next to the basic statistical analysis, the Bias Score (BS) and the Heidke Skill Score (HSS) have been used as verification scores. They estimate the performance of the “estimators” and the different ensembles. To calculate these metrics, initially the reference and estimated rainfall rates are averaged over the eight different ground stations. Then, the following occurrences are counted: the number of hits (defined as H: estimator > threshold and reference > threshold), the number of false alarms (defined as F: estimator > threshold and reference < threshold), the number of misses (defined as M: estimator < threshold and reference > threshold), and the number of correct nulls (defined as C: estimator < threshold and reference < threshold), for data pairs exceeding four different rainfall rate thresholds: 0.5, 1.0, 1.5 and 2.0 mmhr^{-1}.

BS is defined as the ratio of the number of occurrences that estimated rain rates exceed a specified threshold versus the respective number from the reference rain rates:

$$BS = \frac{H+F}{H+M} \quad \text{[Eq. 5]}$$

HSS is defined as the number of correct estimated occurrences minus the number of correct estimated occurrences by chance divided by the total number of estimated occurrences minus the number of correct estimated occurrences by chance:

$$HSS = \frac{2(H\times C - F\times M)}{(H+M)(M+C)+(H+F)(F+C)} \quad \text{[Eq. 6]}$$

Theoretically, the range of HSS is $-\infty$ to 1. A perfect precipitation estimator would have an HSS value equal to 1, while HSS less than or equal to zero indicates that the technique gives mostly a random estimation or has fewer hits than a random estimation. HSS is a widely used score because it is simple, fairly easy to compute and it may explain more than one effect such as probability of detection, false alarm rate, and occurrences by chance.

For validation purposes and statistical error assessment three events with different characteristics have been selected: 01 - 02 August 2002, 05 July 2007 and 17 August 2010. The time step of the reference time series if hourly. The results of each storm event are described in the next sections.

a. The storm event of 1st August 2002

This event lasted only a few hours in the morning of the 01 August 2002. It was characterized by intense rainfall mainly between 06:00 and 10:00 UTC (Figure 9a). Figure 9b documents the total daily accumulated rainfall (mm) of the ground observations (continuous black line), the deterministic forecasts (blue continuous line) and the different ensembles (grey shaded area) The grey area

covers the spectrum of rainfalls from the smallest to the highest estimates and visualizes the uncertainty of rainfall estimation.
In the same way, Figure 9a gives the rainfall rate as time series from eight different locations, which correspond to gauging stations. From these two figures, it is evident that the ensembles envelope the uncertainty of rainfall magnitude compare to the observations, which is very important for flood forecasting.

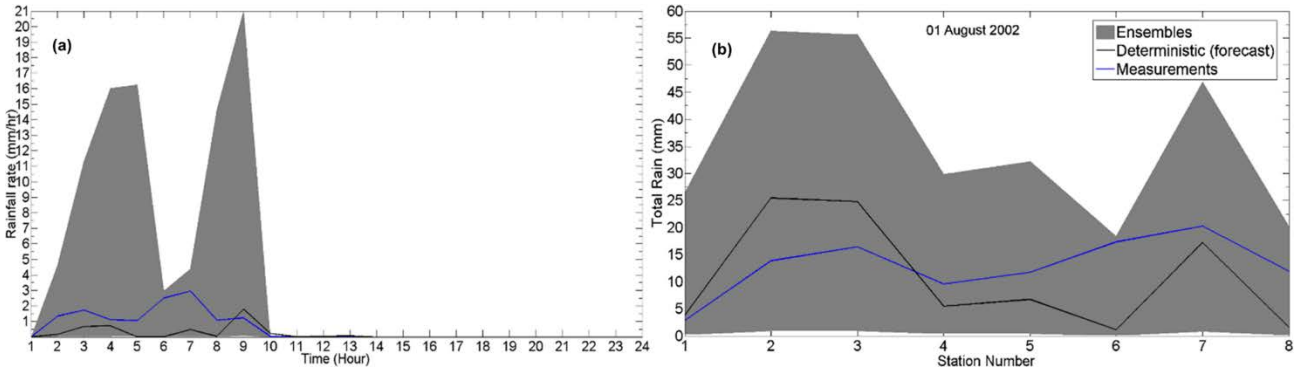


Figure 9: (a) time series of rainfall rates (mmhr^{-1}) and (b) total rainfall (mm) by station for the event of 1st August 2002

Figure 10a and Figure 10b show a bias difference ranging between -60 mm and 75 mm and a mean bias difference between -8 mm and 8 mm of the majority of the ensembles (black squares) with respect to the ground station observations. The blue line indicates the bias difference of the deterministic forecast. Figure 10c indicates a low (between 0.35 up to 0.95) relative rRMSE of almost all ensembles, including the deterministic forecast. It is important to note that the bias is calculated as the total reference rainfall (in mm) minus the estimated, one.

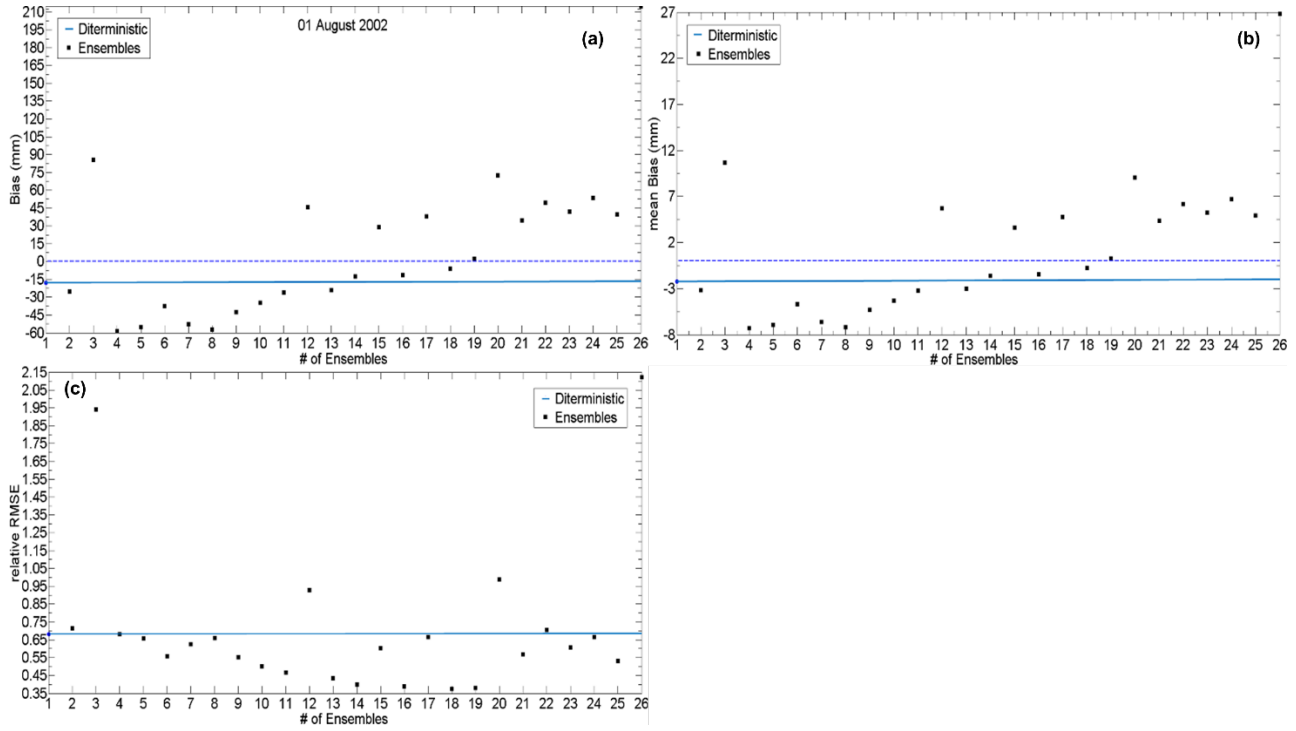


Figure 10: (a) bias difference, (b) mean bias difference and (c) relative root mean square error (rRMSE) of total estimation and reference for the event of 1st August 2002

In addition to the above, the BS in Figure 11a and HSS in Figure 11b have been used in this study to test the details of the deterministic and ensembles forecasts at different hourly rainfall rate thresholds. It is clear that the BS values of the deterministic (the first value from the left side of the x-axis) and the next ca. 12 ensembles representing the lowest rainfall rate thresholds are closest to 1, compared to the remaining ones. This means that those estimates show more similar behaviour with the reference rainfall observations, especially for rainfall rate thresholds greater than 0.5 $mmhr^{-1}$. In addition, the HSS plot shows the highest values (around 0.4) for rainfall thresholds above 0.5 $mmhr^{-1}$. This indicates that the ensembles forecasts, which achieve these scores, not only perform best in estimating rainfall occurrence but also in terms of rainfall magnitudes.

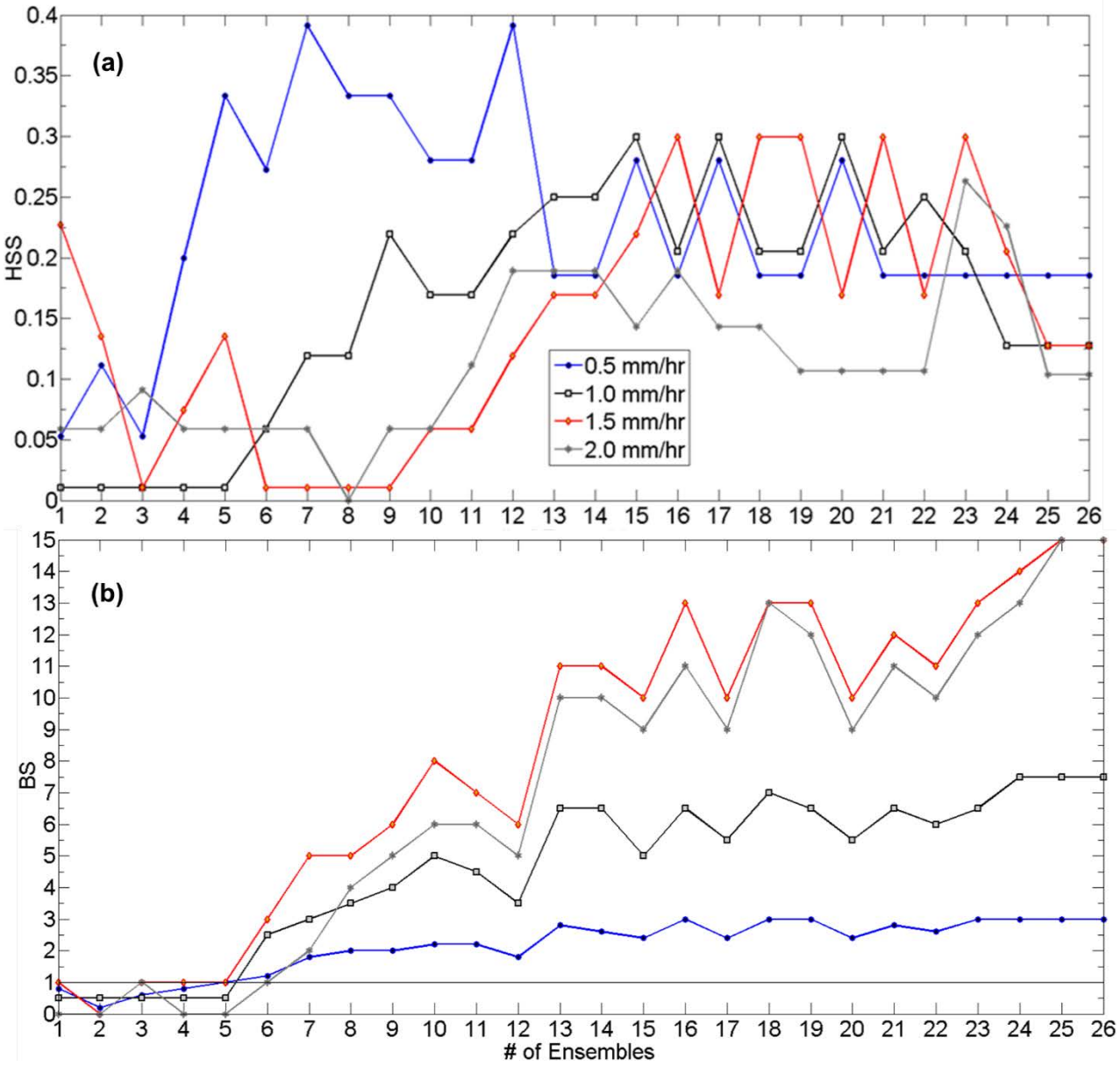


Figure 11: (a) HSS and (b) BS plots for various rainfall rates for the event of 1st August 2002

b. The storm event of 5th July 2007

The event of 5th August 2007 started in the early morning hours and lasted until the end of the day, producing moderate rainfall rates with its peak intensity between 15:00 and 21:00 UTC (Figure 12a).

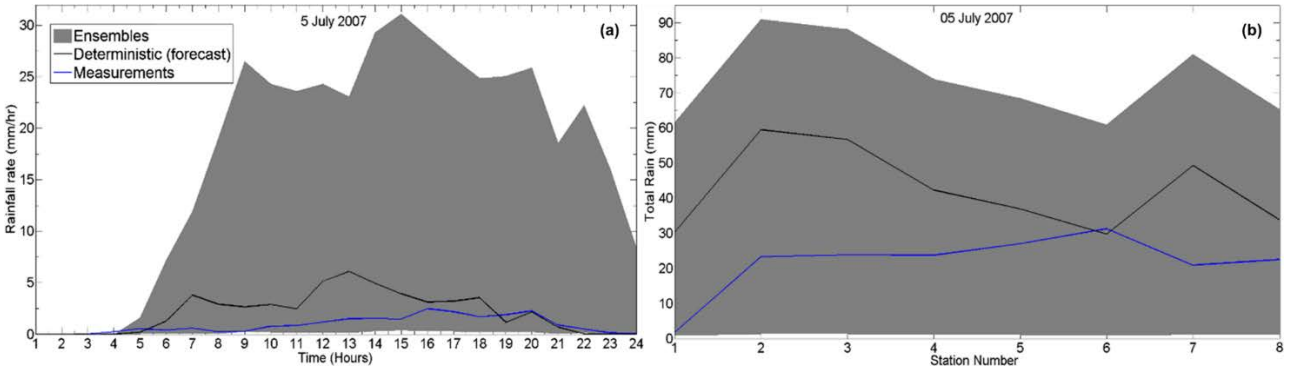


Figure 12: (a) time series of rainfall rates (mmhr^{-1}) and (b) total rainfall (mm) by station for the event of 5th July 2007

The deterministic forecast indicates high bias and mean bias difference as shown in Figure 13a and Figure 13b. The bias difference of the ensemble forecasts range between +/- 90 mm and the mean bias +/- 15 mm. However, it has to be noted that for this event most ensemble forecasts underestimate rainfall. The relative RMSE of most of the ensemble forecast is below the 50% with the deterministic close to 100%, indicating a small error in the variability of the forecasts.

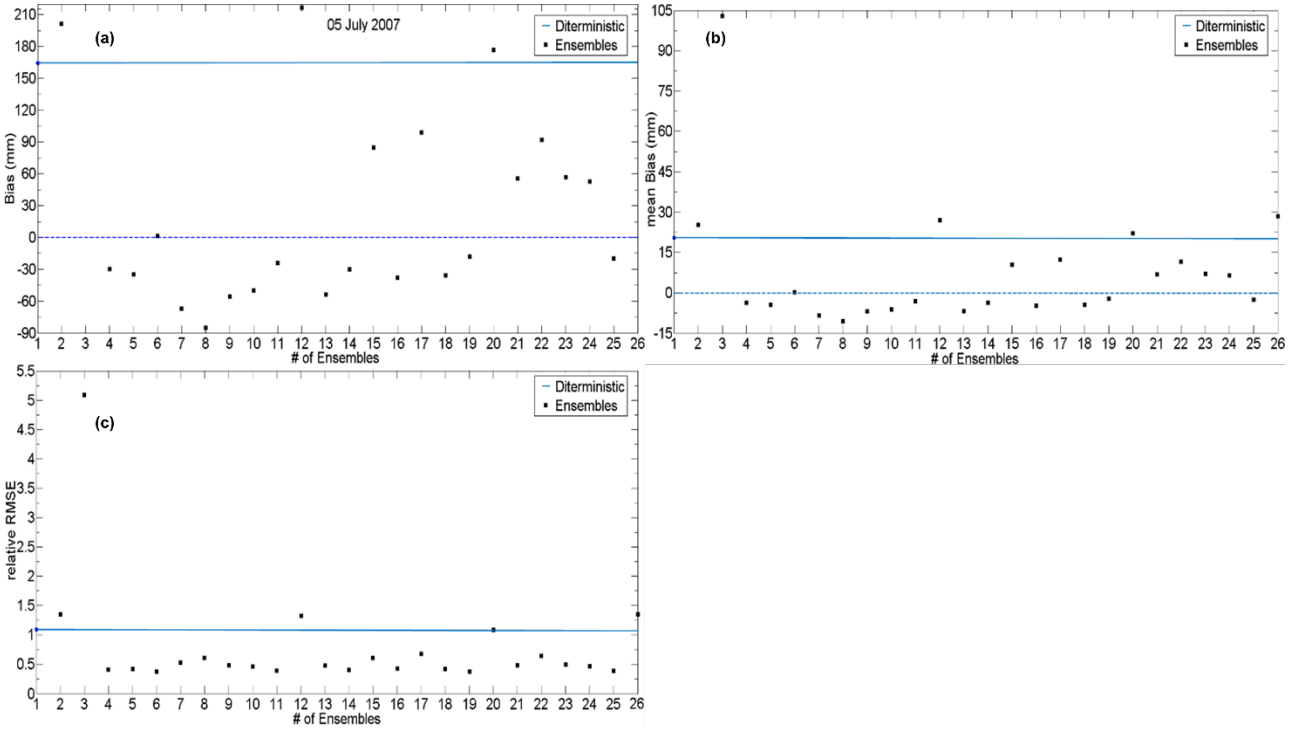


Figure 13: (a) bias difference, (b) mean bias difference and (c) relative root mean square error (rRMSE) of total estimation and reference for the event of 5th July 2007

The HSS in Figure 14a, is always above zero and most of the ensembles above 0.5 and close to 1.0, indicating that the deterministic and ensemble estimators are associated with low random error and high probability of detection. This is associated with the high percentage of detection between the estimator and the ground observations. This is also evidenced by the bias score (in Figure 14b) in which the estimators (both deterministic and ensembles) exhibit a bias score of +/- 25% variability around the value one.

It is also significant that the first 8-9 ensembles, for rainfall rate thresholds above the 2.0 mmhr^{-1}, exhibit a bias score close to one. Moreover, as it was expected, for almost 80% of the estimators, the bias score is close to one, for rainfall rate thresholds above the 0.5 mmhr^{-1}. However, it should be noted that the selected ensembles with smaller rainfall rate thresholds, have better bias scores compared to the ones with higher rainfall thresholds.

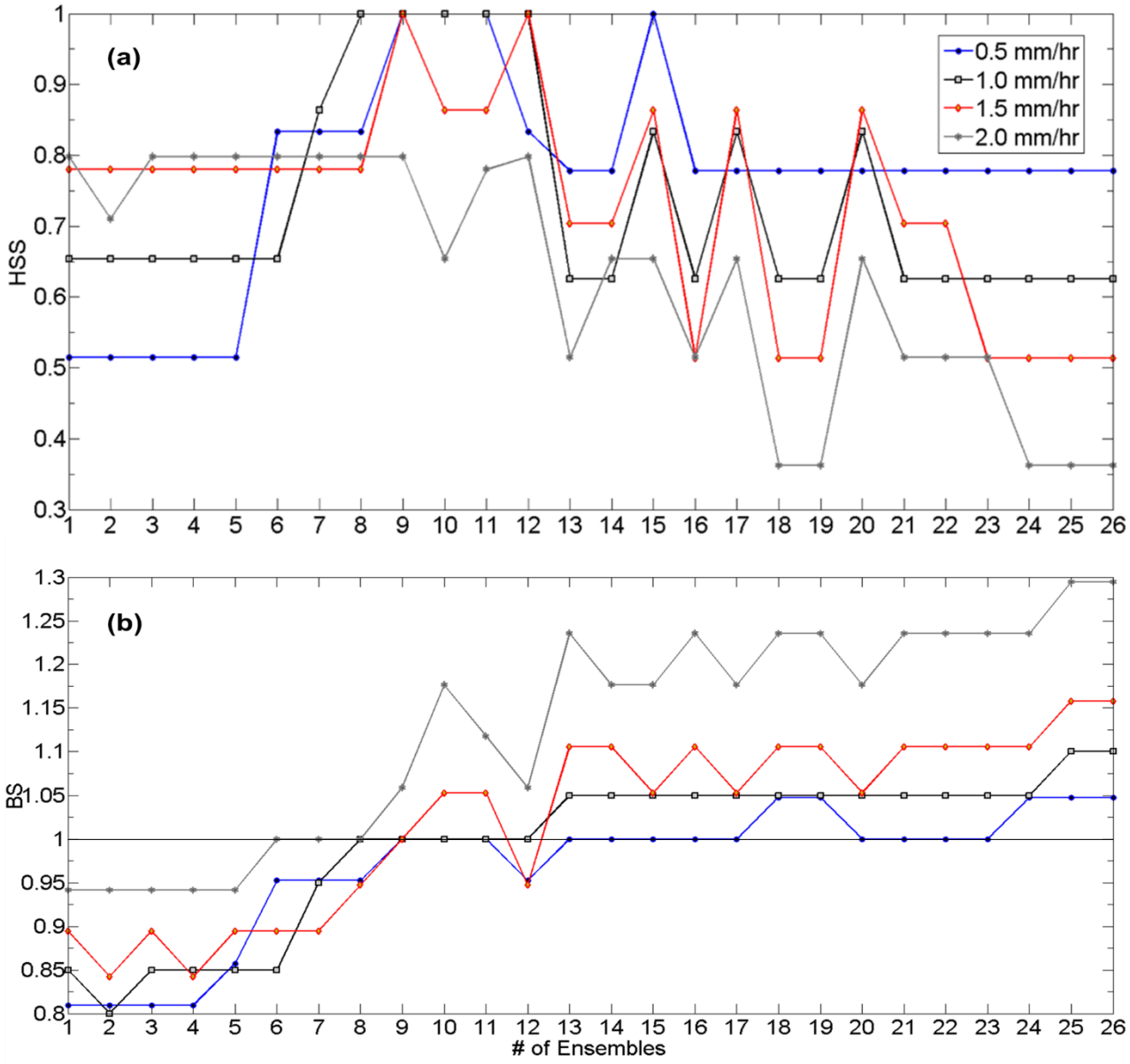


Figure 14: (a) HSS and (b) BS plots for various rainfall rates for the event of 5th July 2007

c. The storm event of 17th August 2010

This selected rainfall event, shown in Figure 15a, started at 06:00 UTC on 17th August 2010. It is associated with moderate rainfall values, as shown in the time series of the rainfall rates in Figure 15a. Although there are no rainfall observations in the late afternoon hours, some ensembles estimate some rainfall. After the 10:00 UTC the deterministic forecast exhibit similar rainfall values to the observations, indicating low bias difference and mean bias difference (Figure 16a and Figure 16b). The range of values for most of the ensembles is +/-40 mm for the bias and +/-5 mm for the

mean bias. In addition, the relative RMSE in Figure 16c is below 80% for most of the ensembles and the deterministic forecast, indicating moderate relative error.

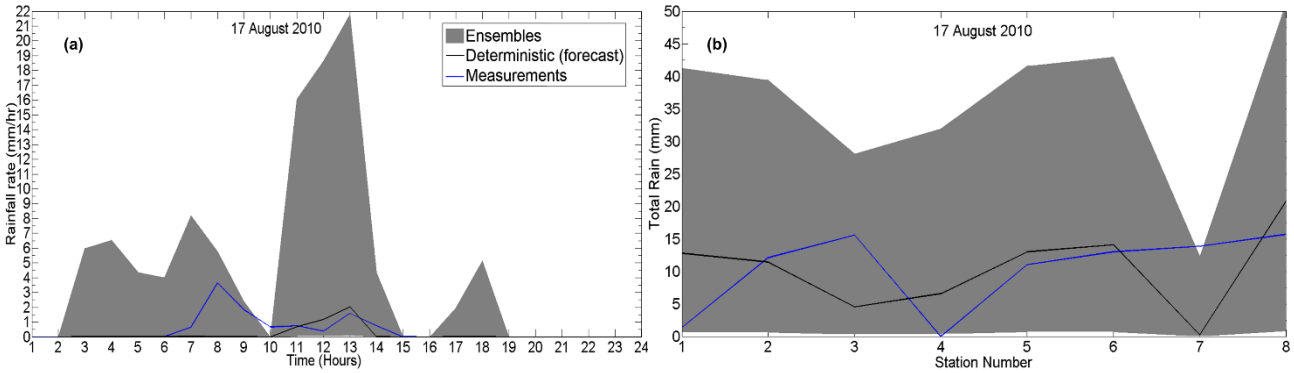


Figure 15: (a) time series of rainfall rates (mmhr^{-1}) and (b) total rainfall (mm) by station for the event of 17th August 2010

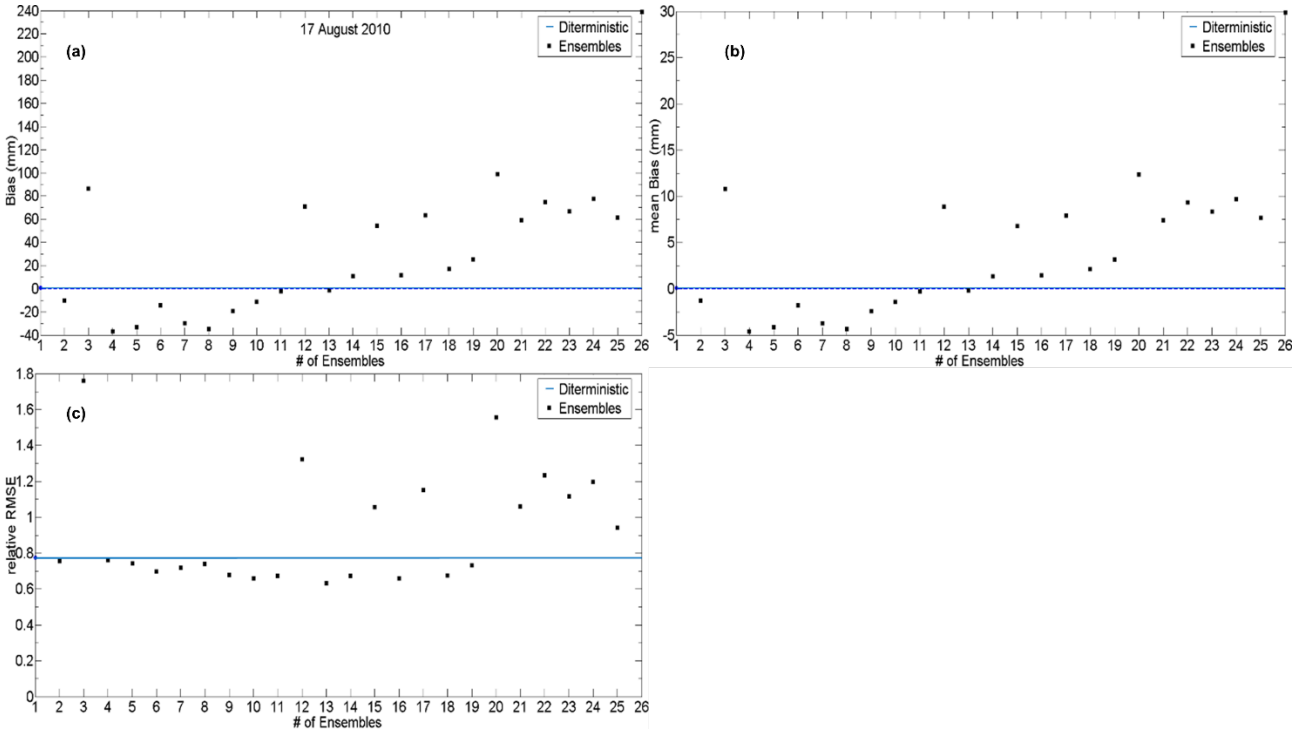


Figure 16: (a) bias difference, (b) mean bias difference and (c) relative root mean square error (rRMSE) of total estimation and reference for the event of 17th August 2010

The HSS score in Figure 17a indicates that almost all ensemble forecasts have values over 0.2 for rainfall rates threshold above 1.0 mmhr^{-1}, while most of the ensembles and the deterministic forecast are also over 0.2 for the rest of rainfall rates thresholds. In addition, the bias score of all ensembles and the deterministic forecast value shown in Figure 17b cover a range of +/-60% around the value one, indicating that most of them have low to moderate bias dependence on rainfall. We also note that most ensembles provide particularly good estimation only for the higher rainfall rate thresholds.

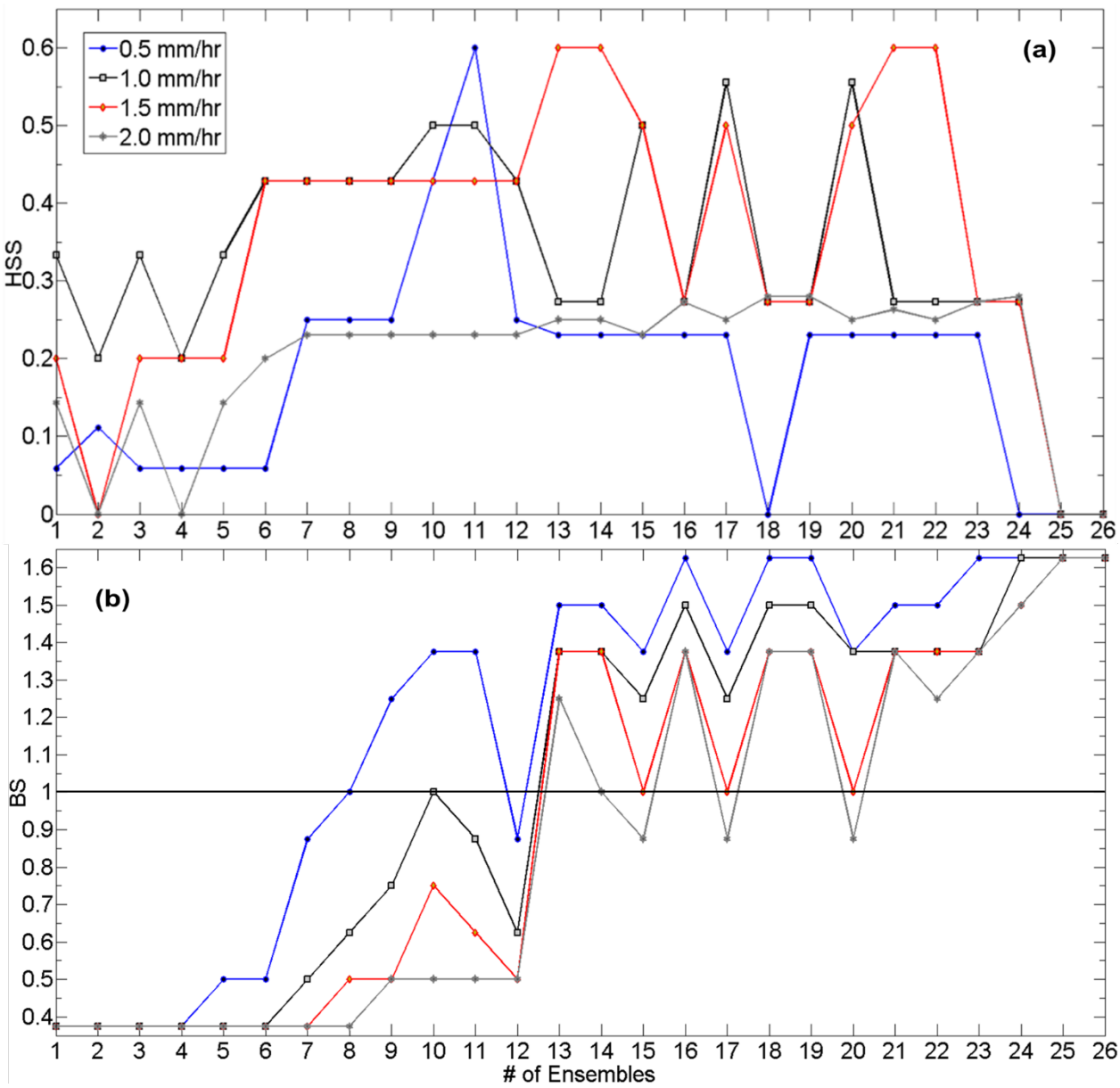


Figure 17: (a) HSS and (b) BS plots for various rainfall rates for the event of 17th August 2010

3.4 Nowcast and Radar precipitation uncertainty

As discussed in section 2 NWP models provide a precipitation forecasting for the following days with temporal resolutions of about 1h and spatial resolutions ranging from 20-50 km to 1-3 km in the high-resolution local models. However, this temporal and spatial resolution might not be enough to capture the local phenomena like thunderstorms leading to urban floods or flash floods in small catchments. In this context is where radar precipitation estimates and nowcasting can play a role in anticipating potential dangers.

Although precipitation accumulation estimates based on radar observations may suffer from different sources of error, they may also be the quickest way to obtain a reliable estimate of the accumulated precipitation, and of the precipitation that may occur in the forthcoming few hours (by means of radar-based nowcasting). This short-term forecasting is crucial for urban floods or flash floods where the response times are usually very short.

Wilson et al. (2010) evaluated eleven different radar nowcasting systems *showing that* systems based on linear extrapolation of the field precipitation were generally more accurate than weather models and systems that perform a combination of meteorological models and extrapolation *(blending)* up to 3 hours from time of forecast.

Radar nowcasting has the resolution of the radar precipitation observations (typically 1 km^2 and 5-10 minutes time step) and the leadtime where the skills remain high is between 1-6 hours depending on the radar domain and the local characteristics. The leadtime is limited compared to NWP models, but the rapid update of the data when using nowcasting (nowcasts are recalculated every time that new radar data is available) qualifies it for usage in EWS for prediction of short-term rain phenomena with good results (see for example Berenguer et al 2005 and Berenguer et al 2011).

Even if this technique provides the best available forecasts for short term, radar-based nowcasting techniques are not free of uncertainty. This is mainly due to the uncertainties in the radar estimates themselves and to the fact that these techniques take into account the movement of the precipitation but not its evolution. Recently, several techniques based on the ensemble concept have been developed to represent the uncertainties in the radar estimates/forecasts (Germann et al., 2006, 2009; Llort et al. 2008, Pegram et al. 2011).

In this section, a methodology to represent uncertainties in radar precipitation fields and to generate radar probabilistic nowcasts is presented, with the aim of generating probabilistic hydrometeorological specific warning products in the identified potential high-risk areas of the Marbella Pilot site.

3.4.1 Deterministic Radar nowcasting

The first operational nowcasting system was implemented in 1974 using data from a weather radar of the McGill University (Austin and Bellon, 1974). Currently there is a variety of operational nowcasting systems.

The simplest approach to make a very short-term rainfall forecast is simply to keep the last available observation constant during forecasting time. This method is known as Eulerian persistence.

To improve such forecasts, the movement of the precipitation field produced according to the latest observations is used. These motion fields not only provide a means for monitoring the precipitation, such as rotation or translation, but also the convergence and divergence of the vectors that represent the precipitation growth and decrease. This methodology is known as Lagrangian persistence.

To calculate the motion field of the precipitation the first studies were made simply by comparing two precipitation fields. In those studies, as described by Wilson (1966), the average displacement vector was obtained that maximizes the correlation between the two fields of precipitation. Rinehart and Garvey (1978) introduced the TREC algorithm, implementing the method of cross - correlation but in different subdomains of the precipitation field (which represent the resolution of the velocity field for the different subdomains), thus obtaining a non-uniform and much more realistic field of motion. Li et al. (1995) presented a postprocessing of TREC called COTREC to eliminate divergences between different subdomains and produce a more comprehensive motion field (see Figure 18).

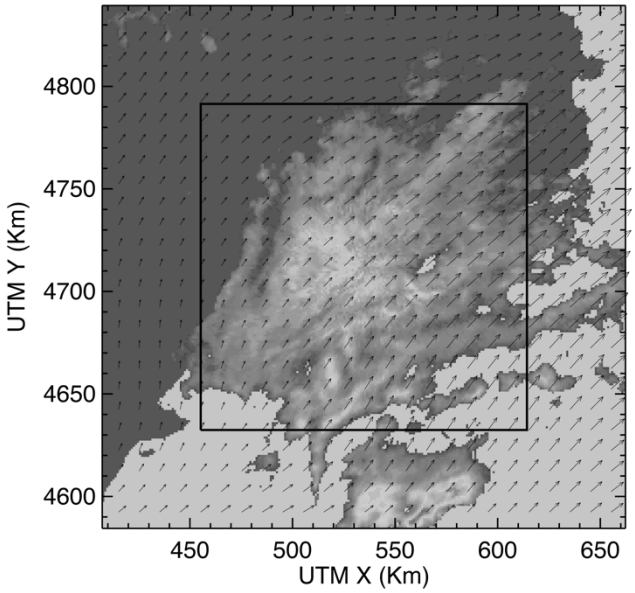


Figure 18: Example of motion field obtained through the COTREC (Li et al. 1995) algorithm.

Once the motion field is known, Germann and Zawadzki (2002) compared two approaches to advect the observations into the future. On the one hand the *forward* process (extrapolating *forward),* where the field of precipitation is advected forward in time and on the other the *backward* process (extrapolating *backwards),* which asks what value in the past should be assigned to the current position; The authors concluded that the backward process is better in the sense that it is exhaustive (a value can be predicted for every single pixel of the precipitation field). See Figure 19 as an example of the nowcasting technique. The Figure shows, for two different events, the nowcasting done at different leadtimes compared to the real measurements.

The two main sources of uncertainty in the radar nowcasting are the uncertainty in the motion field (or evolution of it) and the uncertainty related to the evolution of the precipitation field (last observations are not evolved in the process).

One of the first attempts to characterize these two sources of uncertainty through probabilistic algorithms forecast was carried out by Andersson and Ivarsson (1991). They proposed to carry out probabilistic forecasting by sampling different areas around a point of interest and advecting those fields to the future. Based on this idea, Germann and Zawadzki (2004) proposed the method known as local Lagrangian.

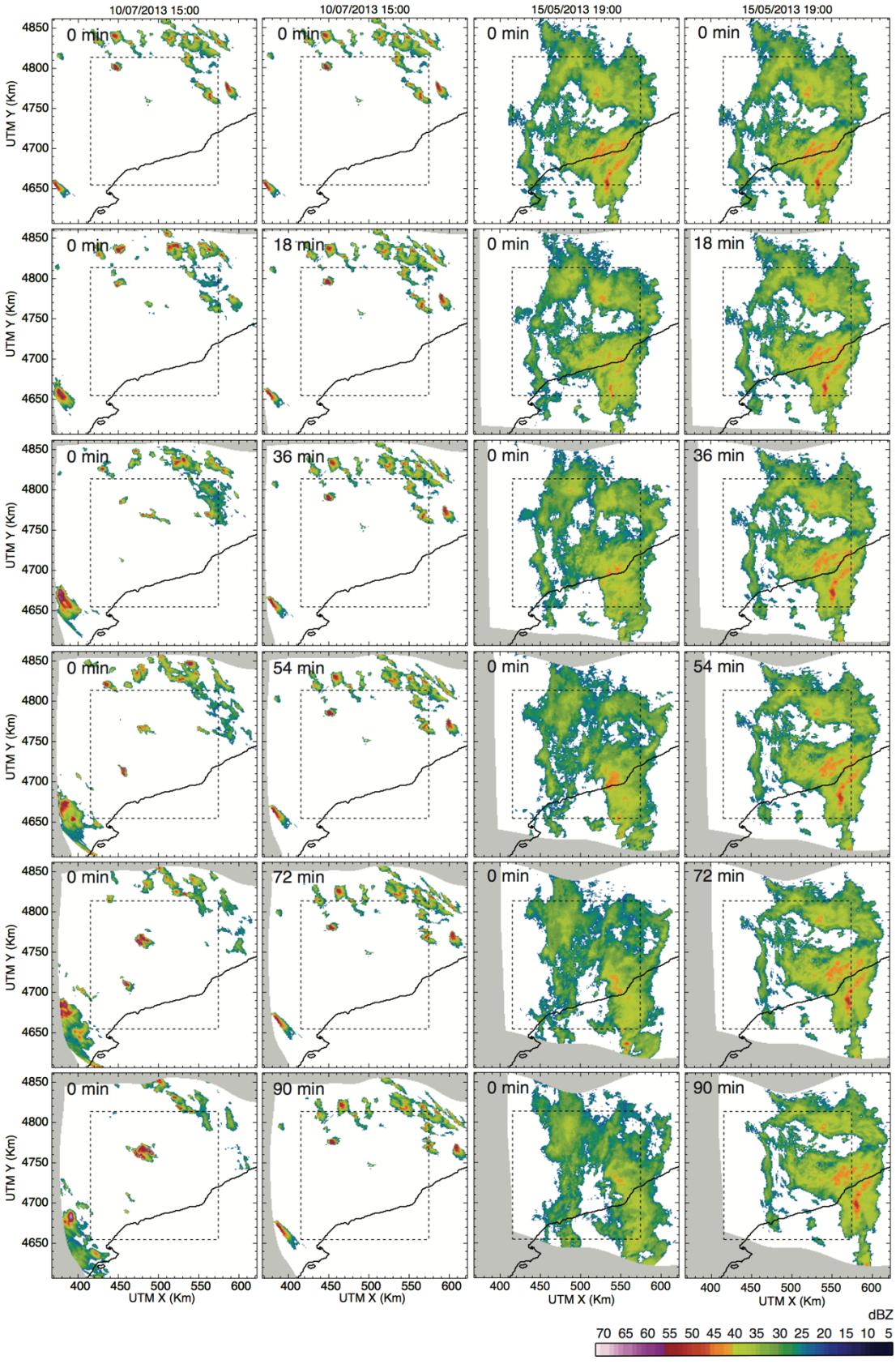


Figure 19: Example of nowcasting. First and third columns show the observations, the second and the fourth the forecasts at different leadtimes.

3.4.2 Probabilistic Radar Nowcasting

In the Local Lagrangian method proposed by Germann and Zawadzki (2004) a single distribution function is no longer obtained for the entire domain, but for each point of the domain. Apart from the local (spatial) character of the algorithm, a hypothesis is imposed from a temporal point of view too. It consists of thinking that the temporal evolution of the precipitation field remains constant over time (Eulerian persistence).

In this method, a set of forecasts probability distribution functions (pdf) are determined for each point of the domain as a function of subdomain size and threshold precipitation. It is therefore necessary to have a methodology to choose for each forecast lead time the optimal subdomain in order to minimize the error. Germann and Zawadzki (2004) described a methodology that uses the *Conditional Square root RANKED Probability Score* (CSRR), based on the Rank Probability Score (Wilks, 1995), as an estimate of the goodness of all probabilistic forecasts made. This estimator provides a probabilistic error value between forecasts and observations associated with each subdomain. As shown in the paper of Germann and Zawadzki (2004), the choice of the optimal subdomain for each forecast leadtime is chosen according to the minimum value of the probabilistic error (CSRR) for each subdomain sample.

Figure 20 shows by the example of two events how optimal sampling of subdomains typically increases over forecast leadtime, reflecting the dependence of the subdomain size from the optimum predictability of rainfall.

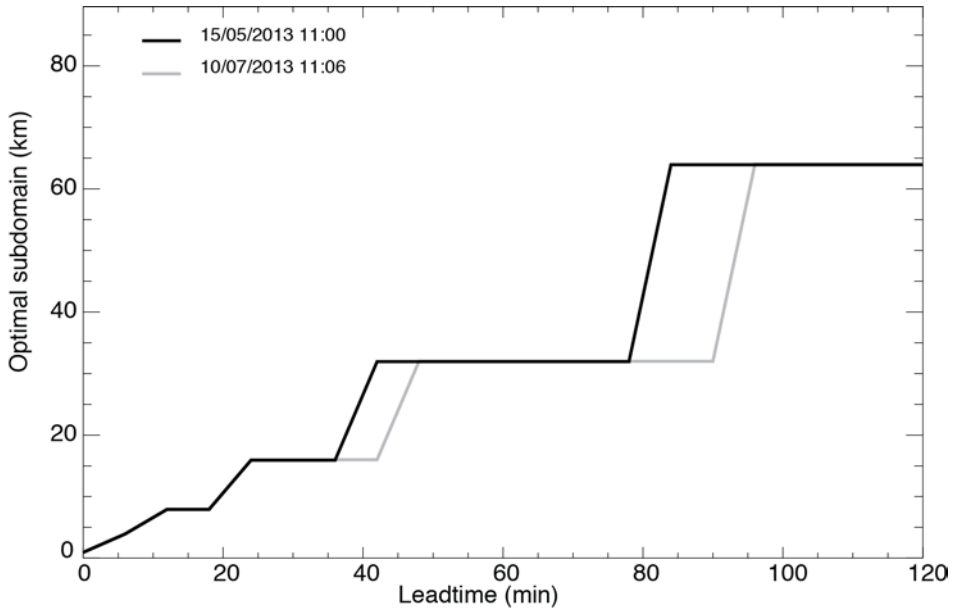


Figure 20: Estimation of the optimal subdomain size as function of the leadtime by the example of two events.

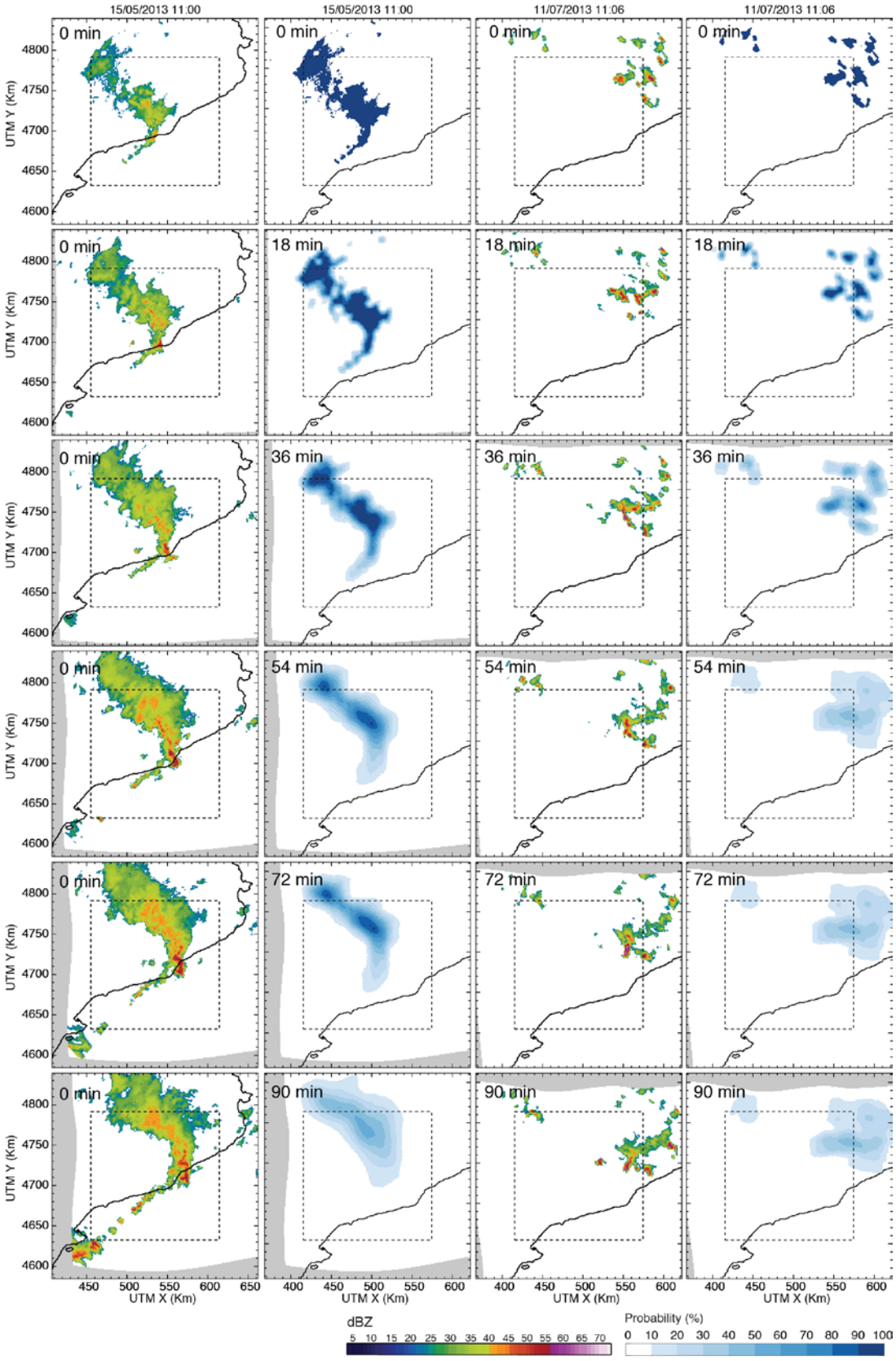


Figure 21: Example of Local Lagrangian nowcasting. First and third columns show the observations, the second and the fourth the probabilistic forecasts at different leadtimes. (from Buil 2017).

Figure 21 shows an example of the probabilistic forecasts obtained by the Local Lagrangian technique with forecasts every 6 minutes up to a maximum of 2 hours. In this case a precipitation threshold of 1 mm/h has been used. Two events typical for the Mediterranean climate (10-11/07/2013 and 15/05/2013) have been selected. In the first two columns of Figure 21, observations and probabilistic forecasts are presented for the period from 05/15/2013 11:00 am to 1:30 pm (90 minutes). The most intense precipitation of the first available observation is located in the central zone of the domain and in a small area to the northwest registering weaker intensities. The precipitation field moves in a northwest direction with increasing intensity, although two different zones can be observed, one convective and the other stratiform. In the Local Lagrangian forecasts, it is observed that at 18 minutes it is quite accurate and with precipitation probability values (>1mm/h) of around 80-100%. This reflects little uncertainty in the forecasts since the optimum sub-domain sample is 8x8 km^2 as shown in Figure 20. The forecast calculated based on the observation of about one hour (54 minutes) shows how the forecast uncertainty increases and leads to a smoothing of the probability field (optimal sample subdomain of 32x32 km^2) and a decrease in the maximum probability values. At 90 minutes the uncertainty in the forecasts has increased, in this case so much that the optimum sub-domain sample has reached the maximum size, i.e. 64x64 km^2. The predicted probability field has been further smoothed with maximum values between 50-60%. The second case, shown by the two columns on the right in Figure 21, is convective. It shows convective cells formed independent of each other and spread in two areas of the domain, one in the northwest and the other in the northeast. Throughout the two hours one can observe how the rainfall activity concentrates gradually and moves to the south very slowly in parallel to the increase of precipitation intensity (reaching values higher than 20 mm/h). The propagation of forecast uncertainty is similar to the 15/05/2013 11:00 event, with a 80-100% accuracy for the first 18 minutes and lower values later.

The smoothing of probabilistic forecasts with lower probability values as forecast lead time increases is due to the very nature of the forecasting algorithm. This algorithm is conceived with the idea of characterizing the uncertainty associated to the predictions of Lagrangian persistence through the minimization of error (CSRR) associated to a subdomain. This optimal subdomain is unique for all precipitation thresholds due to the definition of CSRR.

The main drawback of the Local Lagrangian methodology is that it predicts a certain probability of precipitation only in those places where in the last observation sampled at the optimal subdomain size, the probability is not negligible. Another drawback is that it does reproduce with a representative probability value the initiation of convective nucleus.

Buil 2017 did a comparison of several probabilistic nowcasting techniques, including those based on the generation of ensembles, concluding that the Local Lagrangian technique outperformed the other techniques in almost all cases. Consequently, and since the main objective was to implement in real time a technique able to provide probabilistic short-term rainfall forecasts based on radar, the Local Lagrangian technique has been chosen for feeding the EWS developed for the Marbella Pilot site.

3.5 Model chain uncertainty

This subtask will explore error propagation in a model chain used for flood forecasting, In Figure 22 is shown an example of such a model chain that considers pluvial and fluvial flooding as well as flooding from the sea. The model chain involves rainfall forecasts (see above), rainfall-runoff and urban drainage models, forecast of open sea boundaries, and a hydrodynamic storm surge and flood inundation model. The resulting uncertainty in the flood forecast is due to a number of uncertainty sources:

- Uncertainties in the system data used for setting up the different models, e.g. uncertainties in the DEM, bathymetry, and drainage network
- Uncertainties in forecasts of model forcing (or boundary conditions), i.e. rainfall, open sea boundaries, and runoff boundaries from upstream catchments

- Model structural uncertainties, including process descriptions and model conceptualisations in the entire model chain
- Model parameter uncertainties

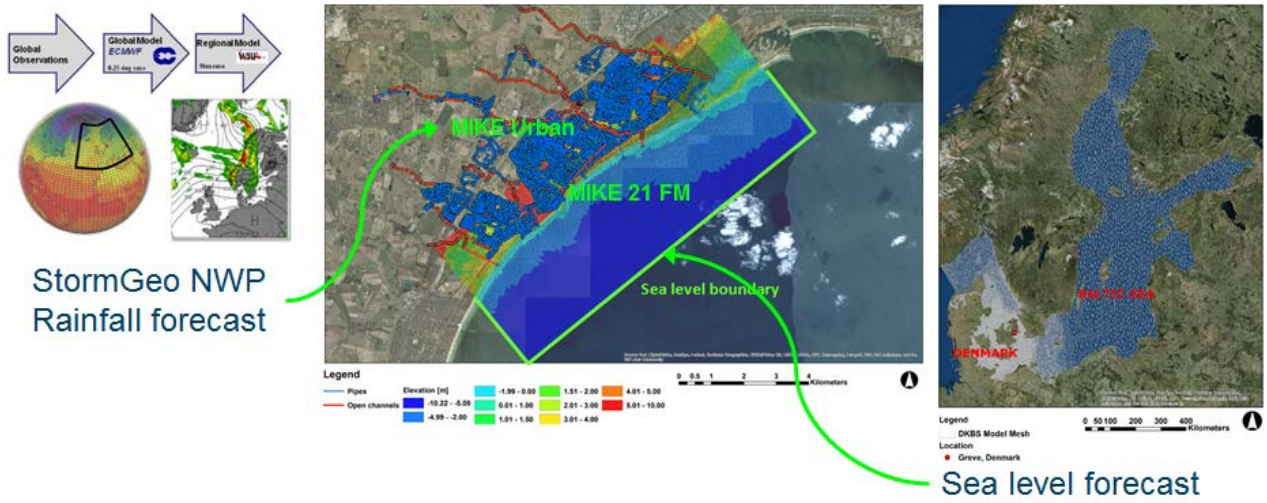


Figure 22: Example of model chain considered in the project.

Ideally, the joint probability distribution of all the important uncertainty sources should be quantified and used in the model chain to derive probability distributions of the model predictions of interest. However, such an approach is not feasible due to complex, non-linear system dynamics and model interactions. Instead Monte Carlo based techniques have been put forward. In these methods the different uncertainty sources are represented by an ensemble estimate, and the ensemble is propagated through the model chain to produce an ensemble of model predictions from which probability distributions can be estimated. A major shortcoming of the Monte Carlo based approach is the slow convergence to the true probability distribution, thus requiring a large ensemble that is computationally intractable and may be infeasible for real-time applications. More efficient sampling strategies exist, including Latin Hypercube Sampling, importance sampling such as the generalised likelihood uncertainty estimation (GLUE) procedure (Beven and Binley, 1992), and Markov Chain Monte Carlo sampling (e.g. Kuczera and Parent, 1998).

A computational efficient alternative for estimation of model prediction uncertainty is based on statistical post-processing. This method does not rely on probabilistic descriptions of the different uncertainty sources but lumps all uncertainties together by considering only the resulting uncertainty in the model predictions of interest. An uncertainty model is estimated from historical time series of model predictions (forecasts) and corresponding observations. This statistical model is then used as a post-processor where prediction uncertainty is estimated conditioned on the model predictions. Different post-processors have been developed and used in hydrological and meteorological model applications, including the Hydrological Uncertainty Processor (Krzysztofowicz and Kelly, 2000), Bayesian Model Averaging (Raftery et al., 2005), Model Conditional Processor (Todini, 2008), and Quantile Regression (Weerts et al., 2011).

3.5.1 Model prediction uncertainty using statistical post-processing

In this project statistical post-processing was applied using quantile regression (QR) for estimation of forecast uncertainty of the open sea boundary condition of the model chain shown in Figure 22. To our knowledge this is the first time QR is used for coastal water level model predictions. With

respect to the rainfall boundary condition of the model chain in Figure 22, post-processing of rainfall forecasts has previously been analysed by René et al. (2013) and has not been applied in this study.

Below is given a brief description of the QR methodology applied. For more details and applications of QR in hydrological modelling see Weerts et al. (2011), Lopez Lopez et al. (2014) and Muthusamy et al. (2016).

The relationship between an observation and corresponding forecast can be expressed as:

$$s_n = \hat{s}_n + e_n$$

where s_n is the observations, $\hat{s}_n$ is the deterministic forecast, and n is forecast lead time. The error e_n may be estimated by means of a probabilistic error model by conditioning the residual (dependent variable) to the deterministic model forecast (independent variable)

$$e_n = f[\hat{s}_n]$$

Once this relation is found, the error model can be applied without any additional data requirements than the model forecast $\hat{s}_n$. A linear functional form is here assumed, i.e.

$$e_n = a_n \hat{s}_n + \mathrm{b}_n$$

QR determines the quantiles of the model residual conditioned on the model forecast. The regression parameters for a given quantile τ are estimated from

$$min \sum_{i=1}^{N} \rho_\tau \left[e_{n,\mathrm{i}} - (a_{n,\tau} \hat{s}_{n,i} + \mathrm{b}_{n,\tau}) \right]$$

where *N* is the sample size available for estimation, and ρ_τ is a weighting function

$$\rho_\tau = \begin{cases} (\tau - 1) e_{n,i} & if \;\; e_{n,\mathrm{i}} \leq 0 \\ \tau e_{n,\mathrm{i}} & if \;\; e_{n,\mathrm{i}} > 0 \end{cases}$$

Individual QR equations are derived for different forecast lead times using samples of historical forecasts and corresponding observations. The QR equations are then used to estimate forecast quantiles conditioned on the model forecast for different lead times and quantiles, i.e.

$$s_{n,\tau} = \hat{s}_n + a_{n,\tau} \hat{s}_n + \mathrm{b}_{n,\tau}$$

In the application of QR described in Section 4.2 a Normal Quantile Transformation (NQT) is applied. In this case the residual and model prediction variables are transformed using NQT and QR is then applied on the transformed variables. The QR equations based on NQT are back-transformed and used as post-processor.

3.5.2 Uncertainty analysis using Monte Carlo based methods

A typical method to assess the contribution of uncertainty of input variables into models' output is the Global Sensitivity Analysis (GSA). This method quantifies the contribution of variance of input variables into output variance. A typical application of GSA method is presented below. The method is based on Monte Carlo Markov Chain (MCMC) sampling techniques and uses Variance based method for Global Sensitivity Analysis.

In the application, data provided from a high-resolution digital elevation model have been used for estimating uncertainty. In addition, the deterministic algorithm (2D hydraulic code) is looked as black box (Saltelli et al. 2000). The model function relies on arbitrary input variables with known distribution while the output is considered as random variable.

The process of global sensitivity analysis goes through (i) uncertainty parameters definition, (ii) uncertainty propagation and (iii) results variability study (Marrel et al., 2012). For the present analysis, the Sobol method is used (Sobol 1990). The Sobol indices are defined as:

$$S_i = Var[E(Y|X_i)]/Var(Y)$$

Where S_i is the Sobol index of parameter *i*, *E* is the Expectation, *Y* is the output. First-order Sobol indices indicates the contribution to the output variance of the main effect of each input parameters.

Global sensitivity analysis has been applied in 1-D hydraulic modelling studies while for 2D models is under investigation. In addition, the GSA asks for a special protocol and development of adapted tools with very important computational resources available to use.

In the application, the investigation was conducted for uncertainties associated with HR topographic data, used in hydraulic modelling (Abily, et al.2015). The following assumptions are made: measurement errors in high-resolution topographic data and inclusion of data in 2D hydraulic codes.

In our application, a small industrial area was chosen covering 17.5 km^2. In GSA method, the hydraulic parameters of the model are the same for all simulations. The changes concern the digital elevation model (DSM) for each simulation. Regarding DSM characteristics, the given data is classified with horizontal and vertical mean accuracy of 0.2 m. The number of 3D polylines over this area (17.5 km^2) is over 1 200 000. The 3D classes are selected for creation of DSM since they are affecting flow direction. There are 12 classes in the DSM. They include buildings, concrete vertical structures above 2 m (walls), sidewalks, road gutter, etc. The 12 selected classes have been aggregated in three groups: (i) buildings, (ii) "concrete" vertical structures (walls) and (iii) street concrete features.

Tools used for delivering a numerous simulations are Promethée and 2D hydraulic Code (FullSWOF_2D). Promethee is a parametric environment developed by IRSN (http://promethee.irsn.org/doku.php). This environment allows parameterisation of any numerical code. On the other side, the FullSWOF_2D code is open source code (GPL- compatible license CeCILL-V2) available on the website: http://www.univ-orleans.fr/mapmo/soft/FullSWOF/. This code solves SWE using finite volume method over structured computational domain using hydrostatic reconstruction method (Delestre 2010).

The method GSA aims to rank the input parameters effect on output variability. The steps for performing GSA have four steps as presented in Figure 23.

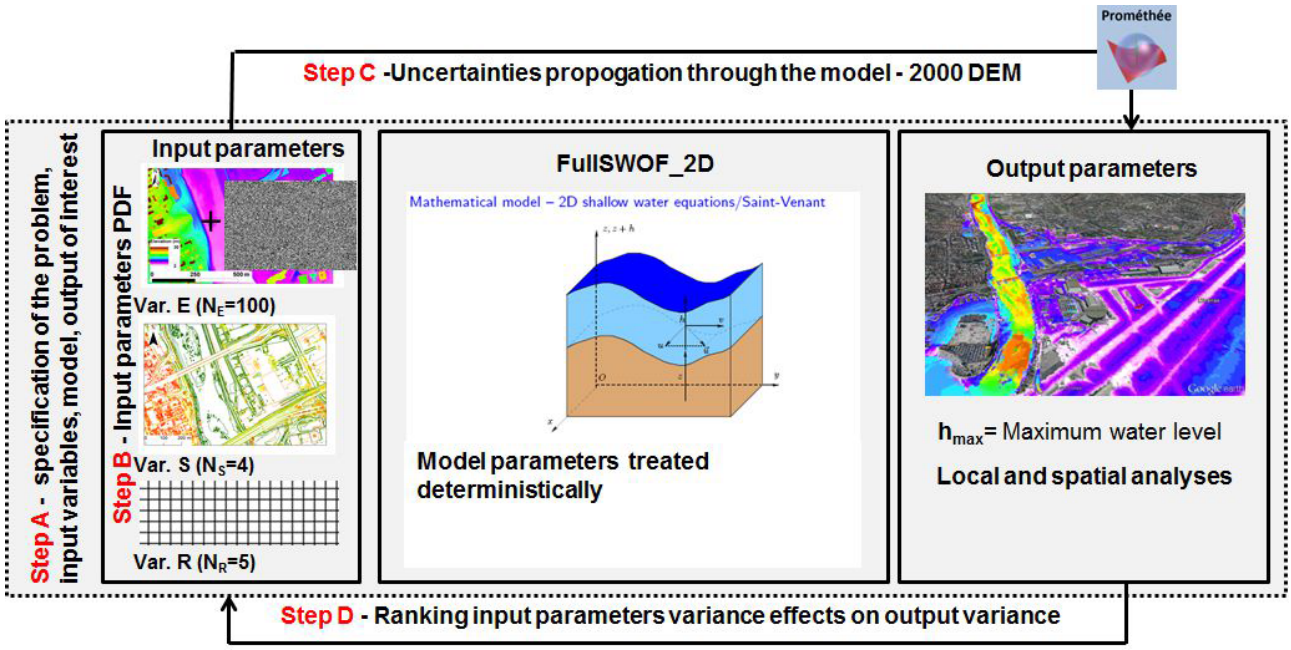


Figure 23: Process of GSA method applied to HR DSMs inclusion in 2D free surface modelling codes. (Abily et al 2015)

The GSA method initiates with problem definition, selecting uncertain paremeters and assigning to them probability density functions. Towards this, one spatial parameter Var.E was used to represent topographical error. The error is randomly introduced in every point of HR DEM (1m), following normal distribution. The two variables Var.S and Var.R were also taken into account. Var.S is ordinal

parameter representing of flow direction affecting the ground structure, where Var.S1 is just DEM, Var.S2 is Var.S1 with buildings included, Var.S3 is VarS2 with walls and Var.S4 is Var.S3 with concreate street features. Var.R represents decisions made by the modeler with discrete values from 1 to 5.

In total 2000 simulations were performed and used in GSA. Uncertainty propagation (step C) is done using Monte Carlo approach. In Step D, Sobol index calculation was conducted via post treatment scripts, written in R language.

For the analysis 1500 simulations were used. The results are presented in Figure 24 below. Averaged results from 1500 simulations for maximum water depths are presented in *Figure* 24a. The output of local analysis is shown in *Figure* 24b,c,d. The convergence of mean is presented in *Figure* 24b. The confidence interval of 95% is obtained from 900 samples. *Figure* 24c illustrates the importance of variance of maximum water levels.

Global sensitivity analysis used for computation of Sobel index is presented in *Figure* 24d. The confidence interval has been estimated by running 1000 random simulations retrieved from the database. The most influential parameters are Var.S and Var.R.

Finally the Sobol index maps are presented in Figure 25. As mentioned, the two parameters are most influential (Var.s and Var.R). The dense urbanized parts of DEM introduces high variability (for the donain of Var.S) in model output. The analysis also shows that for the low output variance Var.r Sobol index is highest. As a conclusion the Var-R Sobol index value is important at the edge of flood extent, close by river ben in not densely urbanized areas.

The index maps presented in Figure 25 represent the relative weight of each uncertain parameter on variability of calculated overland flow. The results highlight uncertainty introduced by modeller choices.

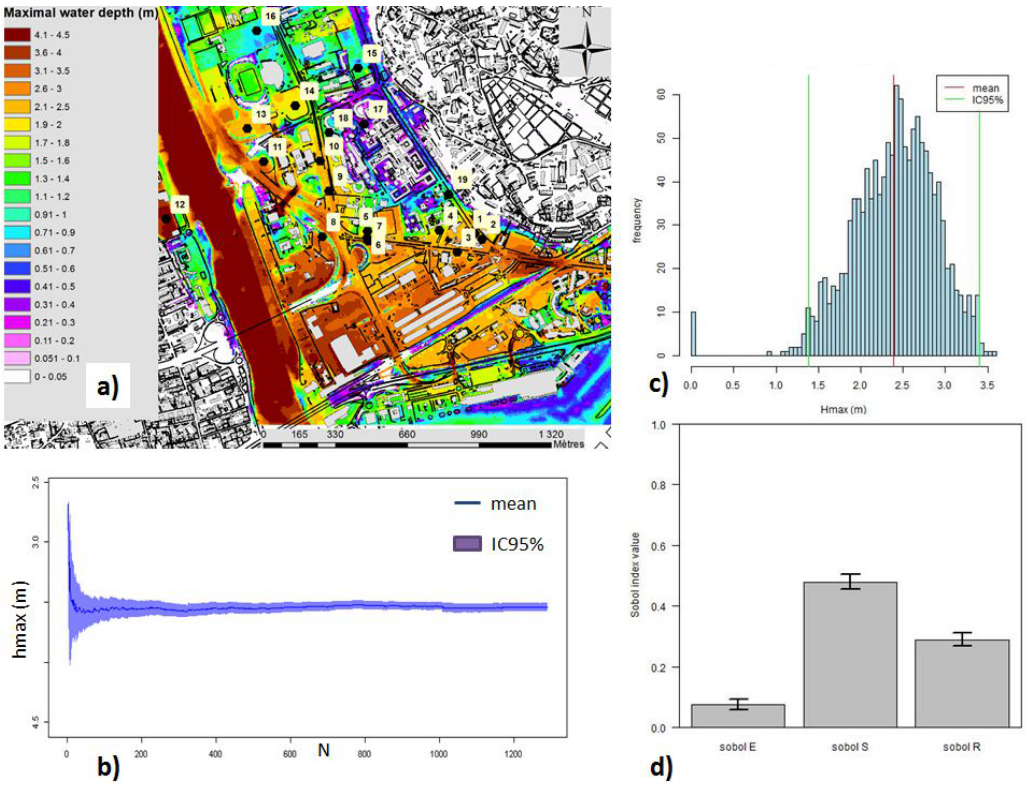


Figure 24: Average of the 1,500 maximal water depth flood maps computed over a selected sub-domain (a); local results at point 1 illustrating convergence check (b), results distribution (c) and Sobol index (d)

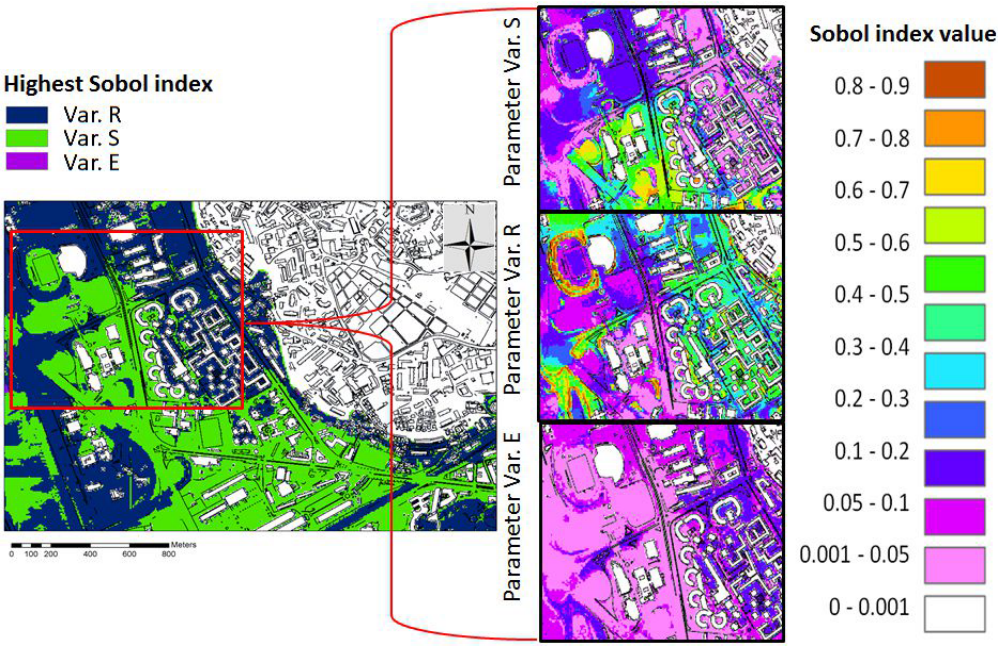


Figure 25: Sobol index maps ranking the 3 input parameters variance proportionated to the output (maximal water depth) variance.

4 Application of the methodologies

4.1 Marbella case study

Marbella is a city that belongs to the province of Malaga and it is part of the “Costa del Sol”. It is located in the south of Spain, in the region called Andalucía (Figure 26). The city covers an area of 117 km^2 and (according to the Spanish Statistical Office in 2015) 138,679 habitants live there with a density of almost 1,200 hab/km^2.

The city presents a typically Mediterranean topography with high slopes on the upper part of the basin and very flat areas close to the sea.

Three main river streams cross the surrounding area: Guadalmina (28 km long, 68 km^2 catchment), Guadaiza (22 km long, 45 km^2 catchment) and Rio Verde (35 km long, 150 km^2 catchment).

As a Mediterranean city, it benefits from a wet and warm winter and a dry and hot summer, with a mean annual temperature of 18 °C.

The mean annual rainfall is 625 mm, most of which is observed during the autumn and winter. While the duration of these events is not very long the rainfall intensity is often very high.

Many beaches and recreational ports are located along the 44 km coastline of Marbella city and its surroundings. The main economic activity is tourism targeting wealthy tourists and therefore the local infrastructure is characterized by high economic value. This fact is demonstrated by the presence of high-class resorts, recreational ports, golf clubs and many other luxurious facilities.

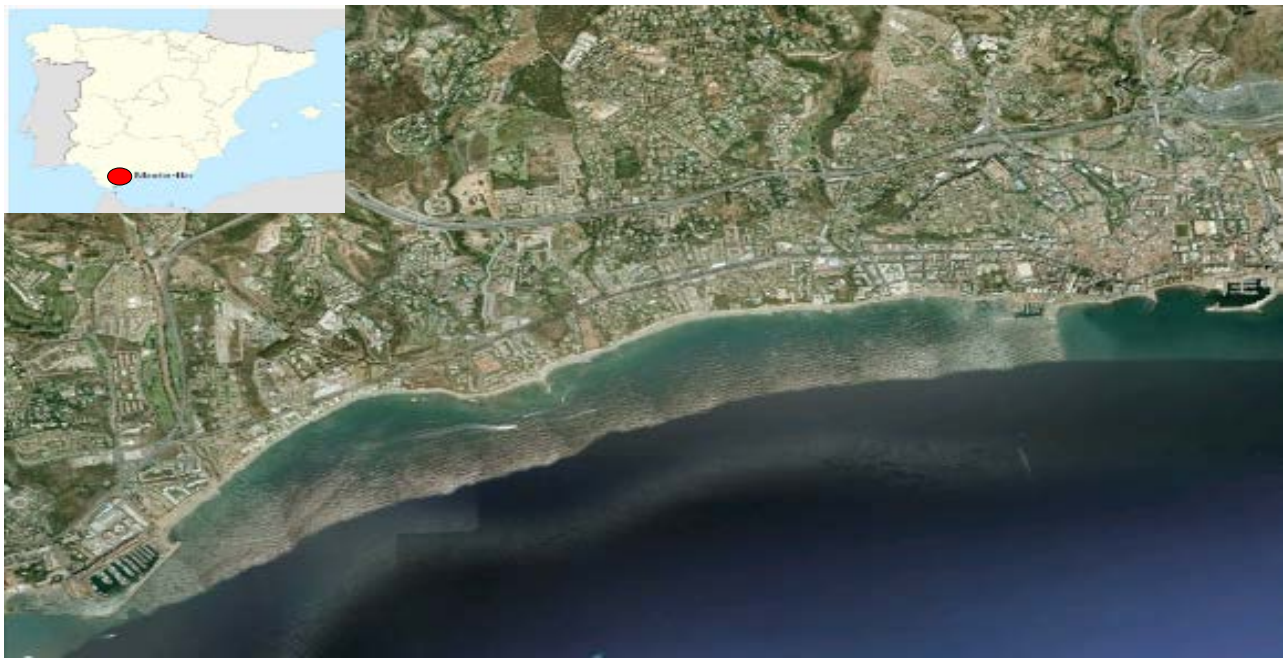


Figure 26: Map showing the location of Marbella.

The municipality is highly vulnerable to flooding from extreme rainfalls. Therefore, the development of a Flood forecasting system has highest priority for the city of Marbella.

Flood forecasting system based on radar nowcasting

An Early Warning System [EWS] for flood forecasting has been implemented in Marbella in order to issue flood warnings. It uses radar precipitation estimates to calculate precipitation forecasts for the

following 2h. Next to the forecasted rainfall intensities the system calculates also temporal accumulations for different time supports (30 min, 1h, etc.). The system is capable to adjust its predictions automatically based on incoming precipitation characteristics (direction and velocity) in order to issue warnings (sent by SMS and e-mail) and in order to reduce false alarms. Details of the radar-based EWS can be found in Llort et al. (2014). Figure 27 and Figure 28 show the radar-based EWS (24h accumulated precipitation and rainfall intensity respectively). These images show one of the most significant events observed in 2016, on the 25th November (used also for calibration purposes in the hydraulic model), where warnings based on forecasted precipitation were issued from the developed EWS.

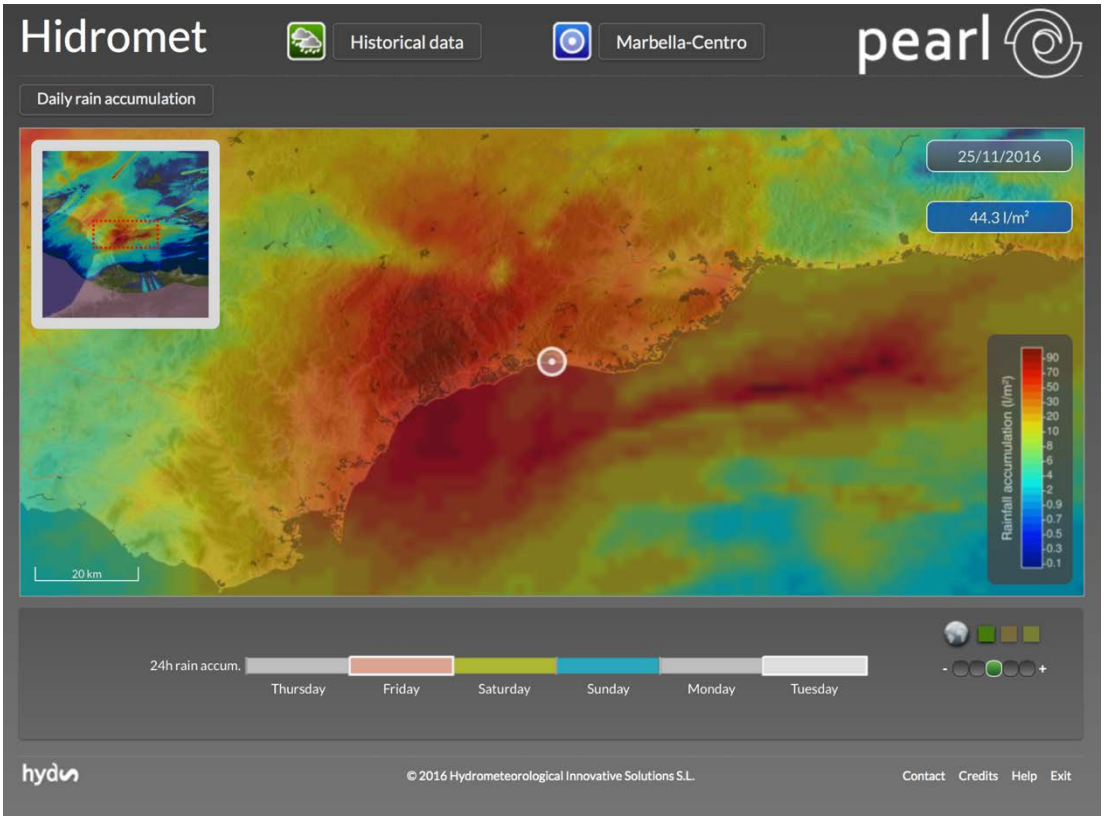


Figure 27: Daily rain accumulation (observed by the radar network) for the 25th of November 2016.

Figure 28: Rainfall intensity registered by the radar network on the 25th of November 2016.

Precipitation information from both rain gauges and radar (observed and forecasted) are used to feed a hydraulic model to issue warnings over specific elements of the sewer network in real time. In order to achieve this advanced techniques such as GPUs have been used for speeding up the calculation. Details of the radar-EWS as well as its link with the hydraulic model will be available in Deliverable 4.5 and are provided by Russo et al. (2017).

4.1.1 Methodology

The experimental module of probabilistic radar nowcasting has been tested in Marbella using the aforementioned EWS. The methodology used is presented in section 3.4 applied by the AEMET (Agencia Estatal de Meteorología). The steps performed are described in the following subsections.

4.1.1.1 Data acquisition

In order to validate the developed probabilistic nowcasting prototype, data from the AEMET radar network is acquired in real time. A conversion of the information to suitable input format is performed.

4.1.1.2 Radar data quality control and calibration

Due to different errors affecting radar precipitation estimates (see for reference Zawadzki 1984), radar data undergo a quality control process aiming at improving the quality of the measurements.

The first applied quality control algorithms deal with the correction of non-meteorological echoes and the underestimation of the precipitation due to beam blockages (interaction of the radar measurement process with the topography and other elements). The schemas in Figure 29 illustrate those errors.

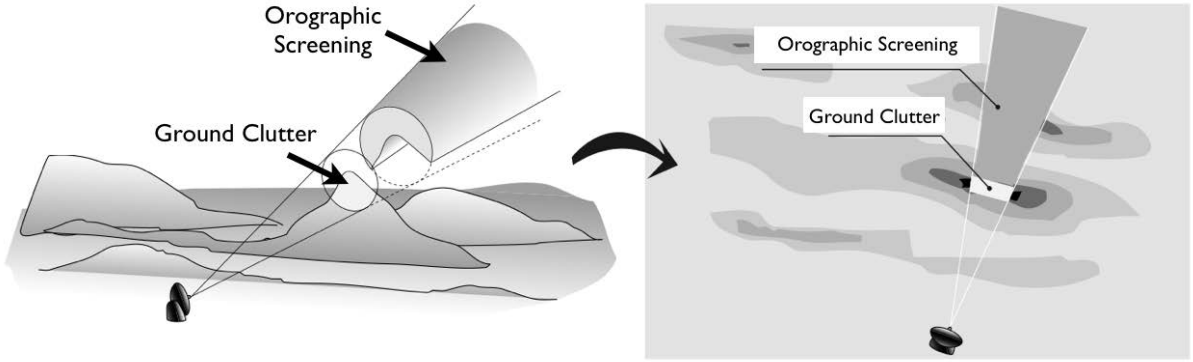


Figure 29: Schemas explaining the occurrence of errors due to the intersection of the radar beam with the topography (ground echoes and orographic screening). (From Sánchez-Diezma, 2001).

In this application, both the areas affected by orographic screening (blockage of the radar beam by the orography) and the areas affected by ground echoes are corrected. Figure 30 shows the areas affected by orographic screening and the areas affected by ground clutter as calculated using the AEMET radar coverage. Those areas are not considered in the radar Nowcasting.

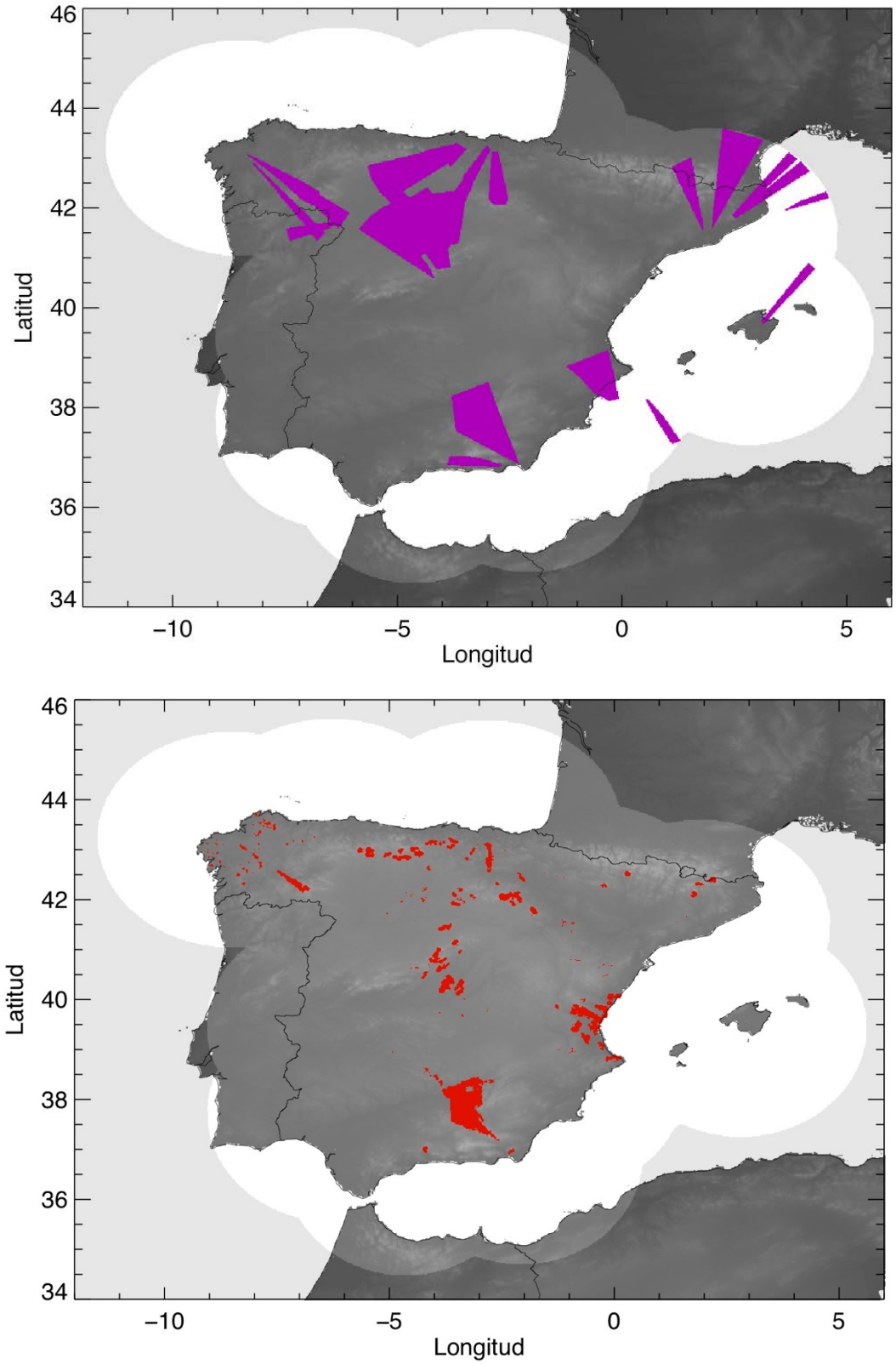


Figure 30: AEMET radar network coverage and areas affected by orographic screening (in purple, top figure) and areas affected by ground clutter (in red, bottom figure).

In a second step, a statistical climatological calibration is calculated for each radar, in order to correct the under/over-estimation due to equipment calibration failure. For this, long-term radar accumulations are calculated and compared to rain gauge estimates in order to derive the differences between them and estimate a linear correction factor to be applied to the radar estimates.

Meteorological radars do not measure directly the rain rate but the “Reflectivity”, i.e. the amount of energy that is reflected by the droplets. Rainfall intensity and reflectivity are related through the Drop Size Distribution [DSD], which is unknown.

Conversion from reflectivity to instant rainfall is done using a climatological Z-R relationship. In this case, a Z-R relationship derived from the DSD studies, introduced by Marshall and Palmer (1948), is used with $Z=200 \cdot R^{1.6}$. More detailed technical information about the algorithms that are applied can be found in Sánchez-Diezma (2001).

4.1.1.3 Calculation of radar Nowcasting

By means of cross correlation techniques, the precipitation movement field is calculated from the last radar observations. This movement field is used to extrapolate the last observations in order to provide forecasted precipitation fields for the next hours (see Berenguer et al. 2005 and Berenguer et al. 2011 for a complete description of the technique).

Forecasting of precipitation with radar Nowcasting techniques cannot be used for large *leadtimes,* as the forecasting ability decreases quickly with time. However, this technique provides the best possible precipitation forecasting for the next 1-6 hours, and is capable to adjust the forecasts in real time keeping them aligned with the last observations.

Figure 31 shows an example of the Nowcasting and the accumulation algorithms. The illustrations show 1h accumulations for the event of 21st-22nd of April 2011 at different times for the full radar coverage of the AEMET radar network. The red arrows represent the precipitation motion field calculated in each time step and used to calculate the radar Nowcasting.

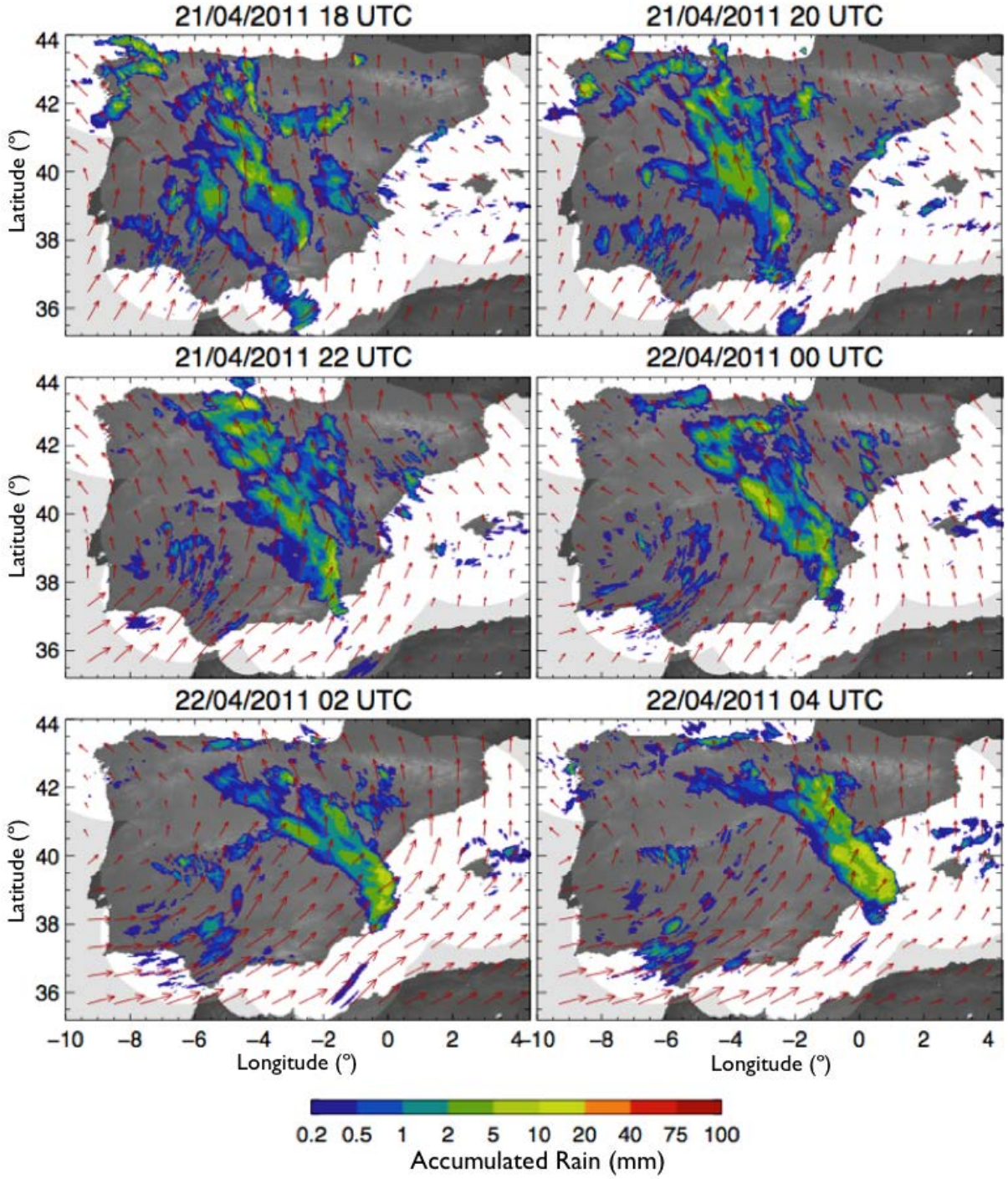


Figure 31: Example of results of the Nowcasting and the accumulation algorithms.

4.1.1.4 Generation of rain accumulation fields

In order to calculate rainfall accumulation from the radar instantaneous precipitation estimates (both observed and forecasted), the movement of the precipitation field is taken into account. A direct sum of the instantaneous precipitation fields would provide non-realistic accumulations because the precipitation between radar scans would not be taken into account. Sánchez-Diezma (2001) provides several examples of this problem, an example of which is shown in Figure 32. With the applied technique, the movement field between two subsequent observations is estimated, and then

used to calculate all the intermediate states between observations (minute by minute rainfall fields). The final accumulation field is then computed taking into account all the intermediate information.

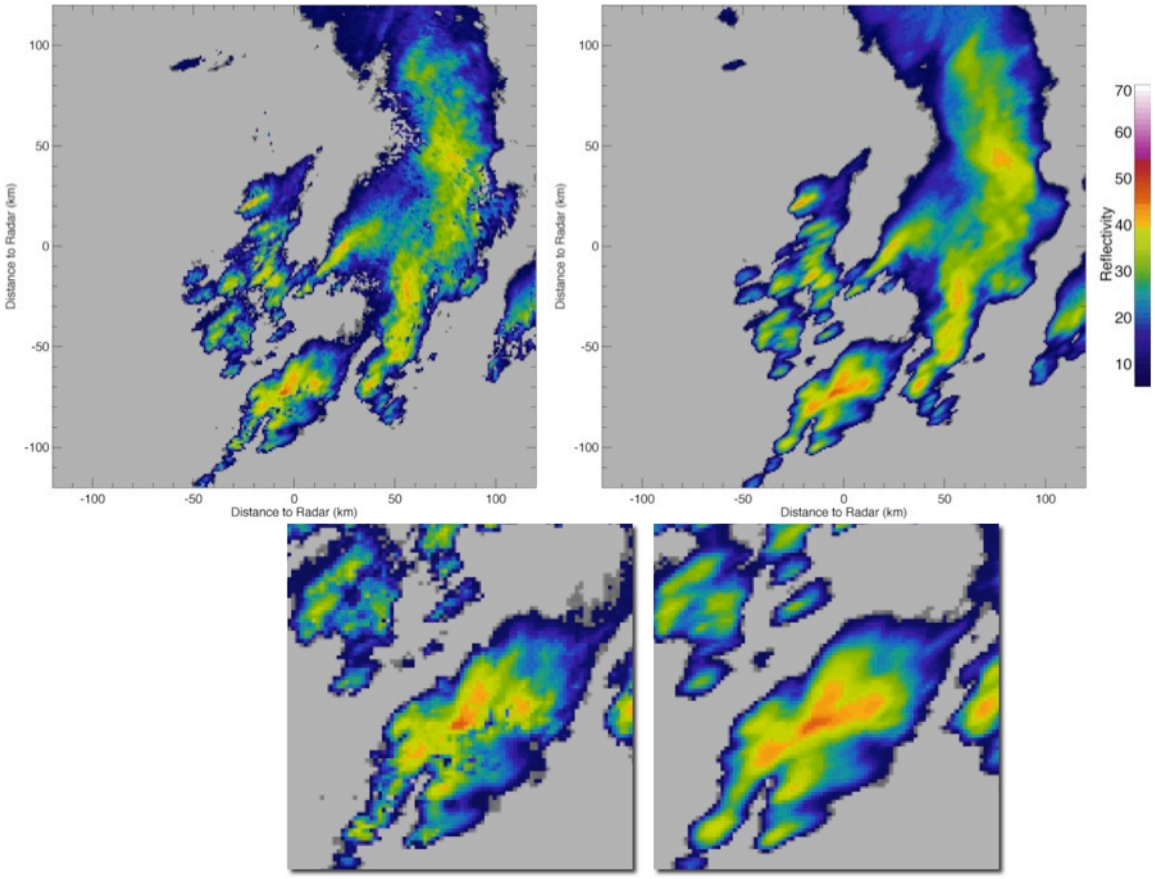


Figure 32: Estimation of rainfall accumulation using different techniques: taking into account the motion field to calculate the rain accumulations (right) and without taking it into account (left). The effect is clearer in the areas affected by convective cells (bottom: zoom of the top image).

4.1.2 Results

No formal evaluation of the results has been done in the Marbella implementation. An extensive comparison of probabilistic nowcasting techniques can be found in Buil (2017), showing that the Local Lagrangian technique outperforms other options.

The process following the described methodology is running in real time in the Marbella Case study and the corresponding EWS is publicly accessible from the PEARL Project page. Details of the EWS system are provided in PEARL Deliverable 4.5 and have been presented in (Russo et al., 2017).

4.2 Greve case study

The coast of Greve Municipality in Eastern Denmark is one of the case study areas in the PEARL Project. The case area is used as a laboratory for implementing and evaluating methods for coastal flood risk management, which are identified and developed during the project.

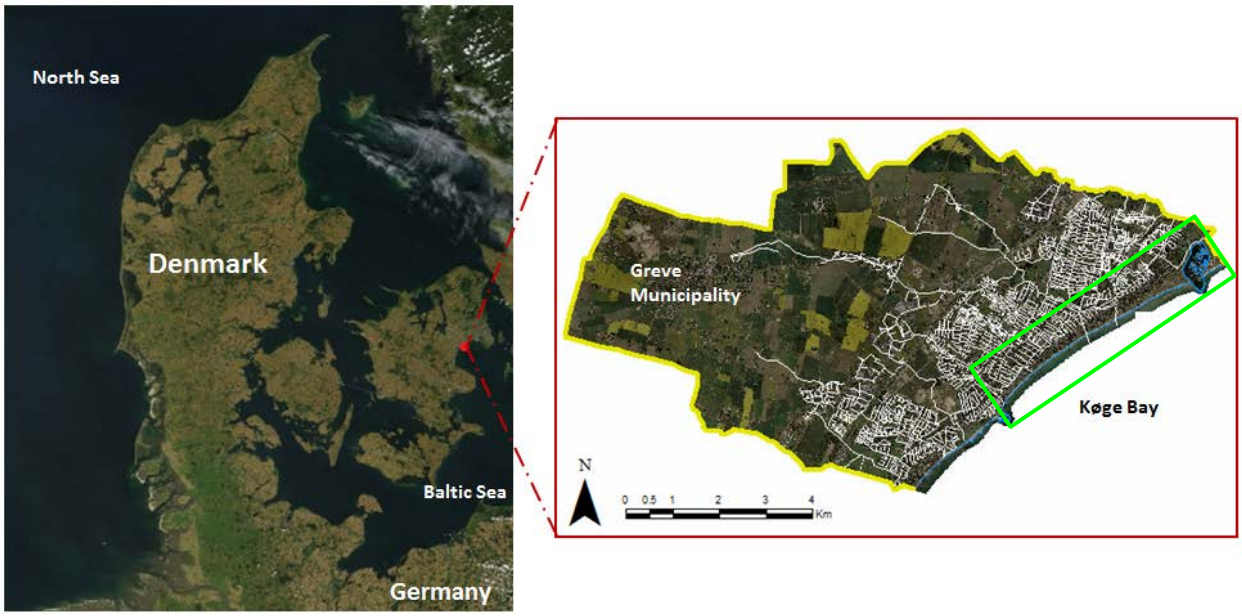


Figure 33: Map showing the location of Greve Municipality in Eastern Denmark, with the extents of the case area outlined in green.

Greve is located about 20 km southwest of the Danish capital, Copenhagen. The municipality has a total area of around 60 km^2, with 9 km of coast to the southeast bordering the Baltic Sea at Køge Bay (Figure 33). Its coast is the most densely built-up part of the Municipality, and comprises the case study domain. The model area is characterized by relatively flat terrain with elevations varying between 2 and 6 m above MSL.

The municipality is vulnerable to flooding from extreme rainfall and sea levels. In July 2007, a series of rain events with an estimated return period of 500 years caused severe and destructive flooding in Greve (Greve Kommune, 2007). Storm surges and general sea level rise also pose a threat to the area due to its location in Køge Bay (Vestergaard, 2011). Køge Bay is 1 of 10 flood risk areas identified in Denmark in the implementation of the EU Flood Directive (Naturstyrelsen og Kystdirektoratet, 2011).

Flood forecasting system

An online real-time coastal flood forecasting system has been developed for Greve under the PEARL Project. The system employs a 1D-2D hydrodynamic flood model, which uses NWP rainfall forecasts and 3D hydrodynamic model water level forecasts as boundary conditions, to forecast flooding in the area (Figure 34). The 1D-2D model was built using MIKE Flood FM, a modelling system integrating one-dimensional and two-dimensional models into one dynamically coupled model (DHI, 2014a). It comprises a 1D model of the drainage system coupled to a 2D surface flow mesh model of the coastal area. Representation of the drainage network in the flood model ensures simulation of pluvial and fluvial flooding, and consideration of the drainage network's influence in conveying flooding from the sea further inland.

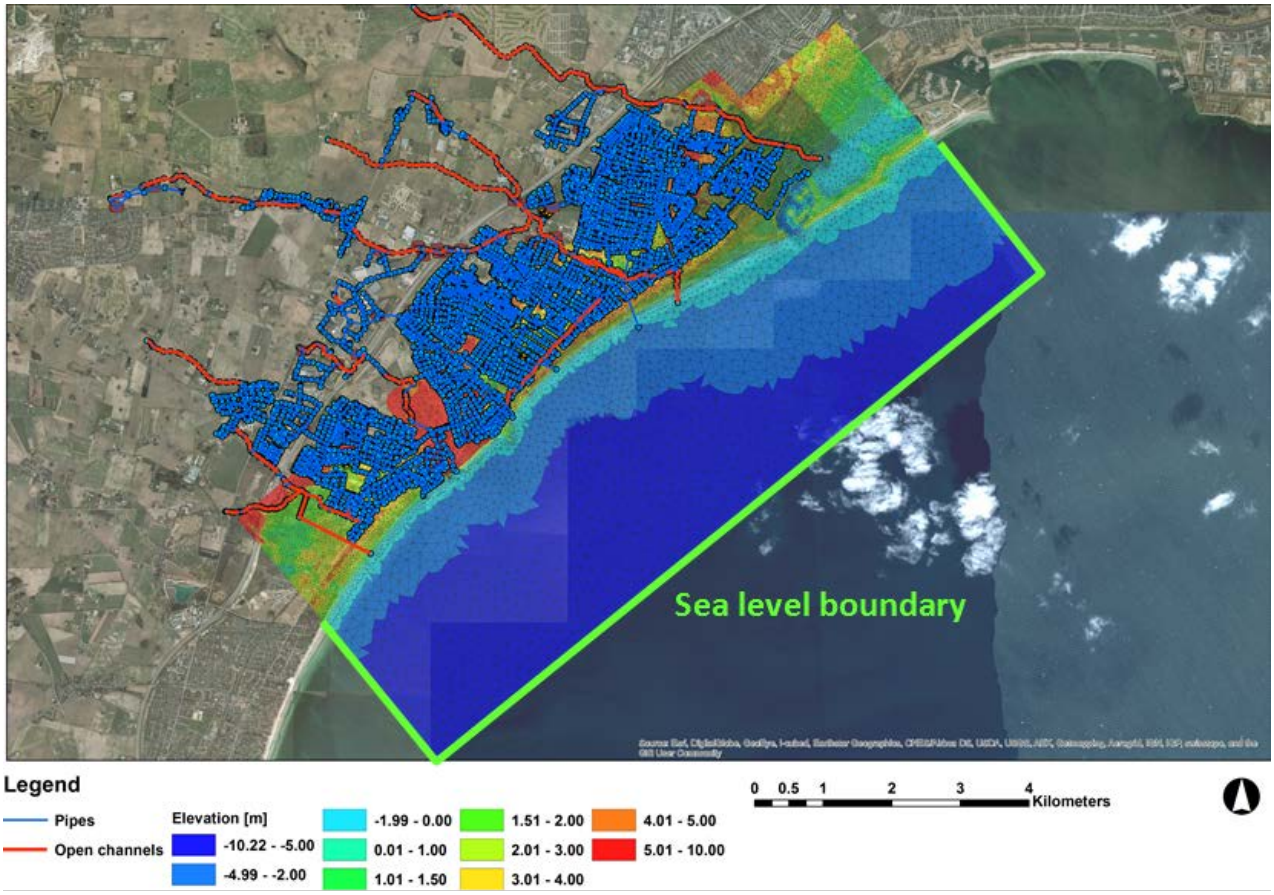


Figure 34: A map showing the 1D-2D coastal flood forecast model for Greve. A 1D model of the drainage system comprising of streams (red lines) and stormwater sewers (blue lines and points) is linked to a 2D mesh model of the coastal area (coloured areas). The triangular elements of the terrain mesh and open boundaries of the 2D model are also shown.

The 1D network model is relatively comprehensive and detailed, covering most of Greve. The model employs a conceptual runoff model to calculate rainfall runoff in the area, and a 1D fully hydrodynamic pipe flow model simulating water levels and flows in the drainage system. Forecasted rainfall data obtained from StormGeo (SormGeo, n.d.) are used to calculate runoff, which is then applied as input to the network model. Sea water level forecasts are applied as boundary conditions to the 2D MIKE 21 FM model (DHI, 2014b) at its open boundaries to the sea. The 2D and 1D models are linked at points where flow exchange between the two systems can occur, such as at drainage outlets, stormwater inlets, unsealed manholes, and open channels.

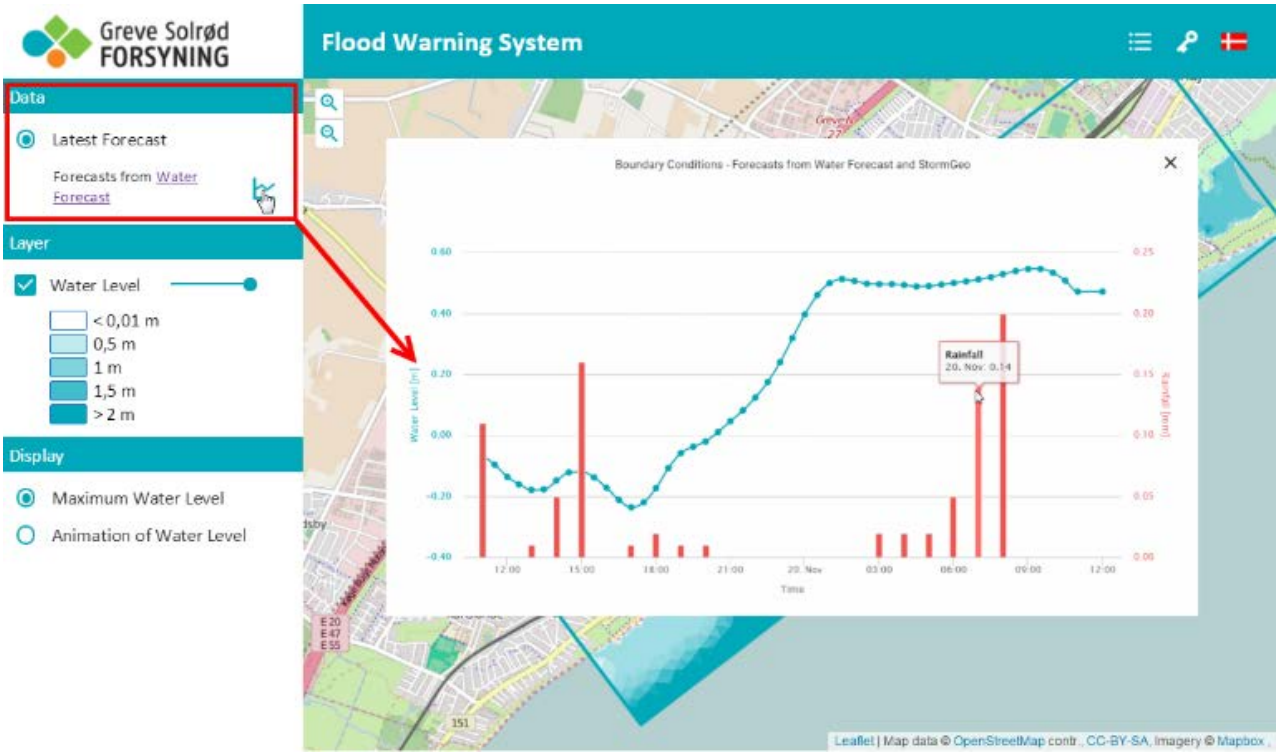


Figure 35: A deterministic flood forecast is issued by the system based on forecasted sea water level (blue lines) and rainfall (red bars) time series in the study area (http://greve.dhigroup.com/).

The system provides deterministic forecasts of spatially-distributed water levels and flows in the model domain. A forecast is issued based on rainfall and sea level forecasts used as boundary conditions (Figure 35). One-day flood forecasts are issued by the system every 4 hours, and computation results are shown on a website, wherein static maximum water depths and time-varying water depth animations are plotted on a map. The deterministic forecast comprises a single estimate of future conditions, which do not include various sources of modelling uncertainty.

Probabilistic flood forecasting

Probabilistic forecasting provides information on the likelihood, in terms of probabilities of occurrence, and may be used for uncertainty consideration in flood forecasting. This method gives model prediction uncertainties, which communicates forecast uncertainties by providing a confidence interval on the forecast. With this type of information, better decisions may be taken regarding issuance of flood warning and/or evacuation. Probabilistic forecasts allow decision-makers to set their own levels of uncertainty for launching emergency response options (Verkade et al., 2013).

In this study, also presented in Diez (2016), improvement of the real-time coastal flood forecasting system in Greve was investigated using probabilistic modelling. A probabilistic flood model, which can quantify forecast uncertainties, was implemented and evaluated for the case area. The technique applied provides information on total prediction uncertainty rather than on each individual uncertainty source, such as input, model parameter, or model structure uncertainties.

4.2.1 Methodology

The probabilistic forecast methodology applied in the case study is based on the QR method described in Section 3.5.1. QR equations are estimated from available historical records of water level observations at a harbour in the area (Mosede harbour) and corresponding water level

forecasts. However, before QR is applied, the model forecasts are bias corrected and updated with observations at time of forecast, and the updated model forecasts are used in QR. The derived QR equations are extrapolated from the harbour to the open sea boundary (see Figure 34), which are used to force the 2D hydrodynamic model. Finally, quantile forecasts of water level at the open sea boundary are applied in the model chain to produce probabilistic forecasts. In extreme storm surge situations where flooding occurs the forecast system then provides probabilistic forecasts of flooding.

Bias correction and model updating

Bias correction involves correcting the difference (i.e. bias) between forecasts and measurements at the time of forecast, and due to auto-correlation in the bias, the entire model forecast is updated. In this study, the deterministic model forecast was bias-corrected and updated using the following procedure:

1. Analyse the correlation between the initial residual (i.e. observation – forecast at time of forecast) and the residuals for the following lead times. A linear regression equation is estimated for each lead time *n*
2. Plot the correlation coefficient obtained in (1) for every lead time. Identify the data pattern, and fit a function to the data to obtain a smoother regression.
3. Correct the deterministic model forecast $\hat{s}_n$ for different lead times using:

$$\hat{s}_{corrected_n} = \hat{s}_n + f(e_0)$$

where $f(e_0)$ is the smoothed function that estimates the residual at lead time *n* based on the residual at time of forecast.

Error model extrapolation to model boundaries

In the study, the error model is derived based on data for one of the two harbours in the area where forecast and observed data are available. Data availability was largely controlled by real measurements, since forecasts could be obtained throughout the domain with the hydrodynamic model.

Extrapolation of the error model at the harbour to the open sea boundaries is necessary in order to use the probabilistic forecasts as forcing data for the hydrodynamic model. This is done by analysing the functional relationship between the simulated water level values at the harbour and the model boundaries.

For the Greve system, there was a linear relationship between simulated water levels at the harbour and the model sea boundaries (see Figure 43). The error model at a boundary is given by:

$$e_{n,\tau}^{bound} = \hat{e}(n,\tau)a_n + b_n$$

where $e_{n,\tau}^{bound}$ is the adjusted error model at each boundary, $\hat{e}(n,\tau)$ is the error model at the harbour, and a_n and b_n are the regression parameters obtained from a linear regression between water levels at the harbour and at the model boundaries. This relation is calculated for every model boundary for each lead time *n* and quantile of interest τ.

Probabilistic flood mapping

The method for obtaining probabilistic flood maps for the case study involves the application of the estimated QR model extrapolated to the model boundaries, used as forcing data for the 1D-2D flood model, from which a set of (probabilistic) maps of water levels are obtained (Figure 36).

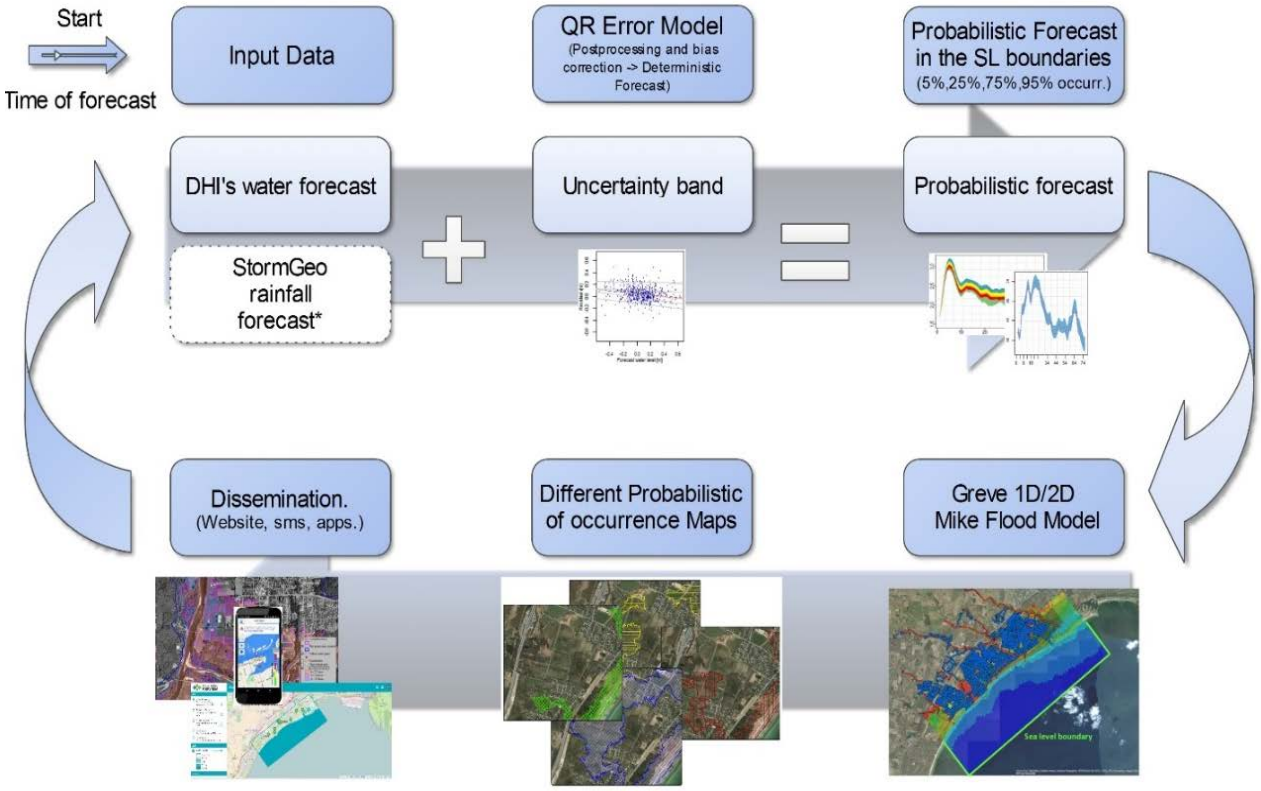


Figure 36: Illustration of the process for obtaining probabilistic forecasts and flood maps in the Greve Case Study.

The analysis starts when forecasts of sea water levels at the open sea boundaries are collected from DHI's Water Forecast system (deterministic forecast). Then, the calibrated QR Model is extrapolated to the boundaries and added to the deterministic forecast. Probabilistic water level forecasts for different probabilities of exceedance (or quantiles) are obtained. Then, water level time series for different probabilities of occurrence/quantiles (usually 5%, 25%, 75% and 95%) are created for a lead time of interest in a file format useable by the flood model as sea level boundaries. Corresponding flood maps for maximum water depths are obtained from flood model calculations. The process is repeated for next time of forecast.

Data

The statistical post-processing method used for developing the forecast error model in the case study requires a long time series of (water level) forecasts and observations in the model domain. However, shortage in available data, as well as lack of extreme event data within the available dataset, were recognized as limitations to the study.

Water level observations at two harbour stations were available from Greve Municipality (Figure 37). Although measurements were available at two harbour stations, statistical analysis and development of the error model were only performed for one location (Mosede harbour), as measured and forecasted levels at both harbours were found to be very similar.

Water level forecasts from the Greve flood model were re-generated by re-running the flood model (i.e. hindcasting) using past saved forecasted sea level boundary conditions obtained from the DHI Water Forecast system. A summary of data that were available for the development of the error model in Greve is shown in Table 4.

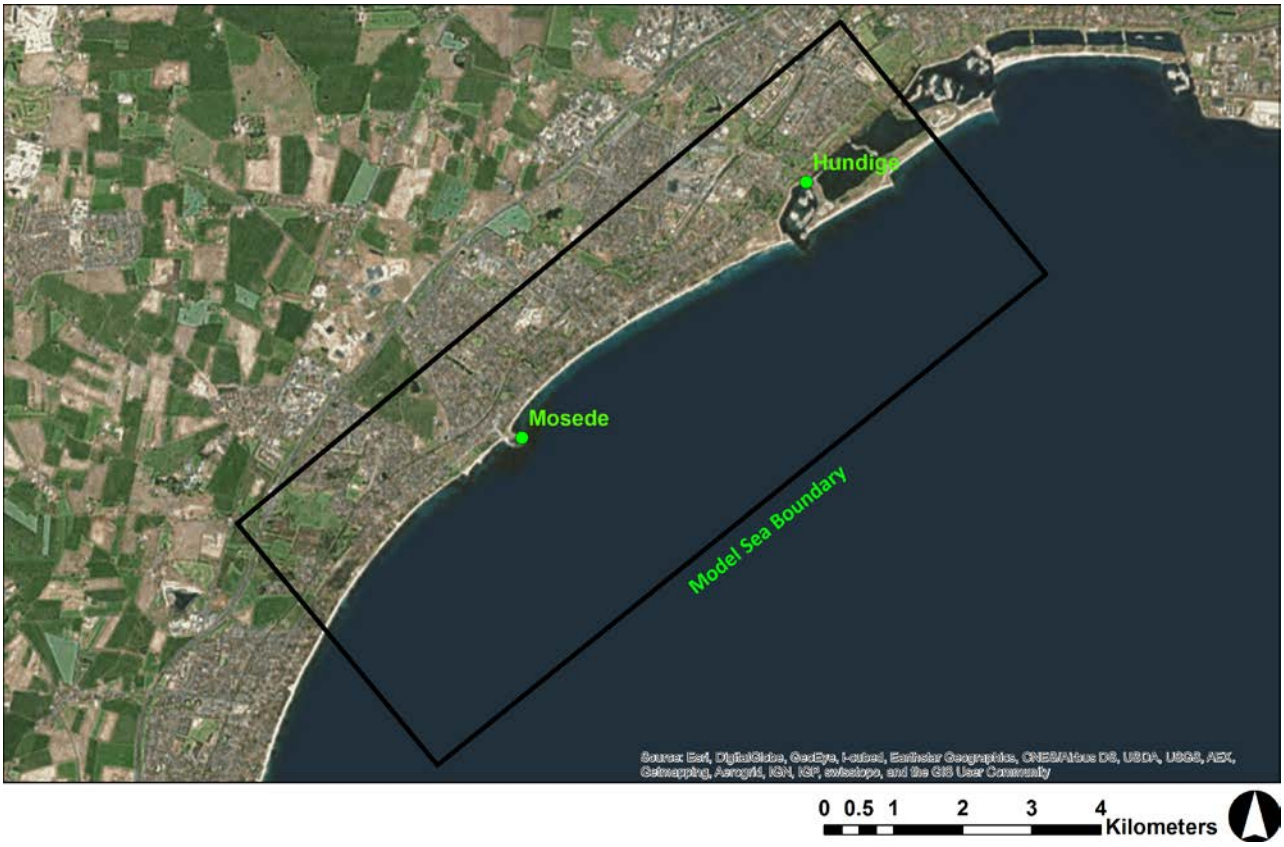


Figure 37: The map shows water level measurement stations at Mosede and Hundige harbour in Greve, and the flood model sea boundary. The coastal flood model domain is outlined in black.

Table 4: Summary of available data for statistical post-processing in the Greve case study area.

Location	Item	Source	Frequency [min]	Available Period
Model sea boundary	Forecasted boundary water level	DHI Water Forecast (DKBS model)	30	24 Apr 2015 – 4 Apr 2016
Mosede Harbour	Measured water level	Greve Municipality	5	25 Jul 2009 – 14 Apr 2016
Hundige Harbour	Measured water level	Greve Municipality	5	19 Feb 2015 – 14 Apr 2016

Given the available dataset, the analysis was conducted using only about 1 year (24 Apr 2015 - 4 Apr 2016) of data, which was further split into training/calibration and verification purposes (Figure 38). This amounted to a total of 559 samples for training and 105 samples for verification.

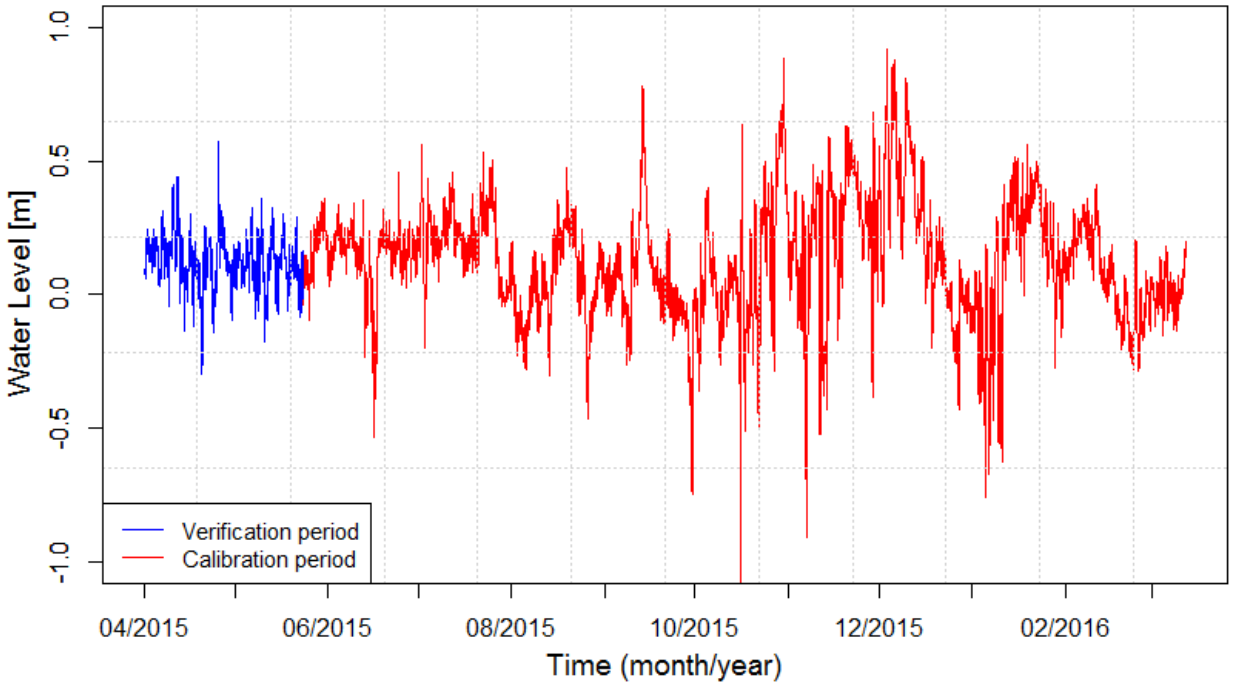


Figure 38: Plot of observed water levels at Mosede Harbour, showing the division of the available dataset into training/calibration (red) and verification (blue) periods.

4.2.2 Results

Results that are presented in this section are based on Diez (2016).

Quantile regression

The QR method based on NQT was applied to the data at Mosede harbour. To evaluate the effect of the bias correction and model updating QR was applied to both the raw data and the bias corrected/updated data. Figure 39 shows the probabilistic water level forecasts resulting from the QR error model without bias correction for four lead times during the verification period (18 April 2015 - 16 June 2015). The deterministic forecast (red line) tends to underestimate the observations. The uncertainty increases with lead time, although not as much as expected, which is due to the relatively large discrepancies between forecasts and observations at the beginning of the forecast.

Figure 40 shows the probabilistic water level forecasts with bias correction during the verification period. The uncertainty band at the smaller lead times has become narrower as a result of the bias correction. Figure 41 shows the RMSE of the water level forecast for different lead times of, respectively, the deterministic forecast and the median quantile forecast based on the QR model with bias correction. The QR bias correction model improves the forecast for all lead, with largest improvements for lead times up to about 6 hours.

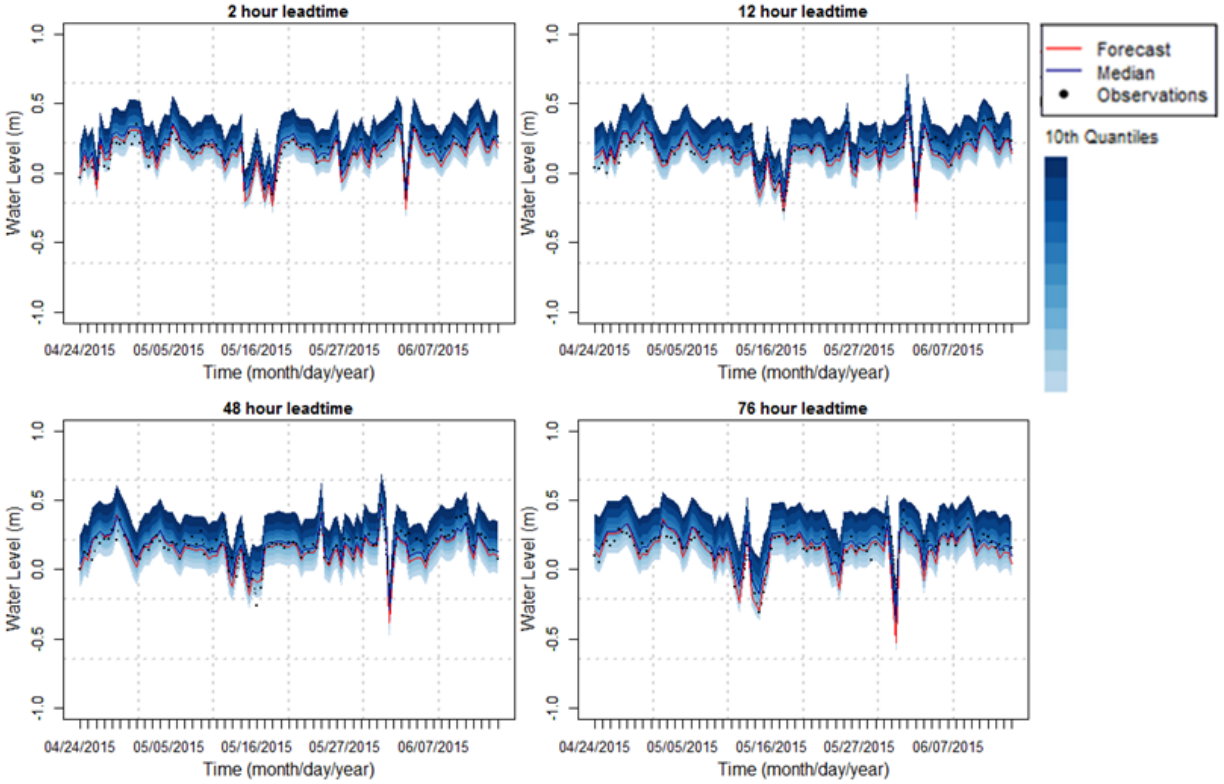


Figure 39: Probabilistic water level forecasts at Mosede harbour without bias correction. The blue shaded areas represent the 90% uncertainty interval, the red lines are the deterministic forecasts, and the black dots, the observations.

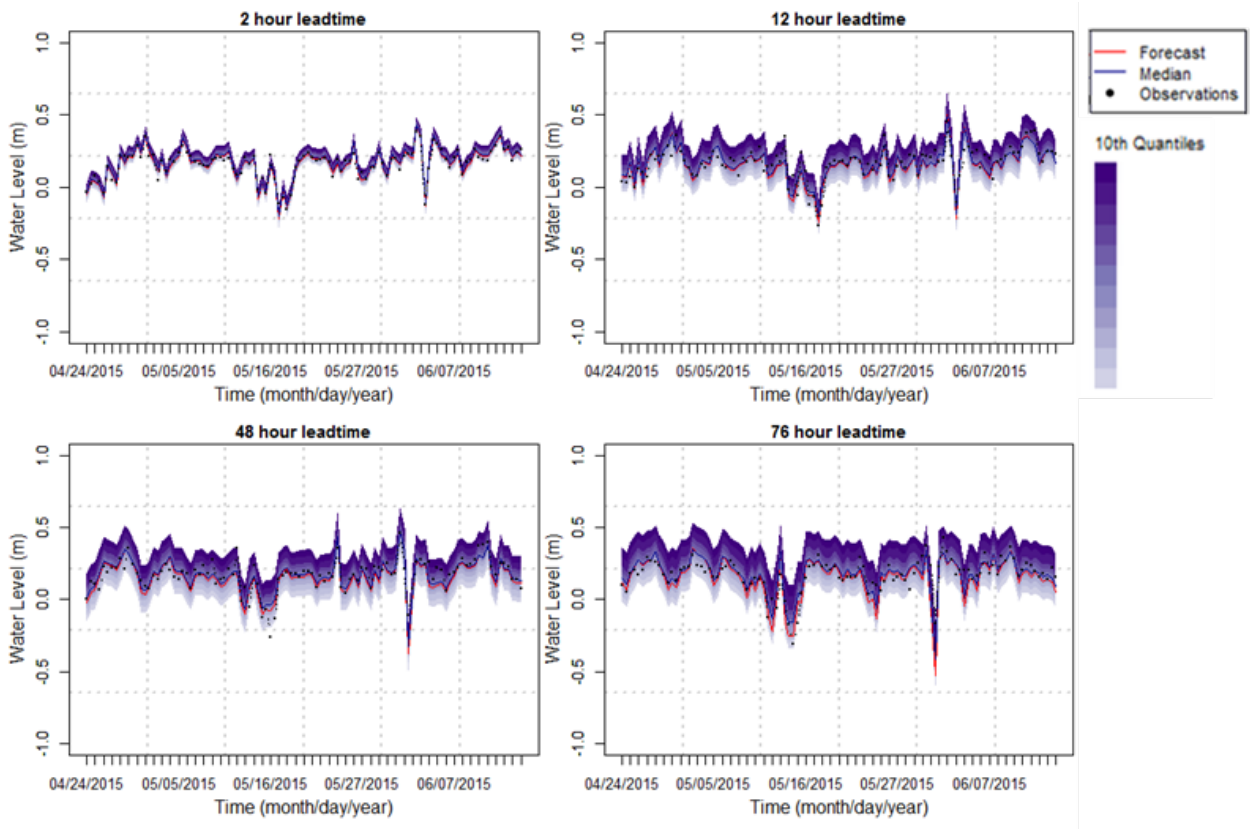


Figure 40: Probabilistic water level forecasts at Mosede harbour with bias correction. The purple shaded areas represent the 90% uncertainty interval, the red lines are the deterministic forecasts, and the black dots, the observations.

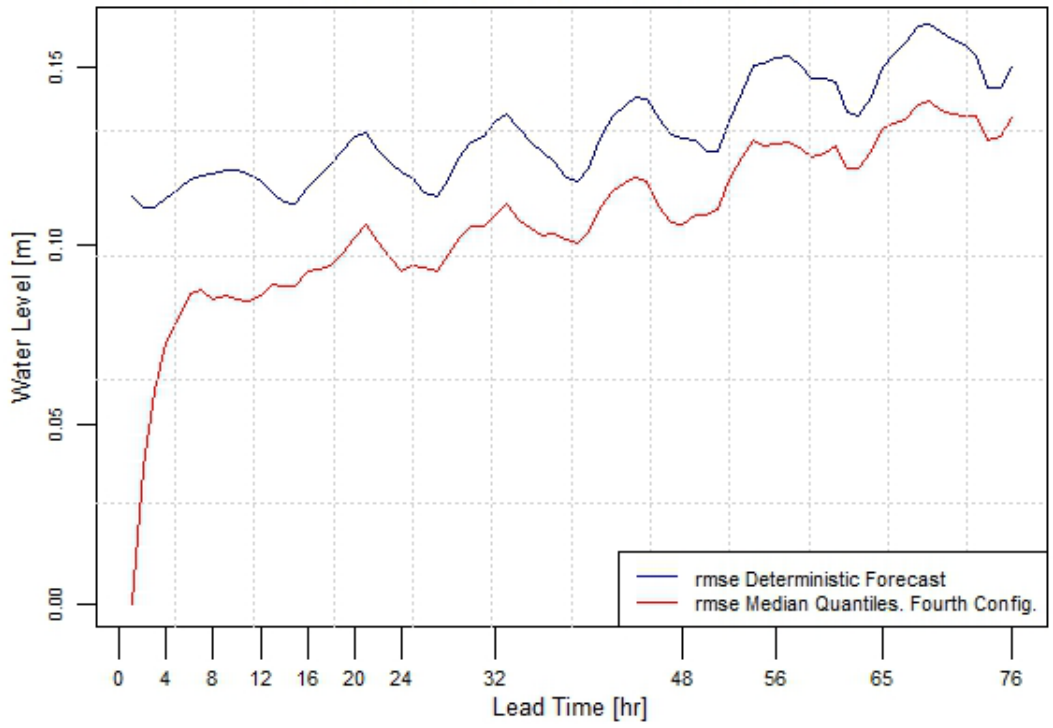


Figure 41: RMSE of the deterministic forceast (black line) and the meadian quantile forecast based on QR with bias correction (red line).

Error model extrapolation to model boundaries

Figure 42 shows plots of forecasted water levels at Mosede harbour and at the open sea model boundary for several forecast datasets, which indicate that they are very similar. Although there are some differences among the time series, they seem to be linearly related. Scatter plots of the water levels at the boundary and Mosede harbour are shown in Figure 43 for three different lead times.

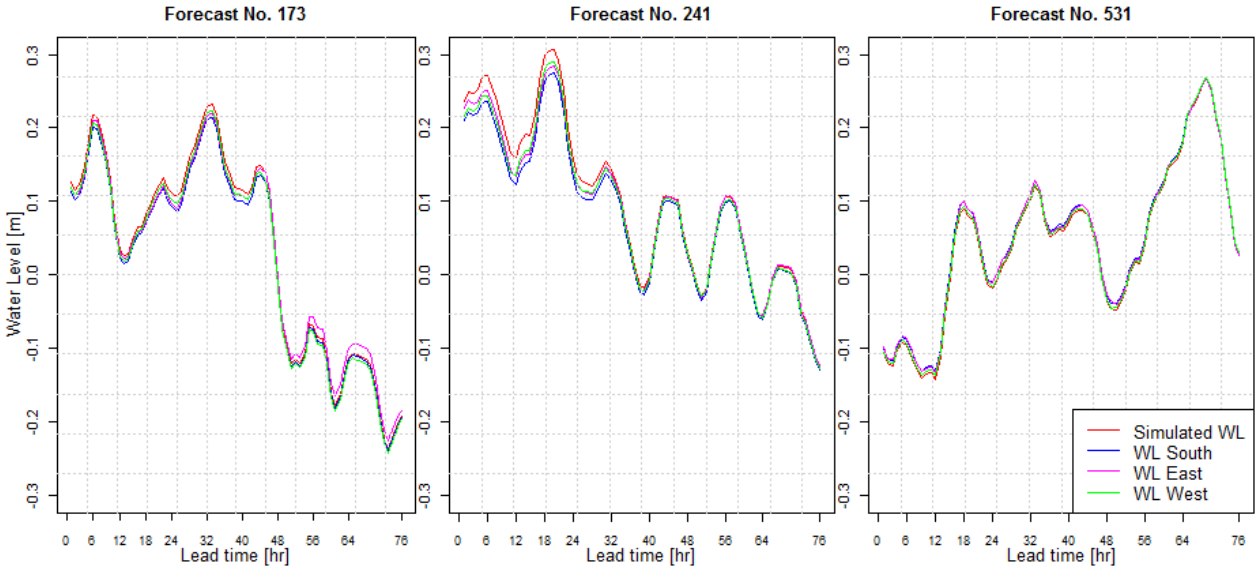


Figure 42: Sea water level boundaries and simulated values at Mosede harbour in selected times of forecast.

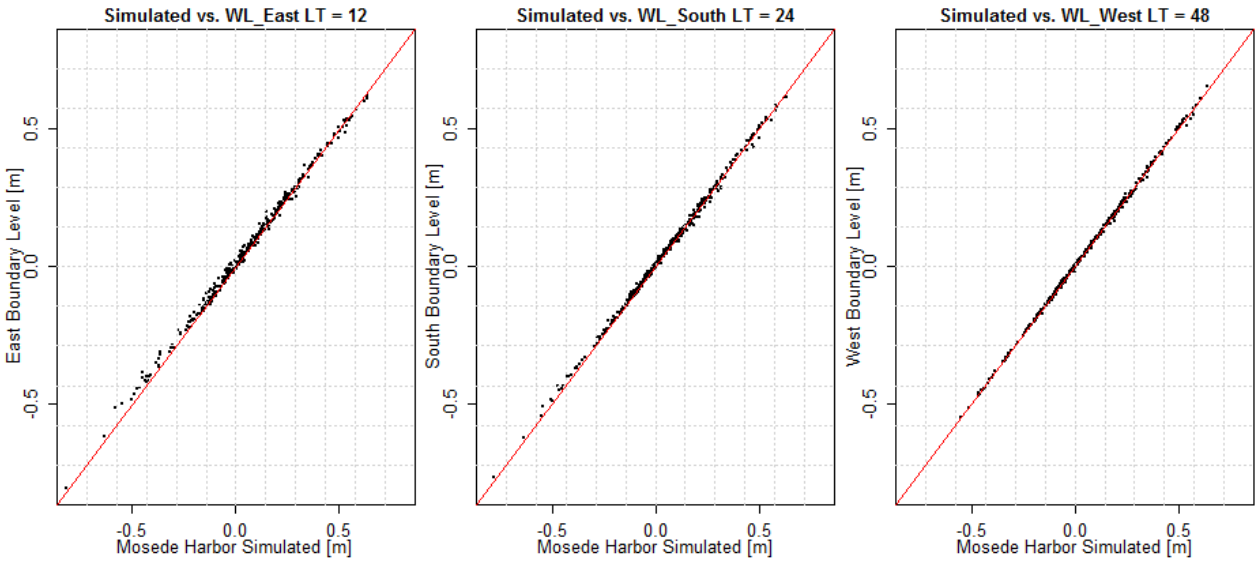


Figure 43: Scatter plots of simulated water levels at Mosede harbour and at model boundaries for 12, 24, 48 hrs lead times.

Linear regression models were estimated for extrapolating the error model from the harbour to the model boundaries, which are then used to derive probabilistic forecasts from deterministic forecasts at the boundaries.

Probabilistic flood mapping

No extreme storm surge events were available in the historical record used in this study. Therefore, in order to illustrate derivation of flood maps from probabilistic sea level forecasts, synthetic extreme sea level time series using a combination of the maximum forecasted water levels from the available dataset in this study and maximum water level projections near Greve from earlier climate change studies were used.

Estimates for mean sea level rise in Denmark by GEUS and DMI (2012) indicate that in Køge Bay mean sea level is projected to rise by around 0.75 ± 0.6 m (Figure 44). In addition, storm surges will change due to changes in storm intensity and patterns. DHI has analysed changes in storm surges in the inner Danish waters based on three regional climate model projections. Results for Køge Bay are shown in Table 5. Based on these studies the most extreme estimates of storm surge levels for different return periods are shown in Table 6.

Table 5: Changes in storm surge signal [m] in Køge Bay in 2100 based on the three regional climate model projections for different return periods.

Return Period [yrs]	HIRHAM_ARPEGE	HIRHAM_ECHAM5	HIRHAM_BCM	Average
10	-0.14	0.14	-0.06	-0.02
50	-0.17	0.19	-0.07	-0.02
100	-0.18	0.21	-0.08	-0.02

Figure 44: Estimates of mean water level increase in year 2100 compared to today. The location of Køge Bay, near Greve, is shown. (Source: GEUS and DMI, 2012)

Table 6: Most extreme estimates for sea levels in Køge Bay considering changes in storm surge signal and mean sea level rise for different return periods T.

T [yrs]	Current Extreme Statistics [m]	Year 2100 Upper Estimates [m]
10	1.35	2.84
50	1.47	3.01
100	1.52	3.08

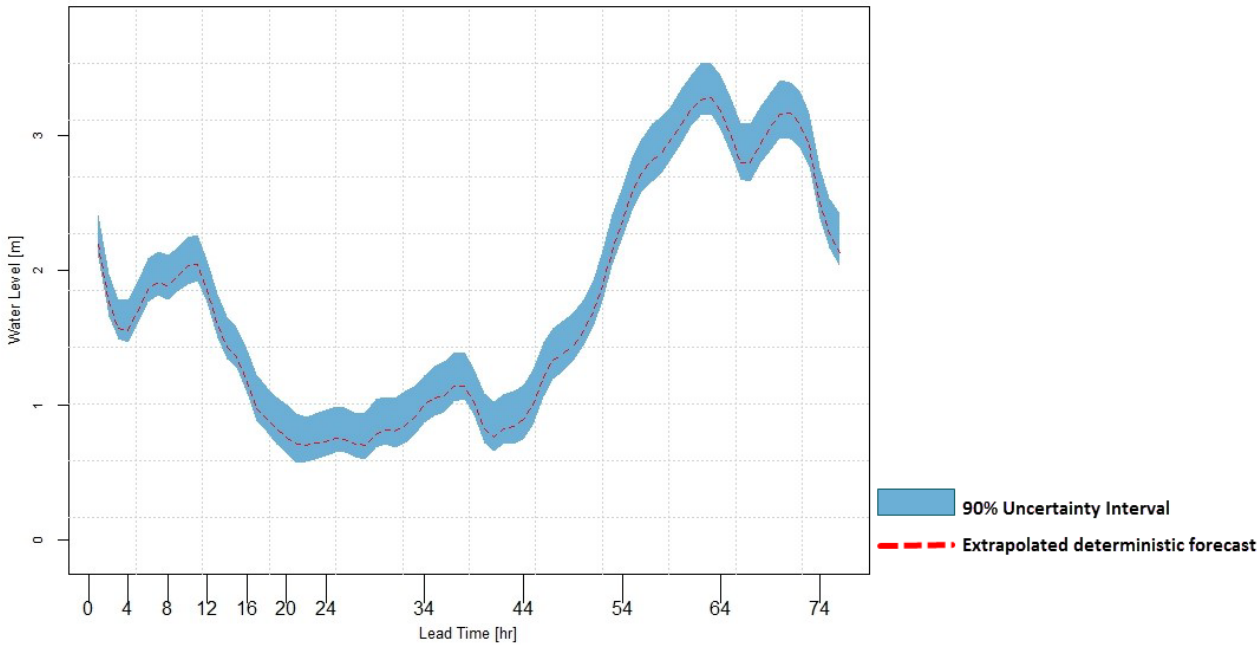


Figure 45: Extrapolated deterministic and probabilistic 'forecast' water levels (bias-corrected) derived using available forecast dataset with the highest water level values and climate change projections for storm surge signal and mean sea level in Køge Bay corresponding to a future 100-year event.

A water level time series from the available forecast dataset was scaled to fit the water level of a future 100-year event, resulting in the water level time series shown in Figure 45. This represents a synthetic extreme sea level event, which was used as input to the Greve coastal flood model in order to obtain a set of probabilistic flood maps for the area (Figure 46).

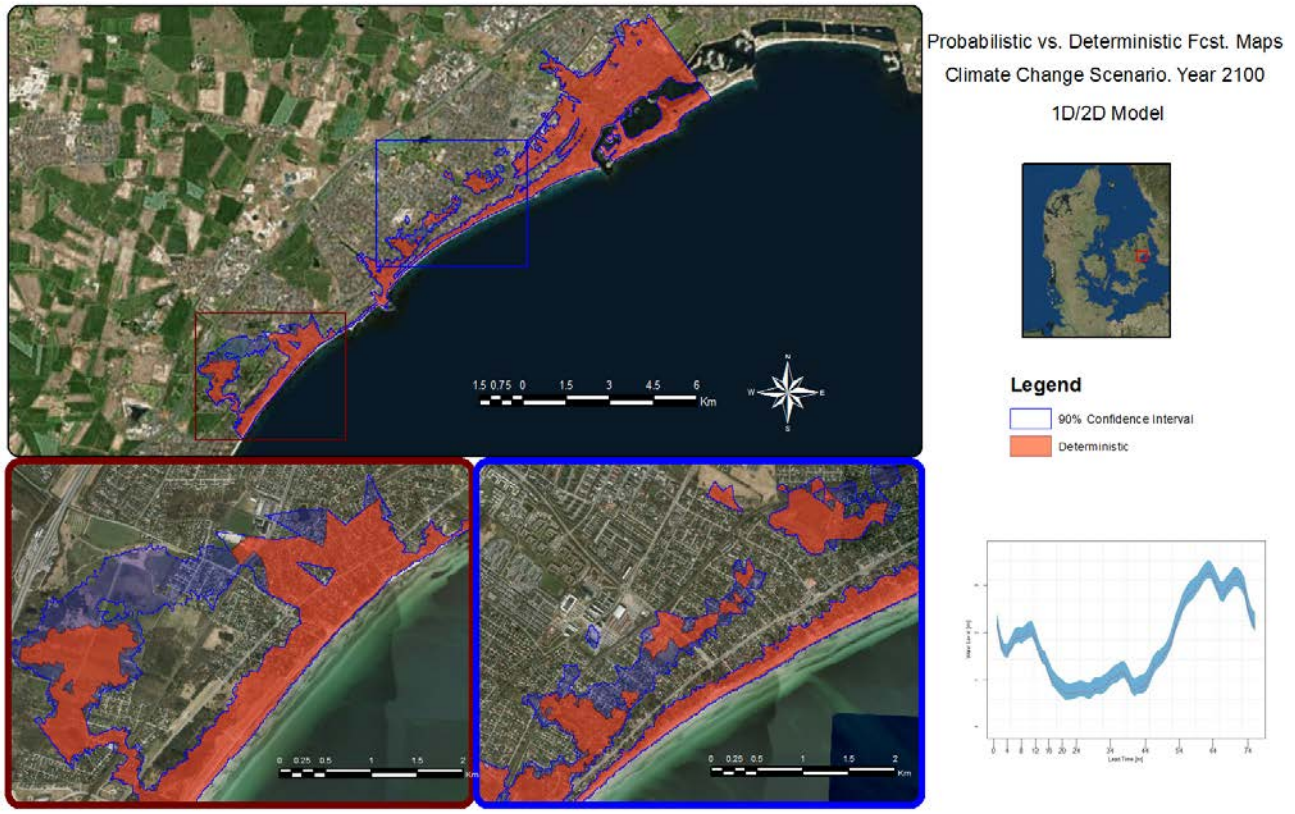


Figure 46: Flood maps obtained from the synthetic deterministic and 90% confidence interval water level time series.

Another way of presenting probabilistic flood maps is shown in Figure 47. Here, maximum flooding from water levels of different probabilities of occurrence are shown together in one map, allowing authorities to make more risk-based decisions in dealing with the potential flood event. This method of flood risk presentation has been successfully used by weather and forecast agencies, such as NOAA (Demargne, 2006). Examining the set of probabilistic flood maps obtained from this exercise, flood characteristics are compared and summarized in Table 5. A difference of 1.93 km^2 (around 22%) in flooded area is observed between flood events with 5% and 95% probability of exceedance.

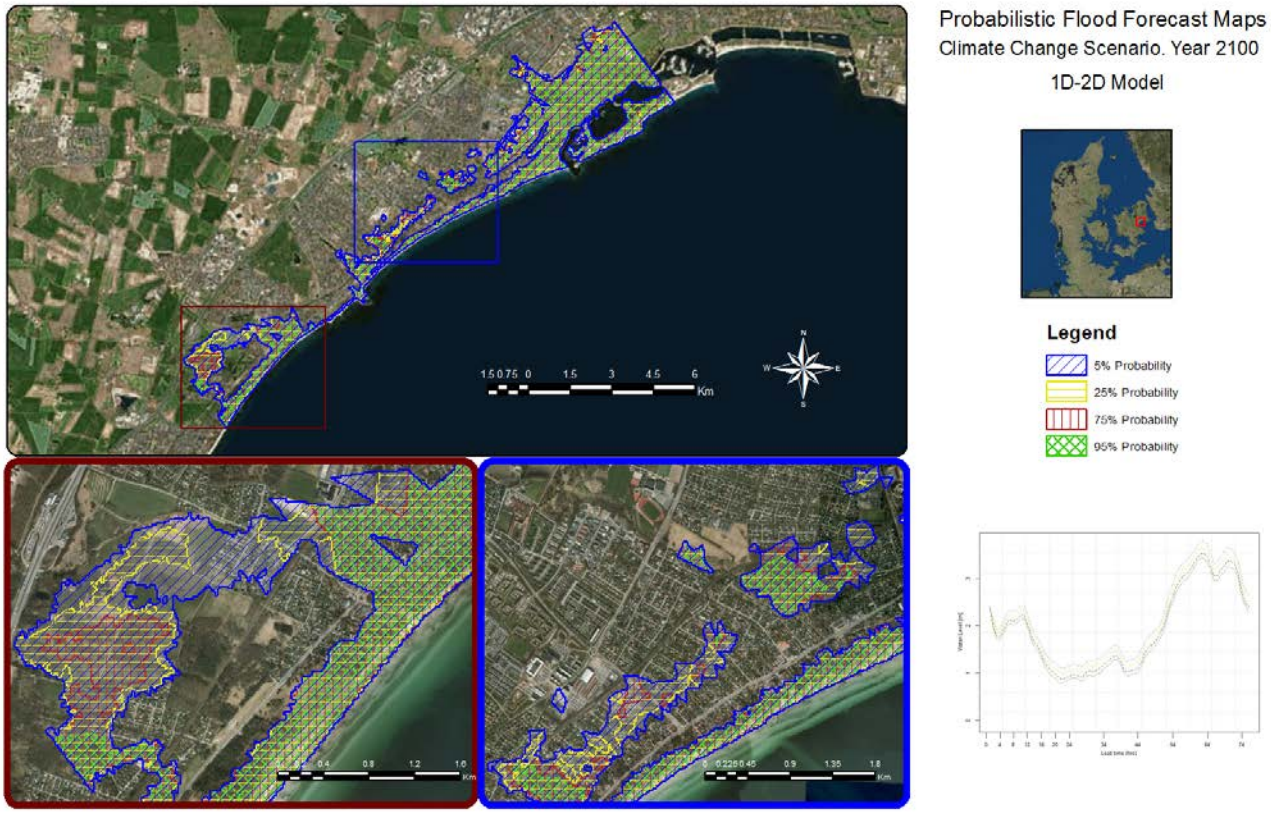


Figure 47: Probabilistic Flood Maps for a selected time of forecast using the synthetic forecast event.

Table 7: Comparison of total flood area [km^2], and number of wet elements for maximum flood maps obtained from deterministic and probabilistic water level events.

Probability	Flood Area	Total Wet Elements
5%	8.40	84 028
25%	7.80	77 921
75%	6.99	69 921
95%	6.47	64 693
Deterministic	7.21	70 178

It should be noted that the synthetic storm surge event used in the probabilistic flood mapping exercise was around four times larger than the largest event in the dataset used to estimate the QR model. This could cause an unrealistic representation of the uncertainty bound since the model was not trained for such large water level predictions. This further highlights the importance of having training/calibration datasets that are long enough and contain a large variability of events for developing the error model.

4.2.3 Summary

Despite the advantages of probabilistic flood maps, they require additional analysis and additional simulation runs. Nevertheless, stochastic models based on QR are relatively easy to apply once the error model has been built. QR, as a post-processor method, requires long training and verification datasets, containing as much variability and range of events as possible. The user should be aware that the efficiency of the stochastic model may decrease with events outside of the training data.

References

——, 2004: Scale Dependence of the Predictability of Precipitation from Continental Radar Images. Part II: Probability Forecasts. Journal of Applied Meteorology, 43, 74-89.

Abily, Morgan, et al.2015. "Uncertainty related to high resolution topographic data use for flood event modeling over urban areas: toward a sensitivity analysis approach." ESAIM: Proceedings and Surveys 48: 385-399.

Alemohammad, Seyed Hamed, Doctor of Philosophy Thesis, submitted-accepter February 2015: “Characterization of Uncertainty in Remotely-Sensed Precipitation Estimates,” Department of Civil and Environmental Engineering, Massachusetts Institute of Technology.

Alfieri, L., Salamon, P., Pappenberger, F., Wetterhall, F., Thielen, J., 2012. Operational early warning systems for water-related hazards in Europe. Environ. Sci. Policy 21, 35–49. doi:10.1016/j.envsci.2012.01.008

Andersson, T., K.-I. Ivarsson, 1991: A Model for Probability Nowcasts of Accumulated Precipitation Using Radar. Journal of Applied Meteorology, 30, 135-141.

Arheimer, B., Lindström, G., Olsson, J., 2011. A systematic review of sensitivities in the Swedish flood-forecasting system. Atmos. Res. 100, 275–284. doi:10.1016/j.atmosres.2010.09.013

Aronica, G., Bates, P.D., Horritt, M.S., 2002. Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. Hydrol. Process. 16, 2001–2016. doi:10.1002/hyp.398

Aronica, G., Hankin, B., Beven, K., 1998. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. Adv. Water Resour. 22, 349–365. doi:10.1016/S0309-1708(98)00017-7

Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T., 2001. A tutorial on Particle Filters for On-line Non- Linear/Non-Gausssian Bayesin Tracking. Ieee 50, 174–188.

Austin, G. L., A. Bellon, 1974: The use of digital weather radar records for short-term precipitation forecasting. Quarterly Journal of the Royal Meteorological Society, 100, 658-664.

Bartholmes, Todini, 2005. Coupling meteorological and hydrological models for flood forecasting. Hydrol. Earth Syst. Sci. 9, 333–346. doi:10.5194/hess-9-333-2005

Berenguer M., D. Sempere-Torres and G.G.S. Pegram, 2011: SBMcast – An ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation, Journal of Hydrology, 404, 226-240.

Berenguer, M., C. Corral, R. Sánchez-Diezma, D. Sempere-Torres, 2005: Hydrological Validation of a Radar-Based Nowcasting Technique. Journal of Hydrometeorology, 6, 532-549, doi:10.1175/JHM433.1

Bergstrom, S., 1976. Development and Application of a conceptual runoff model for Scandinavian catchment. 153.

Beven, K., 2006. A manifesto for the equifinality thesis. J. Hydrol. 320, 18–36. doi:10.1016/j.jhydrol.2005.07.007

Beven, K., 2009. Environmental Modelling: An Uncertain Future, Environmental Modelling : An Uncertain Future ?

Beven, K., 2012. Rainfall-Runoff Modelling, Rainfall-Runoff Modelling: The Primer: Second Edition. John Wiley & Sons, Ltd, Chichester, UK. doi:10.1002/9781119951001

Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrol. Process. 6, 279–298. doi:10.1002/hyp.3360060305

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. 249, 11–29. doi:10.1016/S0022-1694(01)00421-8

Bitew, M. M., M. Gebremichael, L. T. Gebremichael, and Y. A. Bayissa, 2012: "Evaluation of high-resolution satellite rainfall products through streamflow simulation in a hydrological modeling of a small mountainous watershed in Ethiopia." J. Hydrometeor., vol. 13, pp. 338 – 350.

Borga, M., Anagnostou, E.N., Blöschl, G., Creutin, J.-D., 2011. Flash flood forecasting, warning and risk management: the HYDRATE project. Environ. Sci. Policy 14, 834–844. doi:10.1016/j.envsci.2011.05.017

Borga, M., Stoffel, M., Marchi, L., Marra, F., Jakob, and M., 2014: "Hydrogeomorphic response to extreme rainfall in headwater systems: flash floods and debris flows." J. Hydrol. http://dx.doi.org/10.1016/j.jhydrol.2014.05.022 (in press).

Bowler, N.E.H., Pierce, C.E., Seed, A., 2004. Development of a precipitation nowcasting algorithm based upon optical flow techniques. J. Hydrol. 288, 74–91. doi:10.1016/j.jhydrol.2003.11.011

Bray, M., D. Han, Y. Xuan, P. Bates, and M. Williams, 2011: "Rainfall uncertainty for extreme events in NWP downscaling model." Hydrol. Processes, vol. 25, pp. 1397 – 1406.

Buil, A., 2017: Nowcasting probabilístico basado en observaciones de lluvia con radar meteorológico. PhD Thesis. Universitat Politécnica de Catalunya. 173 pp.

Buizza, R., Houtekamer, P.L., Pellerin, G., Toth, Z., Zhu, Y., Wei, M., 2005. A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems. Mon. Weather Rev. 133, 1076–1097. doi:10.1175/MWR2905.1

Burnash, R.J.C., Ferral, R.L., McGuire, R.A., McGuire, R.A., Center, U.S.J.F.-S.R.F., 1973. A Generalized Streamflow Simulation System: Conceptual Modeling for Digital Computers. U.S. Department of Commerce, National Weather Service, and State of California, Department of Water Resources.

Chiang, Y.-M., Hsu, K.-L., Chang, F.-J., Hong, Y., Sorooshian, S., 2007. Merging multiple precipitation sources for flash flood forecasting. J. Hydrol. 340, 183–196. doi:10.1016/j.jhydrol.2007.04.007

Ciarapica, L., Todini, E., 2002. TOPKAPI: a model for the representation of the rainfall-runoff process at different scales. Hydrol. Process. 16, 207–229. doi:10.1002/hyp.342

Clark, M.P., Rupp, D.E., Woods, R.A., Zheng, X., Ibbitt, R.P., Slater, A.G., Schmidt, J., Uddstrom, M.J., 2008. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. Adv. Water Resour. 31, 1309–1324. doi:10.1016/j.advwatres.2008.06.005

Cloke, H.L., Pappenberger, F., 2009. Ensemble flood forecasting: A review. J. Hydrol. 375, 613–626. doi:10.1016/j.jhydrol.2009.06.005

Collier, C.G., 2007. Flash flood forecasting: What are the limits of predictability? Quarterly Journal of the Royal Meteorological Society 133, 3–23. doi:10.1002/qj.29

CRED, 2016. Centre for Research on the Epidemiology of Disasters in Brussels. OFDA/CRED Int. Disaster Database, Univ. Cathol. Louvain, Brussels, Belgium.

Dabbert, W., J. Hale, W. F. Krajewski, A. Crook, and C. Mueller, 2000: "Forecast issues in the urban zone: Report of the 10th prospectus development team of the U.S. Weather Research Program." Bull. Amer. Meteor. Soc., vol. 81, pp. 2047 – 2064.

De Roo, A.P.J., Wesseling, C.G., Van Deursen, W.P.A., 2000. Physically based river basin modelling within a GIS: the LISFLOOD model. Hydrol. Process. 14, 1981–1992. doi:10.1002/1099-1085(20000815/30)14:11/12<1981::AID-HYP49>3.0.CO;2-F

Delestre, 2010. Delestre, O. 2010. Simulation du ruissellement d'eau de pluie sur des surfaces agricoles/ rain water over- land ow on agricultural _elds simulation. PhD thesis, Université d'Orléans (in French), available from TEL: tel.archivesouvertes.fr/INSMI/tel-00531377/fr

Demargne, J., 2006. Probabilistic Flood Mapping. RFC Short-Term Ensemble Workshop. NOAA/National Weather Service.

DHI, 2014a. MIKE FLOOD, ID-2D Modelling, User Manual. DHI: Hørsholm, Denmark.

DHI, 2014b. MIKE 21 & MIKE 3 Flow Model FM, Hydrodynamic and Transport Module, Scientific Documentation. DHI: Hørsholm, Denmark.

Di Baldassarre, G., Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. Hydrol. Earth Syst. Sci 13, 913–921. doi:10.5194/hessd-6-39-2009

Diez, R. F., 2016. From Deterministic to Probabilistic Flood Forecasting of Storm Surges (Case Study: Greve, Denmark). (Master's thesis). University of Nice Sophia Antipolis, Nice, France.

Dinku, T., P. Ceccato, E. Grover-Kopec, M. Lemma, S. J. Connor, and C. F. Ropelewski, 2007: "Validation of satellite rainfall products over East Africa's complex topography." Int. J. Remote Sens., vol. 28, pp. 1503 – 1526.

Du, J., Mullen, S.L., Sanders, F., 1997. Short-Range Ensemble Forecasting of Quantitative Precipitation. Mon. Weather Rev. 125, 2427–2459. doi:10.1175/1520-0493(1997)125<2427:SREFOQ>2.0.CO;2

Ebert, E. E., J. E. Janowiak, and C. Kidd, 2007: "Comparison of near real-time precipitation estimates from satellite observations and numerical models." Bull. Amer. Meteor. Soc., Vol. 88, pp. 47 – 64.

Ebtehaj, M., Moradkhani, H., Gupta, H. V., 2010. Improving robustness of hydrologic parameter estimation by the use of moving block bootstrap resampling. Water Resour. Res. 46, n/a-n/a. doi:10.1029/2009WR007981

Efstratiadis, a., Nalbantis, I., Koutsoyiannis, D., 2015. Hydrological modelling of temporally-varying catchments: facets of change and the value of information. Hydrol. Sci. J. 60, 1438–1461. doi:10.1080/02626667.2014.982123

Efstratiadis, A., Koutsoyiannis, D., 2010. One decade of multi-objective calibration approaches in hydrological modelling: a review. Hydrol. Sci. J. 55, 58–78. doi:10.1080/02626660903526292

England, J.F., Velleux, M.L., Julien, P.Y., 2007. Two-dimensional simulations of extreme floods on a large watershed. J. Hydrol. 347, 229–241. doi:10.1016/j.jhydrol.2007.09.034

Estupina-Borrell, V., Dartus, D., Ababou, R., 2006. Flash flood modeling with the MARINE hydrological distributed model. Hydrol. Earth Syst. Sci. Discuss. 3, 3397–3438. doi:10.5194/hessd-3-3397-2006

European Commission, 2004. Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions – flood risk management – flood prevention, protection and mitigation.

European Commission, 2007. Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks 27–34.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. doi:10.1029/94JC00572

Evensen, G., 2003. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. Ocean Dyn. 53, 343–367. doi:10.1007/s10236-003-0036-9

Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. Ocean Dyn. 54, 539–560. doi:10.1007/s10236-004-0099-2

Evensen, G., van Leeuwen, P.J., 2000. An ensemble Kalman smoother for nonlinear dynamics. Mon. Weather Rev. 128, 1852–1867. doi:10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2

Germann, U., I. Zawadzki, 2002: Scale-Dependence of the Predictability of Precipitation from Continental Radar Images. Part I: Description of the Methodology. Monthly Weather Review, 130, 2859-2873.

Germann U. and I. Zawadzki, 2004: Scale Dependence of the Predictability of Precipitation from Continental Radar Images. Part II: Probability Forecasts. J. Appl. Meteor., 43, 74–89.

Germann, U., I. Zawadzki, B. Turner, 2006: Predictability of Precipitation from Continental Radar Images. Part IV: Limits to Prediction. Journal of the Atmospheric Sciences, 63, 2092-2108.

Germann, U., M. Berenguer, D. Sempere-Torres, and G. Salvade, 2006: Ensemble radar precipitation estimation - a new topic on the radar horizon. Proceedings of Fourth European Conference on Radar in Meteorology and Hydrology. Barcelona (Spain). 559-562. ISBN: 978-848181-227-5.

Germann, U., M. Berenguer, D. Sempere-Torres, and M. Zappa, 2009: REAL - Ensemble radar precipitation estimation for hydrology in a mountainous region. Quarterly Journal of the Royal Meteorological Society, 135, 445-456.

GEUS and DMI, 2012. *Ændringer af havniveauet i Danmark de næste 100-200 år*, Notat til Klima-, Energi- og Bygningsministeriet, GEUS og DMI, Udarbejdet af Karen Edelvang, Andreas Ahlstrøm, Camilla Snowman Andresen, Signe Bech Andersen, Ole Bennike, Jens Morten Hansen, Antoon Kuijpers og Birger Larsen fra De Nationale Geologiske Undersøgelser for Danmark og Grønland (GEUS) og Erik Buch, Katrine Krogh Andersen og Kristine S. Madsen fra Danmarks Meteor ologiske Institut (DMI), 29. februar 2012. Retrieved from https://www.dmi.dk/.../notat_vandstand_geus_dmi.pdf.

Golding, B.., 2000. Quantitative precipitation forecasting in the UK. J. Hydrol. 239, 286–305. doi:10.1016/S0022-1694(00)00354-1

Golding, B.W., 1998. Nimrod: A system for generating automated very short range forecasts. Meteorol. Appl. 5, S1350482798000577. doi:10.1017/S1350482798000577

Greve Kommune, 2007. *Oversvømmelserne i Greve Kommune Juli 2007*. Technical report. Retrieved from http://www.gsforsyning.dk/sites/default/files/oversvommelserne_i_greve_kommune_juli_2007.pdf

Gupta, H.V.,S. Sorooshian, T.S. Hogue and P. Boyle, 2003. Advances in Automatic Calibration of Watershed Models, in Water Science and Application Series, Volume 6, American Geophysical Union

Hapuarachchi, H.A.P., Wang, Q.J., Pagano, T.C., 2011. A review of advances in flash flood forecasting. Hydrol. Process. 25, 2771–2784. doi:10.1002/hyp.8040

Huffman G.J., 1997: “Estimates of root-mean-square random error for finite samples of estimated precipitation” J. Hydrometeorol., vol. 36, pp. 1191– 2101.

Hutton, C.J., Kapelan, Z., Vamvakeridou-Lyroudia, L., Savić, D.A., 2014. Dealing with Uncertainty in Water Distribution System Models: A Framework for Real-Time Modeling and Data

Assimilation. J. Water Resour. Plan. Manag. 140, 169–183. doi:10.1061/(ASCE)WR.1943-5452.0000325

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: “CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution.” J. Hydrometeor., vol. 5, pp. 487 – 503.

Kalman, R.E., 1960. A New Approach to Linear Filtering and Prediction Problems. J. Basic Eng. 82, 35. doi:10.1115/1.3662552

Kavetski, D., Kuczera, G., Franks, S.W., 2006. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. Water Resour. Res. 42, 1–9. doi:10.1029/2005WR004368

Kelly, K.S., Krzysztofowicz, R., 2000. Precipitation uncertainty processor for probabilistic river stage forecasting. Water Resour. Res. 36, 2643–2653. doi:10.1029/2000WR900061

Kidd, C., P. Bauer, J. Turk, G. J. Huffman, R. Joyce, K.-L. Hsu, and D. Braithwaite, 2012: “Intercomparison of high-resolution precipitation products over northwest Europe.” J. Hydrometeor., vol. 13, pp. 67 – 83.

Kim, G., Barros, A.P., 2001. Quantitative flood forecasting using multisensor data and neural networks. J. Hydrol. 246, 45–62. doi:10.1016/S0022-1694(01)00353-5

Kirstetter, P.-E., Y. Hong, J. Gourley, M. Schwaller, W. Petersen, and J. Zhang, 2012a: Comparison of TRMM 2A25 Products Version 6 and Version 7 with NOAA/NSSL Ground Radar-based National Mosaic QPE. Journal of Hydrometeorology.

Kobold, M., Brilly, M., 2006. The use of HBV model for flash flood forecasting. Nat. Hazards Earth Syst. Sci. 6, 407–417. doi:10.5194/nhess-6-407-2006

Kochilakis, G., Poursanidis, D., Chrysoulakis, N., Varella, V., Kotroni, V., Eftychidis, G., Lagouvardos, K., Papathanasiou, C., Karavokyros, G., Aivazoglou, M., Makropoulos, C., Mimikou, M., 2016. A web based DSS for the management of floods and wildfires (FLIRE) in urban and periurban areas. Environ. Model. Softw. 86, 111–115. doi:10.1016/j.envsoft.2016.09.016

Koren, V., Reed, S., Smith, M., Zhang, Z., Seo, D.-J., 2004. Hydrology laboratory research modeling system (HL-RMS) of the US national weather service. J. Hydrol. 291, 297–318. doi:10.1016/j.jhydrol.2003.12.039

Krzysztofowicz, R., 1999. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. Water Resour. Res. 35, 2739–2750. doi:10.1029/1999WR900099

Krzysztofowicz, R., 2001. Integrator of uncertainties for probabilistic river stage forecasting: precipitation-dependent model. J. Hydrol. 249, 69–85. doi:10.1016/S0022-1694(01)00413-9

Krzysztofowicz, R., 2002. Bayesian system for probabilistic river stage forecasting. J. Hydrol. 268, 16–40. doi:10.1016/S0022-1694(02)00106-3

Krzysztofowicz, R., Kelly, K.S., 2000. Hydrologic uncertainty processor for probabilistic river stage forecasting. Water Resour. Res. 36, 3265–3277. doi:10.1029/2000WR900108

Kuczera, G. and Parent, E., 1998. Monte Carlo assessment and parameter uncertainty in conceptual catchment models: the Metropolis algorithm, J. Hydrol. 211, 69-85.

Lean, H.W., Clark, P. a., 2003. The effects of changing resolution on mesocale modelling of line convection and slantwise circulations in FASTEX IOP16. Q. J. R. Meteorol. Soc. 129, 2255–2278. doi:10.1256/qj.02.57

Li, L., W. Schmid, J. Joss, 1995: Nowcasting of Motion and Growth of Precipitation with Radar over a Complex Orography. Journal of Applied Meteorology, 34, 1286-1300.

Liechti, K., M. Zappa, F. Fundel, and U. Germann, 2013: “Probabilistic evaluation of ensemble discharge nowcasts in two nested Alpine basins prone to flash floods. Hydrological Processes HEPS Special Issue, vol. 27, pp. 5 – 17.

Liu YQ, Gupta HV. 2007. Uncertainty in hydrologic modeling: toward an integrated data assimilation framework. Water Resources Research 43: W07401. DOI:10.1029/2006WR005756.

Liu, Y., Freer, J., Beven, K., Matgen, P., 2009. Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error. J. Hydrol. 367, 93–103. doi:10.1016/j.jhydrol.2009.01.016

Llort X., C. Velasco-Forero, J. Roca-Sancho, D. Sempere-Torres, 2008: Characterization of uncertainty in radar-based precipitation estimates and ensemble generation. *Proceedings of Fifth European Conference on Radar in Meteorology and Hydrology (ERAD2008)*. Helsinki (Finland). ISSN: 978-951-697-676-4.

Llort X., R. Sánchez-Diezma, A. Rodríguez, D. Sancho, M. Berenguer, D. Sempere-Torres, 2014: FloodAlert: A simplified radar-based EWS for urban flood warning. Proceedings of the 11th International Conference on Hydroinformatics HIC 2014, New York City, USA. ISBN: 978-0-692-28129-1.

Obligeois, F., V. Andréassian, C. Perrin, P. Tabary, and C. Loumagne, 2014: “When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events.” Hydrol. Earth Syst. Sci., vol. 18, pp. 575 – 594.

Lopez Lopez, P., Verkade, J.S., Weerts, A.H., Solomatine, D.P., 2014, Alternative configurations of quantile regression for estimating predictive uncertainty in water level forecasts for the upper Severn River: A comparison, Hydrol. Earth Syst. Sci. 18, 3411–3428. doi:10.5194/hess-18-3411-2014.

Lorenz, E.N., 1969. The predictability of a flow which possesses many scales of motion. Tellus 21, 289–307. doi:10.1111/j.2153-3490.1969.tb00444.x

Maddox R.A., et al., 2002: “Weather radar coverage over the contiguous United States” Weather Forecasting, vol. 17, pp. 927–934.

Madsen, H., Canizares, R., 1999. Comparison of extended and ensemble kalman filers for data assimilation in coastal area modeling. Int. J. Numer. Meth. Fluids 31, 961–981. doi:10.1002/(SICI)1097-0363(19991130)31:6<961::AID-FLD907>3.3.CO;2-S

Madsen, H., Skotner, C., 2005. Adaptive state updating in real-time river flow forecasting—a combined filtering and error forecasting procedure. J. Hydrol. 308, 302–312. doi:10.1016/j.jhydrol.2004.10.030

Marrel et al., 2012. Global sensitivity analysis of stochastic computer models with joint metamodels. Statistics and Computing, 22(3):833-847

Marshall, J. S. and W. M. Palmer, 1948: The distribution of raindrops with size. *Journal of Meteorology*, 5, 165-166.

Matott LS, Babendreier JE, Purucker ST. 2009. Evaluating uncertainty in integrated environmental models: a review of concepts and tools. Water Resources Research 45: W06421. DOI:10.1029/2008WR007301.

Matsangouras, I.T., P.T. Nastos, and I. Pytharoulis, 2015: “Study of the tornado event in Greece on March 25, 2009: Synoptic analysis and numerical modeling using modified topography,” Atmos. Res., http://dx.doi.org/10.1016/j.atmosres.2015.08.010.

McMillan, H., Freer, J., Pappenberger, F., Krueger, T., Clark, M., 2010. Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions. Hydrol. Process. n/a-n/a. doi:10.1002/hyp.7587

McMillan, H., Jackson, B., Clark, M., Kavetski, D., Woods, R., 2011. Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models. J. Hydrol. 400, 83–94. doi:10.1016/j.jhydrol.2011.01.026

Meischner, P. (ed.) (2004). *Weather radar. Principles and advanced applications*. Springer-Verlag, ISBN 3-540-000328-2, Berlin – Heidelberg, Germany.

Montanari, A., 2005. Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. Water Resour. Res. 41, n/a-n/a. doi:10.1029/2004WR003826

Montanari, A., 2007. What do we mean by “uncertainty”? The need for a consistent wording about uncertainty assessment in hydrology. Hydrol. Process. 21, 841–845. doi:10.1002/hyp.6623

Montanari, A., Brath, A., 2004. A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. Water Resour. Res. 40, n/a-n/a. doi:10.1029/2003WR002540

Montanari, A., Di Baldassarre, G., 2013. Data errors and hydrological modelling: The role of model structure to propagate observation uncertainty. Adv. Water Resour. 51, 498–504. doi:10.1016/j.advwatres.2012.09.007

Montanari, A., Grossi, G., 2008. Estimating the uncertainty of hydrological forecasts: A statistical approach. Water Resour. Res. 44, 1–9. doi:10.1029/2008WR006897

Montanari, A., Koutsoyiannis, D., 2012. A blueprint for process-based modeling of uncertain hydrological systems. Water Resour. Res. 48, n/a-n/a. doi:10.1029/2011WR011412

Moore, R.J., Cole, S.J., Bell, V. a., Jones, D. a., 2006. Issues in flood forecasting: ungauged basins, extreme floods and uncertainty. Front. Flood Res. 305, 103–122.

Moradkhani, H., Hsu, K.-L., Gupta, H., Sorooshian, S., 2005a. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. Water Resour. Res. 41, W05012. doi:10.1029/2004WR003604

Moradkhani, H., Sorooshian, S., Gupta, H. V., Houser, P.R., 2005b. Dual state-parameter estimation of hydrological models using ensemble Kalman filter. Adv. Water Resour. 28, 135–147. doi:10.1016/j.advwatres.2004.09.002

Mugnai, A., Casella, D., Cattani, E., Dietrich, S., Laviola, S., Levizzani, V., Panegrossi, G., Petracca, M., Sanò, P., Di Paola, F., Biron, D., De Leonibus, L., Melfi, D., Rosci, P., Vocino, A., Zauli, F., Pagliara, P., Puca, S., Rinollo, A., Milani, L., Porcù, F., and Gattari, F., 2013: Precipitation products from the hydrology SAF, Nat. Hazards Earth Syst. Sci., 13, 1959-1981.

Muthusamy, M., Godiksen, P. N., & Madsen, H., 2016. Comparison of different configurations of quantile regression in estimating predictive hydrological uncertainty. Procedia Engineering, 154 513-520.

Naturstyrelsen og Kystdirektoratet, 2011, Endelig udpegning af risikoområder for oversvømmelse fra vandløb, søer, havet og fjorde, Naturstyrelsen, Miljøministeriet og Kystdirektoratet, Transportministeriet, ISBN 978-87-7279-298-9 (PDF)

Neuman, S.P., 2003. Maximum likelihood Bayesian averaging of uncertain model predictions. Stoch. Environ. Res. Risk Assess. 17, 291–305. doi:10.1007/s00477-003-0151-7

Nikolopoulos, E. I., E. N. Anagnostou, and M. Borga, 2013: “Using High-Resolution Satellite Rainfall Products to Simulate a Major Flash Flood Event in Northern Italy.” J. Hydrometeor., vol. 14, pp. 171 – 185.

Nikolopoulos, E. I., E. N. Anagnostou, F. Hossain, M. Gebremichael, and M. Borga, 2010: "Understanding the scale relationships of uncertainty propagation of satellite rainfall through a distributed hydrologic model." J. Hydrometeor., vol. 11, pp. 520 – 532.

Nikolopoulos, E. I., S. Crema, L. Marchi, F. Marra, F. Guzzetti, and M. Borga, 2014: "Impact of uncertainty in rainfall estimation on the identification of rainfall thresholds for debris flow occurrence." Geomorphology, vol. 221, pp. 286 – 297.

Pappenberger F, Ghelli A, Buizza R, Bodis K. 2009. The skill of probabilistic precipitation forecasts under observational uncertainties within the generalized likelihood uncertainty estimation framework for hydrological applications. *Journal of Hydrometeorology* 10:807–819.

Pappenberger, F., Bartholmes, J., Thielen, J., Cloke, H.L., Buizza, R., de Roo, A., 2008. New dimensions in early flood warning across the globe using grand-ensemble weather predictions. Geophys. Res. Lett. 35, L10404. doi:10.1029/2008GL033837

Pappenberger, F., Beven, K., Horritt, M., Blazkova, S., 2005a. Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations. J. Hydrol. 302, 46–69. doi:10.1016/j.jhydrol.2004.06.036

Pappenberger, F., Beven, K.J., 2004. Functional classification and evaluation of hydrographs based on Multicomponent Mapping (Mx). Int. J. River Basin Manag. 2, 89–100. doi:10.1080/15715124.2004.9635224

Pappenberger, F., Beven, K.J., Hunter, N.M., Bates, P.D., Gouweleeuw, B.T., Thielen, J., de Roo, A.P.J., 2005b. 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). Hydrol. Earth Syst. Sci. 9, 381–393. doi:10.5194/hess-9-381-2005

Pappenberger, F., Frodsham, K., Beven, K.J., Romanovicz, R., Matgen, P., 2007. Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations. Hydrol. Earth Syst. Sci. 11, 739–752. doi:10.1029/2001WR001056

Pegram G., X. Llort and D. Sempere-Torres, 2011: Radar-rainfall: separating signal and noise fields to generate meaningful ensembles. Atmospheric Research, 100, 226-236.

Pianosi, F., Raso, L., 2012. Dynamic modeling of predictive uncertainty by regression on absolute errors. Water Resour. Res. 48, 1–11. doi:10.1029/2011WR010603

Piotrowski, A., Napiórkowski, J.J., Rowiński, P.M., 2006. Flash-flood forecasting by means of neural networks and nearest neighbour approach - a comparative study. Nonlinear Process. Geophys. 13, 443–448. doi:10.5194/npg-13-443-2006

Quintero F., D. Sempere-Torres, M. Berenguer and E. Baltas, 2012: A scenario-incorporating analysis of the propagation of uncertainty to flash flood simulations, Journal of Hydrology, 460–461, 90-102.

Raftery, A.E., Gneiting, T., Balabdaoui, F., Polakowski, M., 2005, Using Bayesian Model Averaging to Calibrate Forecast Ensembles, Mon. Weather Rev. 133, 1155–1174. doi:10.1175/MWR2906.1.

René, J.R., Madsen, H., Mark, O., 2013, A Methodology for probabilistic real -time forecasting – An urban case study, Journal of Hydroinformatics, 15(3), 751-762, doi: 10.2166/hydro.2012.031

Rinehart, R. E., E. T. Garvey, 1978: Three-dimentsional storm motion detection by conventional weather radar. Nature, 273, 287-289.

Romanowicz, R., Beven, K., 1998. Dynamic real-time prediction of flood inundation probabilities. Hydrol. Sci. J. 43, 181–196. doi:10.1080/02626669809492117

Romanowicz, R., Beven, K., 2003. Estimation of flood inundation probabilities as conditioned on event inundation maps. Water Resour. Res. 39. doi:Artn 1073\rDoi 10.1029/2001wr001056

Russo, B., P. Sánchez, X. Llort, A. Rodríguez, 2017: Advanced urban flood EWS integrating radar nowcasting and 1D/2D modelling in real time**.** Proceedings of the 14th IWA/IAHR International Conference on Urban Drainage.

Sahoo, G.B., Ray, C., 2006. Flow forecasting for a Hawaii stream using rating curves and neural networks. J. Hydrol. 317, 63–80. doi:10.1016/j.jhydrol.2005.05.008

Sahoo, G.B., Ray, C., De Carlo, E.H., 2006. Use of neural network to predict flash flood and attendant water qualities of a mountainous stream on Oahu, Hawaii. J. Hydrol. 327, 525–538. doi:10.1016/j.jhydrol.2005.11.059

Salamon, P., Feyen, L., 2010. Disentangling uncertainties in distributed hydrological modeling using multiplicative error models and sequential data assimilation. Water Resour. Res. 46, 1–20. doi:10.1029/2009WR009022

Saltelli, A., Tarantola, S., and Campolongo, F. 2000. Sensitivity analysis as an ingredient of modeling. Statistical Science, 15 (4), 377-395.

Sánchez-Diezma, R., 2001: Optimización de la medida de lluvia por radar meteorológico para su aplicación hidrológica, PhD Thesis, Universitat Politècnica de Catalunya, 313 pp.

Schröter K., X. Llort, C. Velasco-Forero, M. Ostrowski and D. Sempere-Torres, 2011: Implications of radar rainfall estimates uncertainty on distributed hydrological model predictions. Atmospheric Research. ISSN: 0169-8095. Vol: 100, pp: 237-245.

Schwartz, C. S., et al., 2010: “Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership.” Wea. Forecasting, vol. 25, pp. 263 – 280.

Sene, K., Tilford, K., 2004. Review of transfer function modelling for fluvial flood forecasting. R&D Tech. Rep. W5C- 013/6/TR Defra/Environment Agency Flood Coast. Def. R&D Program. Dep. Environ. Food Rural Aff. London, UK.

Serban, P., Askew, A.J., 1991. Hydrological forecasting and updating procedures, in: IAHS pub.no.201. pp. 357–369.

Shrestha, D.L., Kayastha, N., Solomatine, D., Price, R., 2014. Encapsulation of parametric uncertainty statistics by various predictive machine learning models: MLUE method. J. Hydroinformatics 16, 95. doi:10.2166/hydro.2013.242

Sikorska, A.E., Montanari, A., Koutsoyiannis, D., 2015. Estimating the Uncertainty of Hydrological Predictions through Data-Driven Resampling Techniques. J. Hydrol. Eng. 20, A4014009. doi:10.1061/(ASCE)HE.1943-5584.0000926

Sikorska, A.E., Scheidegger, A., Banasik, K., Rieckermann, J., 2013. Considering rating curve uncertainty in water level predictions. Hydrol. Earth Syst. Sci. 17, 4415–4427. doi:10.5194/hess-17-4415-2013

Smith, E. A., Asrar, G., Furuhama, Y., Zhang, W., 2007: “International Global Precipitation Measurement (GPM) program and mission: An overview.” *Measuring Precipitation from Space: EURAINSAT and the Future*, V. Levizzani, P. Bauer, and F. J. Turk, Eds., Advances in Global Change Research Series, Vol. 28, Springer, 611–653.

Smith, P., Beven, K.J., Tawn, J.A., 2008a. Informal likelihood measures in model assessment: Theoretic development and investigation. Adv. Water Resour. 31, 1087–1100. doi:10.1016/j.advwatres.2008.04.012

Smith, P., Beven, K.J., Tawn, J.A., 2008b. Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach. Water Resour. Res. 44, 1–11. doi:10.1029/2006WR005205

Sobol', I. M. 1990. On sensitivity estimation for nonlinear mathematical models. Matematicheskoe modelirvanie, 2(1), 112-118 (in Russian), MMCE, 1(4) (1993), 407-414 (in English)

Sørensen, J.V.T., Madsen, H., Madsen, H., 2004. Data assimilation in hydrodynamic modelling: On the treatment of non-linearity and bias. Stoch. Environ. Res. Risk Assess. 18, 228–244. doi:10.1007/s00477-004-0181-9

Sorooshian, S. and A. AghaKouchak, 2010: "Advanced Concepts on Remote Sensing of Precipitation at Multiple Scales." workshop held at the University of California Irvine, March 15-17, 2010.

Stampoulis, D. and E. N. Anagnostou, 2012: "Evaluation of global satellite rainfall products over continental Europe." J. Hydrometeor, vol. 13, 588 – 603.

StormGeo (n.d.). Metocean Forecasting. Retrieved from http://www.stormgeo.com/offshore/metocean-forecasting/.

Thielen, J., Bartholmes, J., Ramos, M.-H., de Roo, A., 2009. The European Flood Alert System – Part 1: Concept and development. Hydrol. Earth Syst. Sci. 13, 125–140. doi:10.5194/hess-13-125-2009

Todini, E., 2009. Predictive uncertainty assessment in real time flood forecasting. In *Uncertainties in Environmental Modelling and Consequences for Policy Making* (pp. 205-228). Springer Netherlands.

Tolson, B.A., Shoemaker, C.A., 2008. Efficient prediction uncertainty approximation in the calibration of environmental simulation models. Water Resour. Res. 44, 1–19. doi:10.1029/2007WR005869

Tracton, M.S., Kalnay, E., 1993. Operational Ensemble Prediction at the National Meteorological Center: Practical Aspects. Weather Forecast. 8, 379–398. doi:10.1175/1520-0434(1993)008<0379:OEPATN>2.0.CO;2

UNISDR, 2009. UNISDR Terminology on Disaster Risk Reduction. Int. Strat. Disaster Reduct. doi:978-600-6937-11-3

United Nations Environment Programme Division of Early Warning and Assessment (UNEP-DEWA), "One Planet Many People: Atlas of Our Changing Environment", Nairobi, Kenya, 2006.

United Nations, Global Survey of Early Warning Systems, United Nations report, 2006.

van Leeuwen, P.J., 2009. Particle Filtering in Geophysical Systems. Mon. Weather Rev. 137, 4089–4114. doi:10.1175/2009MWR2835.1

Verkade, J, S., Brown, J, D., Reggiani, P., Weerts, A, H., 2013. Post-processing ECMWF precipitation and temperature ensemble reforecasts for operational hydrologic forecasting at various spatial scales. Journal of Hydrology, 501, 73-91.

Vestergaard, L. F., 2011. *Analyse af oversvømmeslesmønstret ved havvandsstigning i Køge Bugt ud for Greve*. Note to agenda item for meeting in 'Teknik og Miljøudvalget', Copenhagen Municipality. Retrieved from http://www.gsforsyning.dk/sites/default/files/havstrategi_politiskvedtaegt_bilag.pdf.

Vrugt, J. a, Robinson, B.A., 2007a. Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging. Water Resour. Res. 43, 1–15. doi:10.1029/2005WR004838

Vrugt, J. a, Robinson, B. a, 2007b. Improved evolutionary optimization from genetically adaptive multimethod search. Proc. Natl. Acad. Sci. 104, 708–711. doi:10.1073/pnas.0610471104

Vrugt, J.A., Gupta, H. V., Nualláin, B., Bouten, W., 2006. Real-Time Data Assimilation for Operational Ensemble Streamflow Forecasting. J. Hydrometeorol. 7, 548–565. doi:10.1175/JHM504.1

Vrugt, J.A., ter Braak, C.J.F., Gupta, H. V., Robinson, B.A., 2009. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stoch. Environ. Res. Risk Assess. 23, 1011–1026. doi:10.1007/s00477-008-0274-y

Wagener, T., M. J. Lees, H.S. Wheater, MCAT v.3 - Monte-Carlo Analysis Toolbox User Manual, Imperial College of Science Technology and Medicine, 2001.

Walker, J.P., Willgoose, G.R., Kalma, J.D., 2001. One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: A comparison of retrieval algorithms. Adv. Water Resour. 24, 631–650. doi:10.1016/S0309-1708(00)00043-9

Weerts, A. H., Winsemius, H. C., & Verkade, J. S., 2011. Estimation of predictive hydrological uncertainty using quantile regression: examples from the National Flood Forecasting System (England and Wales). Hydrology and Earth System Sciences, 15, 255-265.

Weerts, A.H., El Serafy, G.Y.H., 2006. Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. Water Resour. Res. 42, 1–17. doi:10.1029/2005WR004093

Wilks, D. S., 1995: Statistical Methods in the Atmospheric Sciences. Academic Press, 467 pp.

Wilson, J. W., 1966: Movement and predictability of radar echoes. ERTM-NSSL-28., National Severe Storms Laboratory, 30 pp.

Wilson, J. W., Y. Feng, M. Chen, R. D. Roberts, 2010: Nowcasting Challenges during the Beijing Olympics: Successes, Failures, and Implications for Future Nowcasting Systems. Weather and Forecasting, 25, 1691-1714.

WMO, 2015. Synthesis of the Status and Trends with the Development of Early Warning Systems. Background Paper prepared for the Global Assessment Report on Disaster Risk Reduction 2015. Geneva.

Xie, X., Zhang, D., 2010. Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. Adv. Water Resour. 33, 678–690. doi:10.1016/j.advwatres.2010.03.012

Yeh, H.D., Huang, Y.C., 2005. Parameter estimation for leaky aquifers using the extended Kalman filter, and considering model and data measurement uncertainties. J. Hydrol. 302, 28–45. doi:10.1016/j.jhydrol.2004.06.035

Young, P.C., 2002. Advances in real-time flood forecasting. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 360, 1433–1450. doi:10.1098/rsta.2002.1008

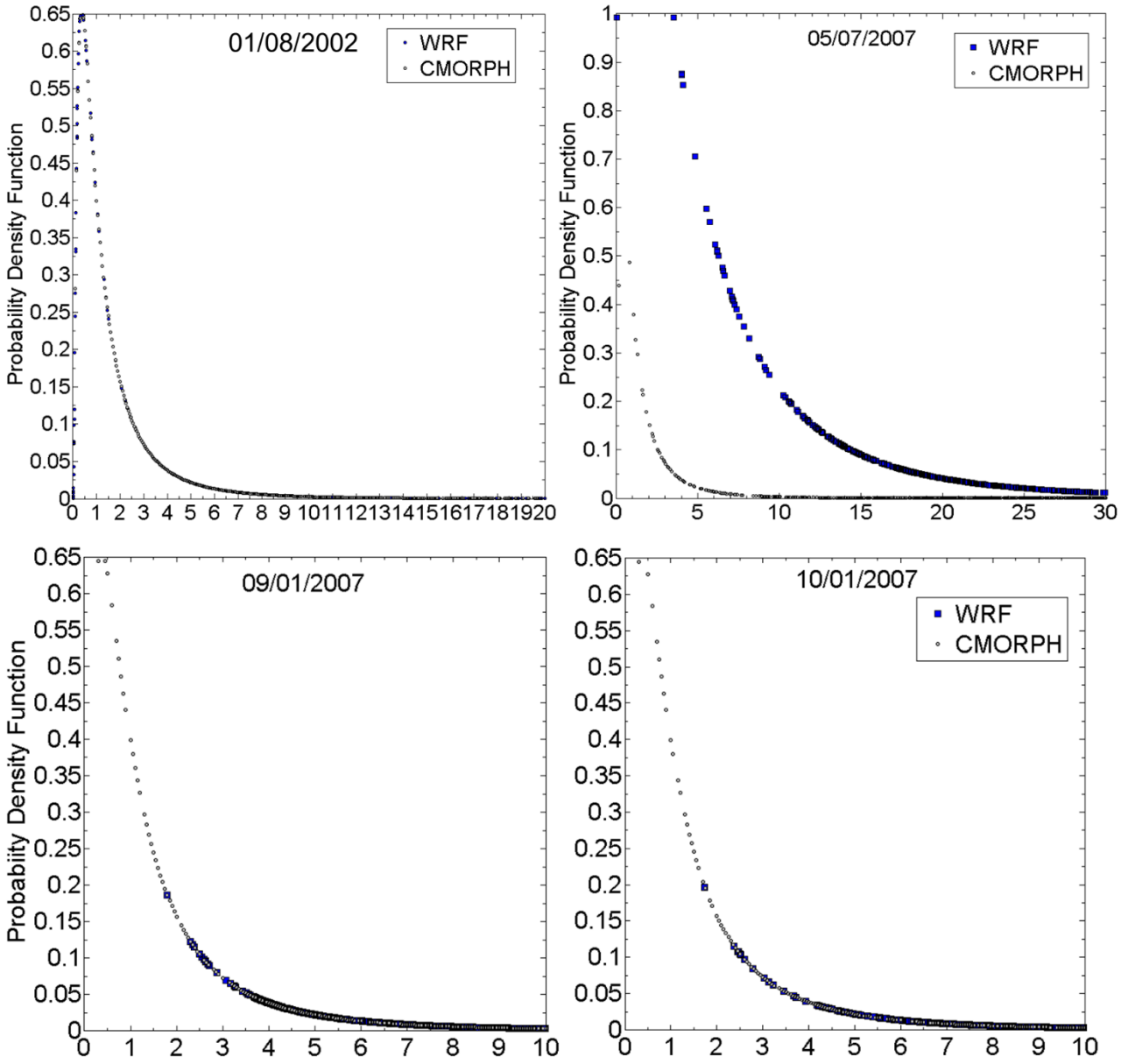
Zawadzki, I., 1984: Factors affecting the precision of radar measurement of rain. 22th Conference on Radar Meteorology, Zurich (Switzerland), 251-256.

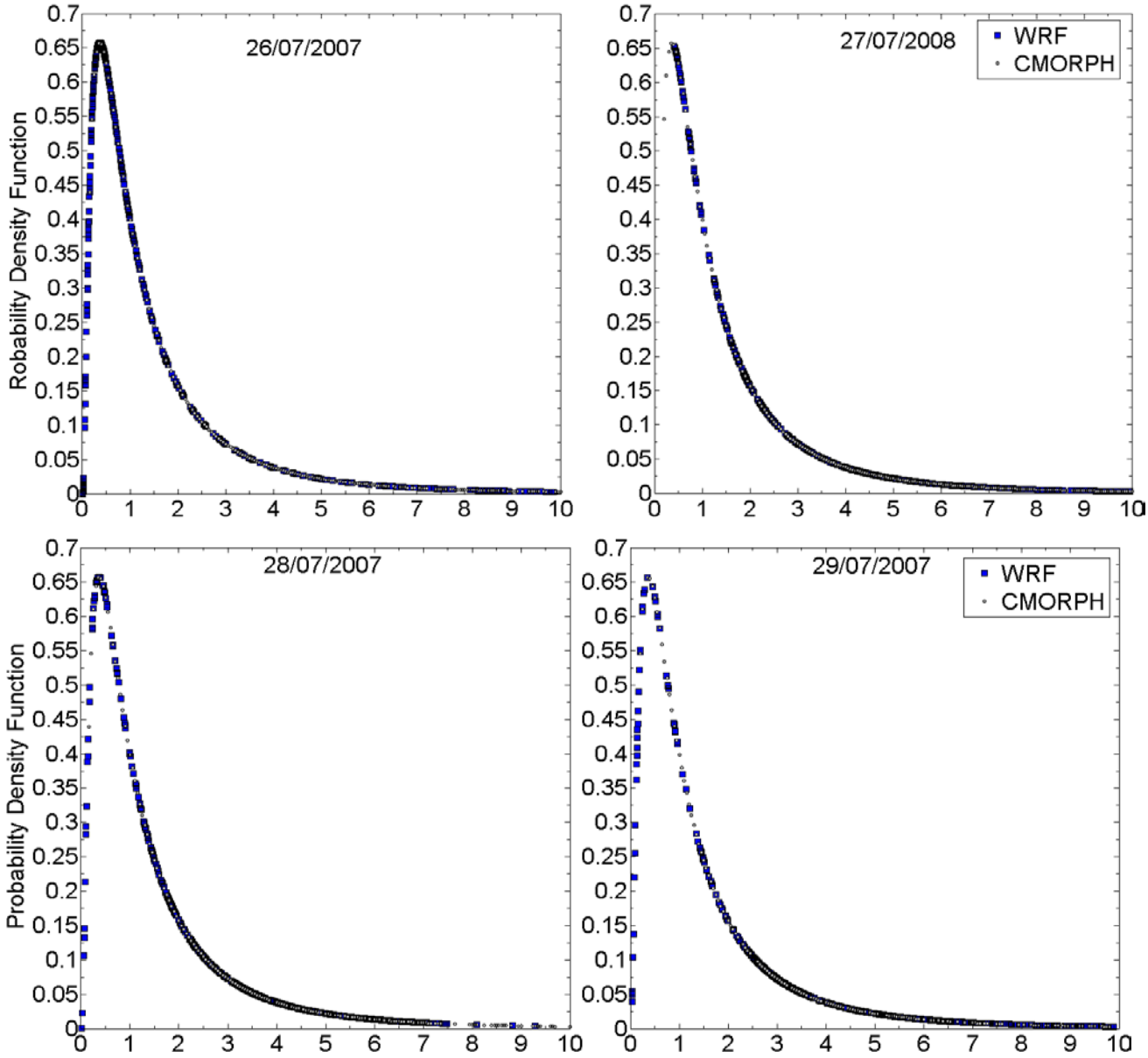
Zhang, X., E. N. Anagnostou, M. Frediani, S. Solomos, and G. Kallos, 2013: “Using NWP Simulations in Satellite Rainfall Estimation of Heavy Precipitation Events over Mountainous Areas. Journal of Hydrometeorology, vol. 14, pp. 1844 – 1858.

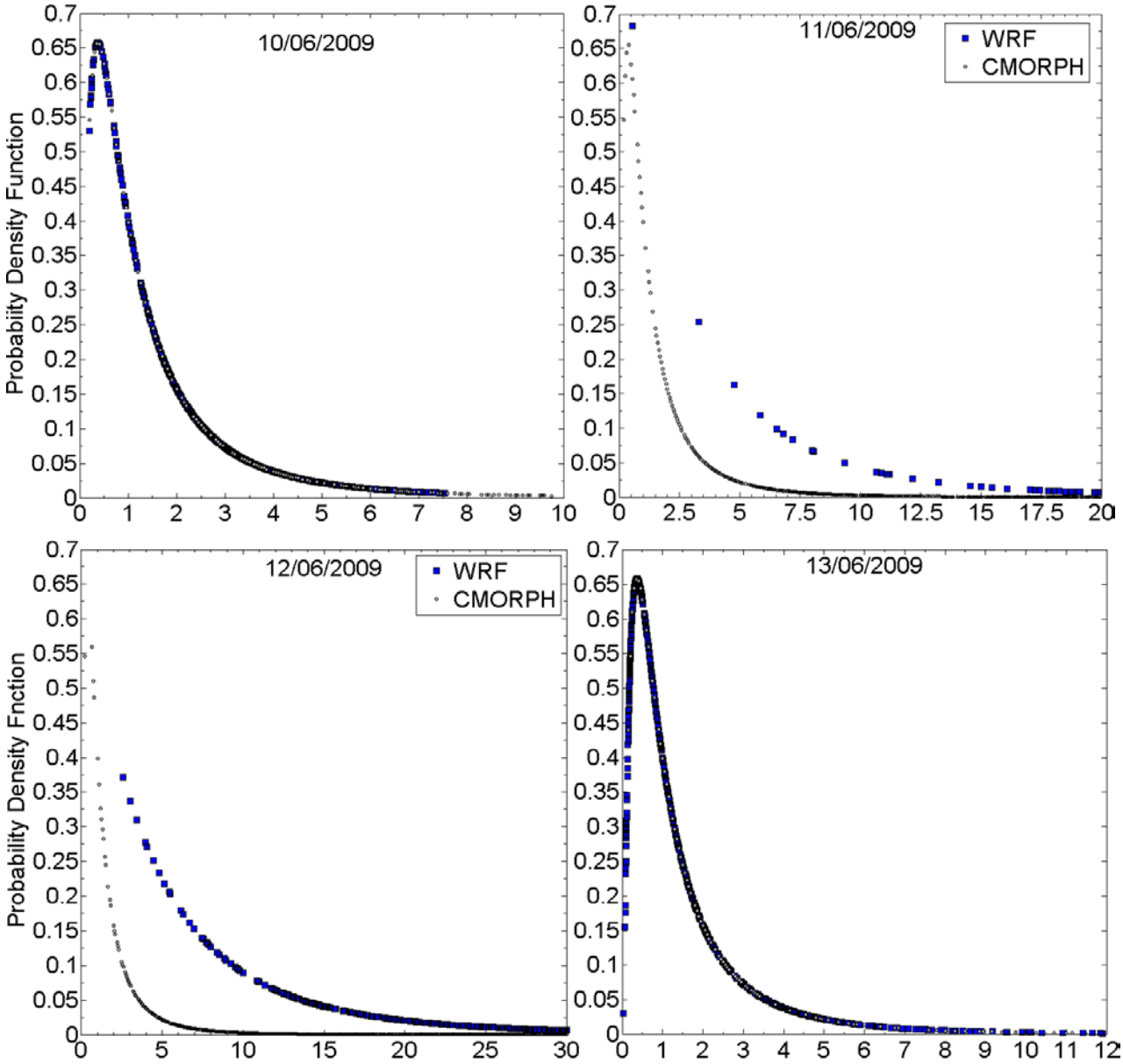
Zoccatelli, D., M. Borga, F. Zanon, B. Antonescu, and G. Stancalie, 2010: “Which rainfall spatial information for flash flood response modelling? A numerical investigation based on data from the Carpathian range. Romania. J. Hydrol., vol. 394, pp. 148 – 161.

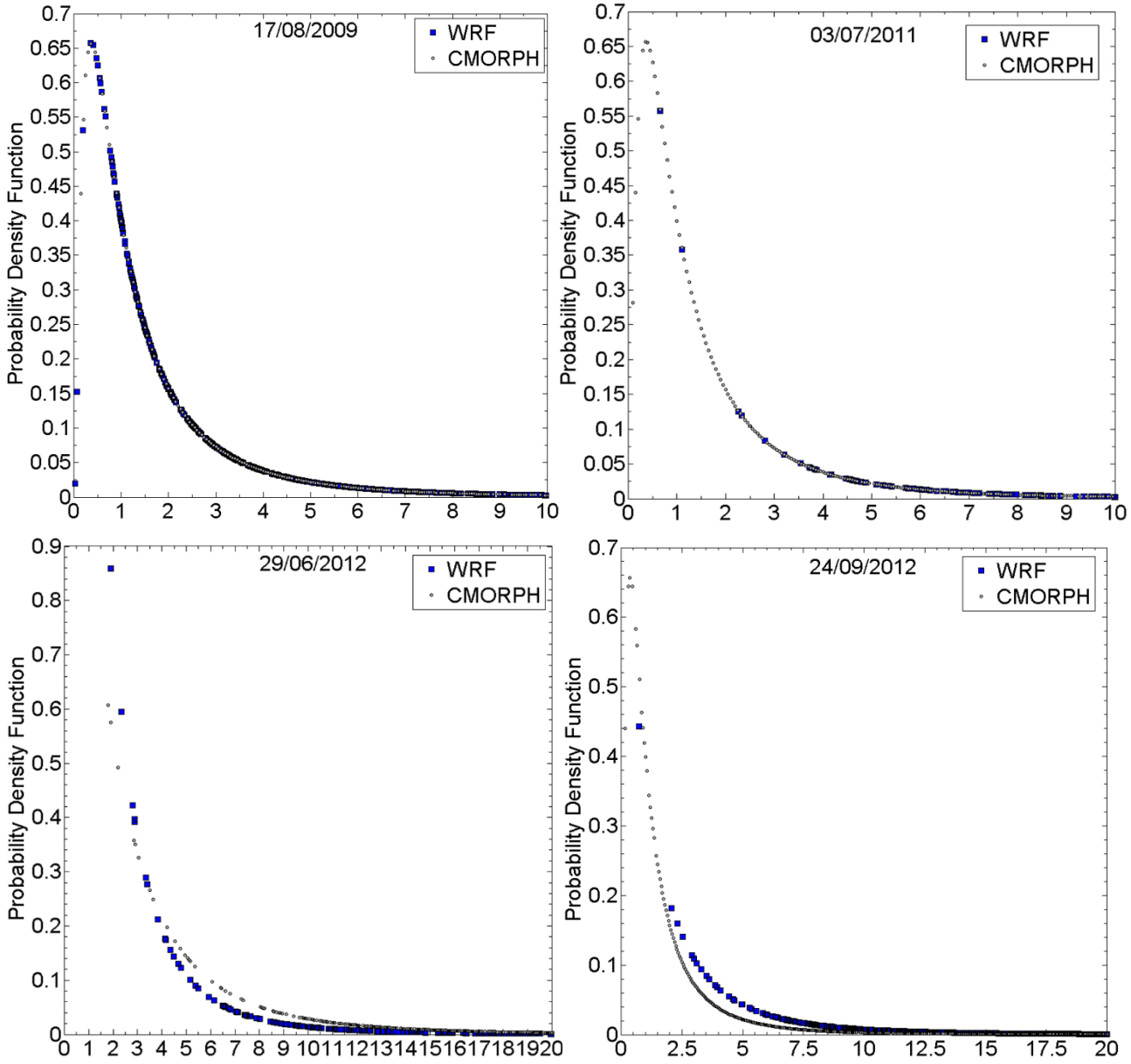
Annex 1

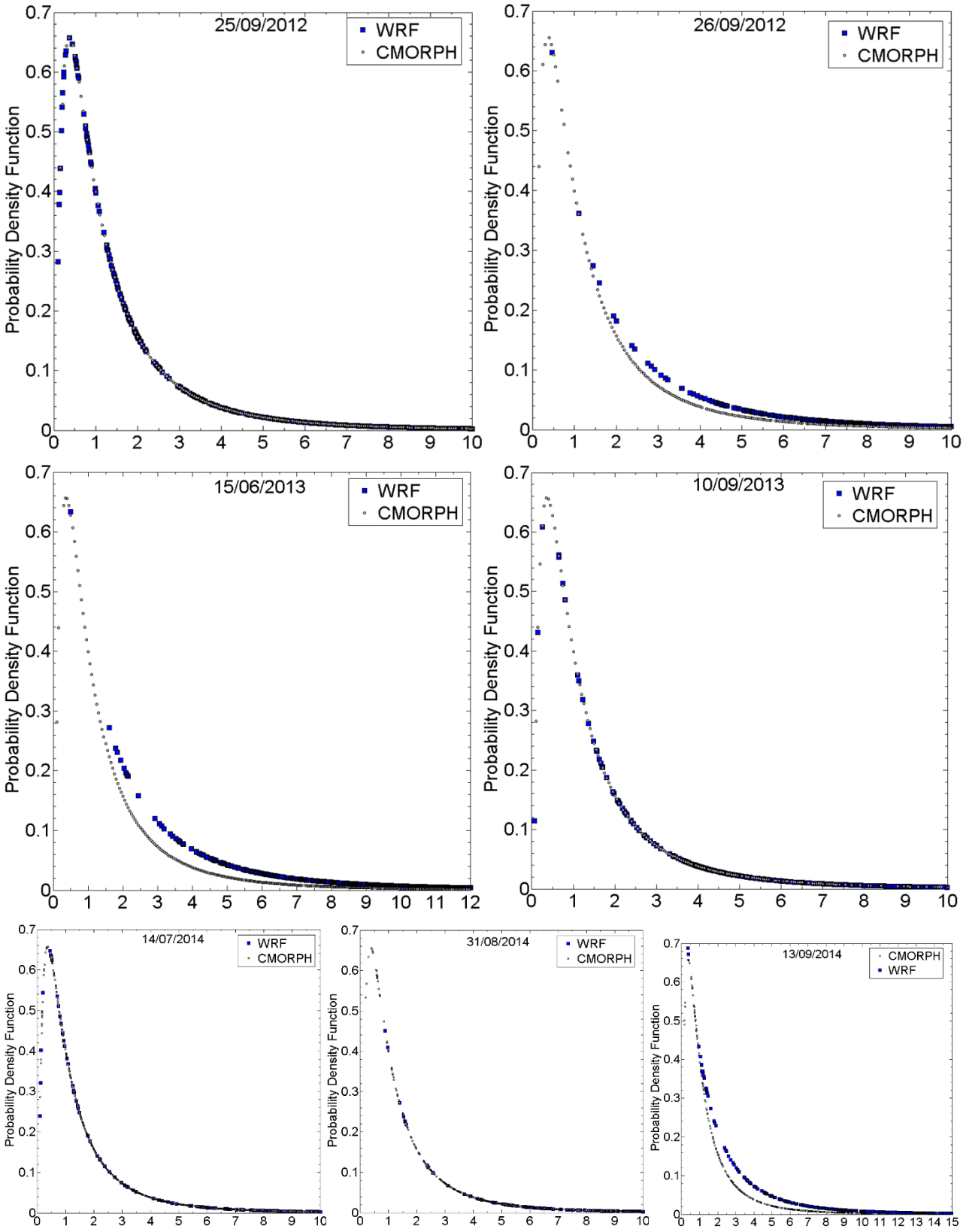
Probability density function (PDF) plots of the WRF precipitation estimations compared to CMORPH observations

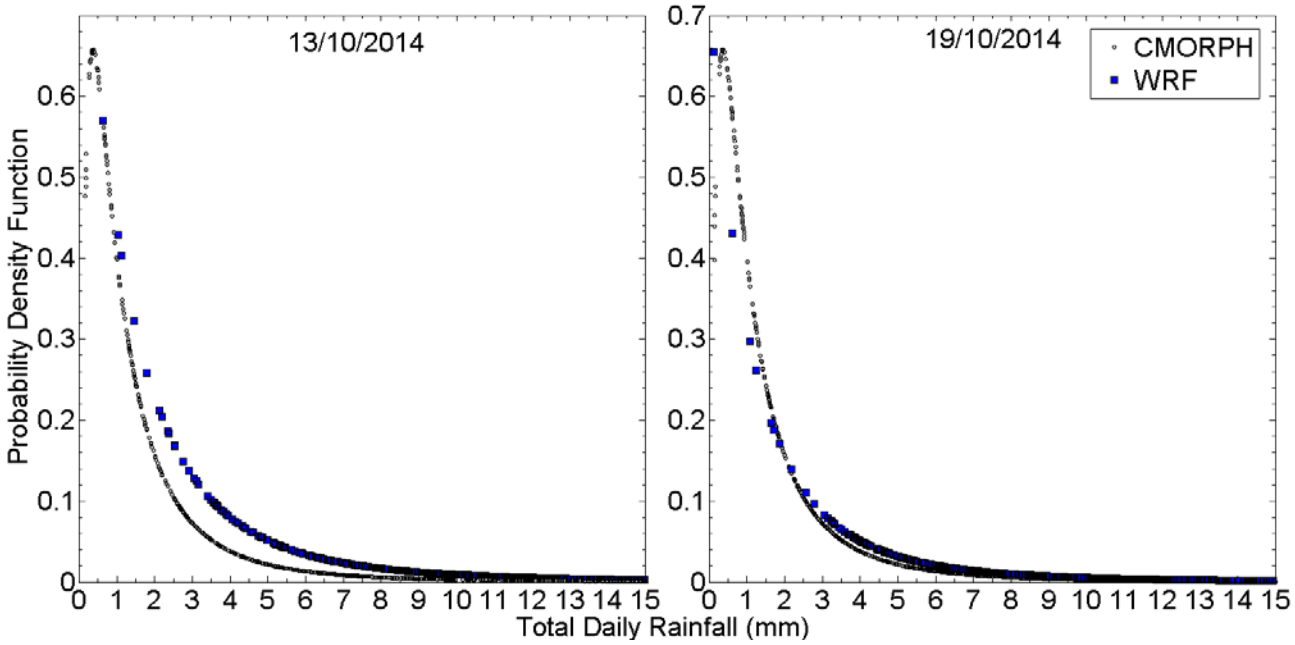



26/07/2007
27/07/2008
28/07/2007
29/07/2007
WRF
CMORPH
Robability Density Function
Probability Density Function


10/06/2009
11/06/2009
12/06/2009
13/06/2009
WRF
CMORPH
Probability Density Function
Probability Density Fnction


17/08/2009
WRF
CMORPH
Probability Density Function
03/07/2011
WRF
CMORPH
Probability Density Function
29/06/2012
WRF
CMORPH
Probability Density Function
24/09/2012
WRF
CMORPH
Probability Density Function


25/09/2012
WRF
CMORPH
Probability Density Function
26/09/2012
WRF
CMORPH
Probability Density Function
15/06/2013
WRF
CMORPH
Probability Density Function
10/09/2013
WRF
CMORPH
Probability Density Function
14/07/2014
WRF
CMORPH
Probability Density Function
31/08/2014
WRF
CMORPH
Probability Density Function
13/09/2014
CMORPH
WRF
Probability Density Function


13/10/2014
19/10/2014
CMORPH
WRF
Probability Density Function
Total Daily Rainfall (mm)