Probabilistic forecasting for urban water management: A case study

Jeanne-Rose Rene'[1] , Henrik Madsen[2], Ole Mark[3]

[1] DHI, Denmark, jerr@dhigroup.com
[2] DHI, Denmark, hem@dhigroup.com
[3] DHI, Denmark, Ole.Mark@dhigroup.com

ABSTRACT

This paper illustrates the application of a probabilistic approach for the estimation of the uncertainty in rainfall forecast from a numerical weather prediction model in combination with a 1D/2D hydrodynamic model for producing probabilistic flood forecasts. The approach quantifies the uncertainty conditioned on the rainfall forecast in the form of probability distribution functions. The method utilized in this paper involves a retrospective comparison at different lead times between archived forecasted rainfall and its corresponding observed rainfall for the second largest city in Denmark, Aarhus. Since there were no large events on record to generate flooding, a synthetic forecast event is used for illustration of the method. The Latin hypercube sampling technique was used to generate ensembles of rainfall for the synthetic rainfall forecast which has been used in conjunction with the 1D/2D hydrodynamic model. For comparison, a direct quantile approach was used to generate rainfall quantiles which were also ingested into the 1D/2D model to enable the selection of a robust approach that can be used in real time.

KEYWORDS

Flood forecasting, numerical weather prediction model, probabilistic rainfall forecasting, real-time modelling, urban flooding

1 INTRODUCTION

It has generally been acknowledged by most literature sources in the field of water management that flood forecasts should be made with a quantifiable estimate of the uncertainty in the forecast (Krzysztofowicz, 2001, Pappenberger and Beven, 2006). For this reason, many studies are moving away from the conventional deterministic approach and towards probabilistic approaches. However, most of the approaches are applicable at the river basin scale and less so at the urban scale. One of the reasons for this is that hydrological and hydrodynamic modelling at the urban scale requires rainfall forecast of high temporal and spatial resolution. A few studies have attempted to assess the feasibility of using quantitative rainfall forecasts as well as probabilistic forecasting schemes in combination with a 1D sewer model in urban flood modelling (Rico-Ramirez et al., 2009, Schellart et al., 2009, Liguori et al., 2012).

More recently, with the increased frequency of floods due to climate change and rapid urbanization, more studies are trying to increase lead-time by investigating the feasibility of using different rainfall forecast products and different approaches for downscaling the rainfall forecast for forecasting flows and water levels in urban drainage system (Schellart et al., 2011, Simões et al., 2011a, Simões et al., 2011b). However, in order to achieve the desired level of accuracy in an operational context of the presented approaches, further investigation must be carried out using a combination of approaches and techniques.

This paper proposes a new approach for probabilistic flood forecasting in urban areas. The approach comprises of three components: (1) estimation of a probabilistic rainfall forecast model which is based on a retrospective comparison of archived historical forecasted rainfall and its corresponding observed rainfall; (2) prediction of rainfall quantiles or rainfall ensembles based on the stochastic model of the rainfall forecast; and (3) prediction of probabilistic flood maps using the probabilistic rainfall forecasts as input for the physically-based 1D/2D hydrodynamic model. The proposed approach is tested using a synthetic extreme rainfall event for a case study in Aarhus, Denmark. Two methods for the generation of probabilistic rainfall forecasts; Latin Hypercube Sampling (LHS) approach and direct quantile approach are compared.

2 METHODOLOGY

2.1 Rainfall data

In this research, two years of continuous hourly data (2009 – 2010) for observed and forecasted rainfall was used for the city of Aarhus, Denmark. The observed data originated from a network of 3 tipping bucket rain gauges (Figure 1) installed in the catchment. The data had an original temporal resolution of 1 minute. The Thiessen polygons method was used for estimating catchment rainfall based on the rain gauge data (Figure 1). The forecasted rainfall data originated from a numerical weather prediction model (StormGeo, 2011) and was used as the source of historical rainfall forecast. A 72 hour, hourly product of (6.2 x 11.1) km resolution updated every 12 hours was used. The data provided covered two rainfall forecasts grids, which fell directly over the study area (Figure 1). Observed and forecasted rainfall data used for estimation of the stochastic rainfall model are shown in Figure 1.

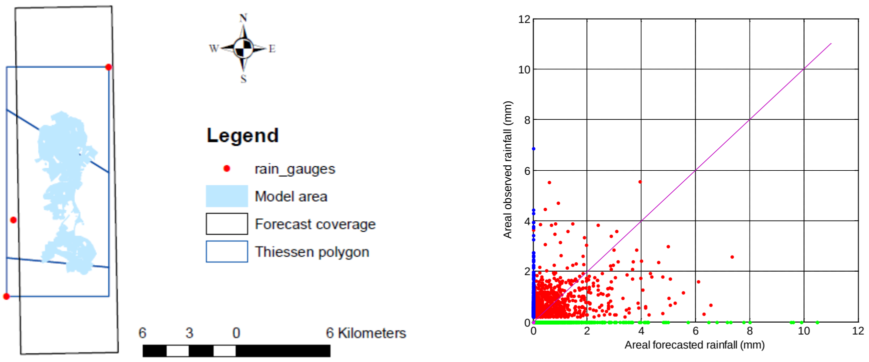

Figure 1: Left: Model area, total forecast area used, Thiessen polygons and rain gauge locations. Right: observed and forecasted hourly areal rainfall for a forecast lead time of 12 hours.

2.2 Stochastic model

This study utilizes the probability distributions obtained from the retrospective comparison of observed rainfall and its corresponding forecasts as described in (René et al., 2012), denoted S and $\hat{S}$, respectively. The method involves decomposing the observed and corresponding forecast data into lead times and approximating the error conditioned on the rainfall forecast in the form of probability distribution functions. These functions are then imposed on the forecasted rainfall value for the corresponding lead time to determine the probability of rainfall given a rainfall forecast, i.e.:

$$\mathrm{P}(\text{rainfall} \mid \text{rainfall forecast})$$

Two stochastic models are defined to reflect the intermittent nature of rainfall. One model comprises of lognormal probability distributions for each lead time when the rainfall forecast is zero. The probability conditioned on a zero rainfall forecast is given by:

$$\mathrm{P}(\mathrm{S} \le x \mid \hat{\mathrm{S}} = 0) = \mathrm{P}(\mathrm{S} \le x \mid \mathrm{S} > 0, \hat{\mathrm{S}} = 0)\mathrm{P}(\mathrm{S} > 0 \mid \hat{\mathrm{S}} = 0) + \mathrm{P}(\mathrm{S} = 0 \mid \hat{\mathrm{S}} = 0) \quad (1)$$

where $\mathrm{P}(\mathrm{S} \le x \mid \mathrm{S} > 0, \hat{\mathrm{S}} = 0)$ is obtained by fitting the data to a lognormal distribution, and the conditional probabilities $\mathrm{P}(\mathrm{S} > 0 \mid \hat{\mathrm{S}} = 0)$ and $\mathrm{P}(\mathrm{S} = 0 \mid \hat{\mathrm{S}} = 0)$ are obtained from the data.

The other stochastic model comprises of probability distributions for each lead time for a rainfall forecast more than zero, denoted $x*$, but in this case for data transformed from its original domain to the normal domain using the Box-Cox transformation method. The probability conditioned on a non-zero rainfall forecast is given by:

$$\mathrm{P}(\mathrm{S} \le x \mid \hat{\mathrm{S}} = x*) = \mathrm{P}(\mathrm{S} \le x \mid \mathrm{S} > 0, \hat{\mathrm{S}} = x*)\ \mathrm{P}(\mathrm{S} > 0 \mid \hat{\mathrm{S}} = x*) + \mathrm{P}(\mathrm{S} = 0 \mid \hat{\mathrm{S}} = x*) \quad (2)$$

where $\mathrm{P}(\mathrm{S} \le x \mid \mathrm{S} > 0, \hat{\mathrm{S}} = x*)$ is obtained by fitting the transformed data to a bivariate normal distribution, the conditional probability $\hat{\mathrm{P}}(\mathrm{S} = 0 \mid \hat{\mathrm{S}} = x^*)$ is estimated from data by fitting a functional relationship of the form:

$$\hat{\mathrm{P}}(\mathrm{S} = 0 \mid \hat{\mathrm{S}} = x^*) = a \exp(bx*) + c \quad (3)$$

and:

$$\mathrm{P}(\mathrm{S} > 0 \mid \hat{\mathrm{S}} = x*) = 1 - \mathrm{P}(\mathrm{S} = 0 \mid \hat{\mathrm{S}} = x*) \quad (4)$$

Using these relationships, the probability distribution of a rainfall conditioned on a rainfall forecast can be found.

2.3 Urban hydrodynamic 1D/2D model

A 1D/2D hydrodynamic MIKE URBAN (DHI, 2011) model is used in this study. The runoff computations are performed using a simple time-area model. The runoff hydrographs are generated and subsequently used as hydraulic loads in the pipe network which overflows onto the 2D surface model once the pipe system becomes surcharged.

The sewer system is a combined system carrying storm water run-off as well as industrial and domestic waste water. The system has been modeled using 1985 manholes, 1722 circular pipes, 184 weirs, 83 basins and 26 pumps. The sewer model has been calibrated by the municipality and is considered to be fit for the purpose for this research by being able to produce realistic flood maps.

The sewer model was coupled with a 2D surface model of 1.6m resolution digital terrain model (DTM) – no buildings included. A calculation grid cell size of 10 x 10 m was selected.

2.4 Experimental setup

The probabilistic forecast model was tested using a synthetic rainfall forecast event. The event was generated by multiplying a 12-hour rainfall forecast on record by a factor of 10 to generate a 12-hour event with the accumulated rainfall of the 95th percentile of the rainfall equivalent to a 100-year event. This approach was selected in order to generate flooding.

2.5 Generation of probabilistic rainfall forecast

Using the established stochastic model, the rainfall forecast for input into the hydrodynamic model was generated using the LHS approach and the direct quantile approach. This is done by imposing the 12 hour-hourly deterministic rainfall forecast from the NWP model on the stochastic model and then the two approaches are used to generate rainfall ensembles and percentiles of rainfall forecasts respectively.

LHS Approach - For the synthetic 12-hour rainfall forecast, the LHS approach was used to generate an ensemble of 12-hour hourly rainfall forecasts. In this case an ensemble size of 50 was used. Each hour in the 12 hour forecast has a corresponding rainfall probability distribution. The approach samples 50 times from each distribution resulting in rainfall forecasts of equal probability. The ensemble forecasts were computed based on the assumption of complete temporal dependence. This implies that the ensemble is generated by pairing values from the same sampling interval across lead-times to generate the 12 hour-hourly time series of rainfall forecasts.

Direct Quantile Approach - This approach uses percentiles of rainfall extracted from the rainfall probability distributions as input to the 1D/2D hydrodynamic model. For a given 12 hour - hourly rainfall forecast, each lead-time has a corresponding probability distribution. The exact value corresponding to a selected quantile can be extracted from each probability distribution (lead-time) to generate a 12 hour - hourly time series. This time series is approximately equal to the corresponding percentile from the rainfall ensembles obtained using the LHS approach.

2.6 2D Result processing

In order to present flood maps with an estimation of the uncertainty for the LHS approach, the results from each computational grid cell from the 2D computational domain for each rainfall ensemble member are used to compute different percentiles. In this paper the 50th and 95th percentile is considered.

Consider a 2D overland model with horizontal and vertical extents $J\Delta x$ and $K\Delta y$ respectively. The computational grid is divided into individual cells each of dimension Δx x Δy for the 2D computation (Figure 2). For each rainfall ensemble, the maximum water levels in each cell are obtained. The results from the rainfall ensembles corresponding to each cell are then used to compute the selected percentiles for that cell.

In the direct quantile approach the 50th and 95th percentile flood maps are obtained directly from the simulations of the 50th and 95th rainfall forecast percentiles.

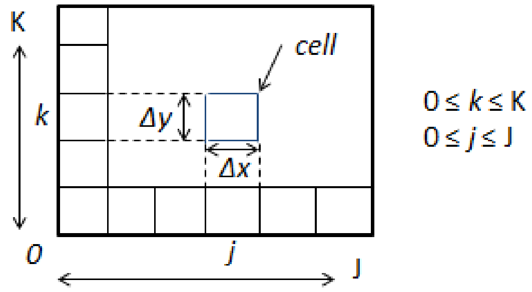

Figure 2: Schematic representation of 2D computational domain for the 2D overland model

3 RESULTS AND DISCCUSION

Rainfall ensembles from the stochastic model

Initial analysis of the data presented in Figure 1 shows that there is a tendency that the forecast overestimates for large events and underestimates for smaller events. Thus, when using the stochastic rainfall model on a large rainfall forecast the model will compensate for this overestimation and reduce the rainfall. The bias correction becomes quite severe for the extreme synthetic rainfall event used in this study (see Figure 3). It should be noted, however, that the synthetic rainfall event is more than 3 times larger than the largest event used for estimation of the stochastic model, and hence the model has not been validated for such events.

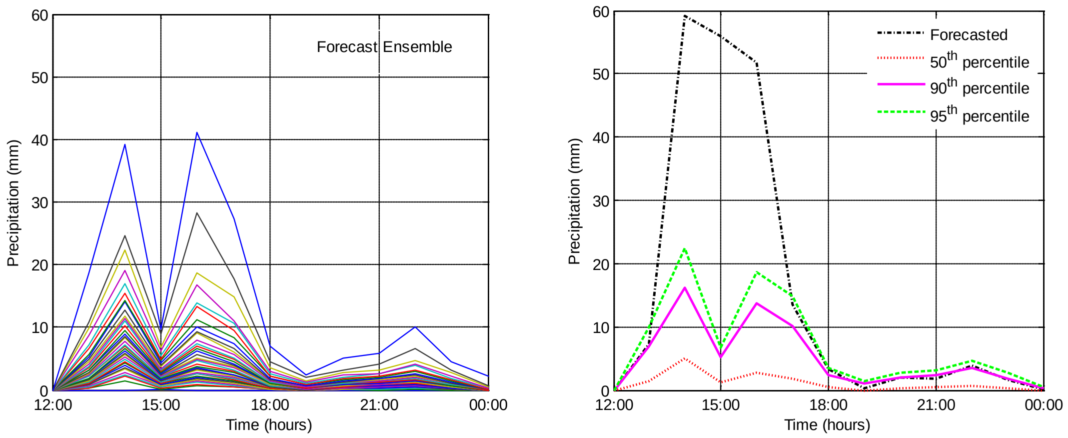

Figure 3: Forecast ensemble from probabilistic rainfall model, forecasted 12-hour rainfall and forecast percentiles estimated from the forecast ensemble

1D/2D hydrodynamic model simulation results

For a selected area in the model domain, flood maps for the 50th and 95th percentile are presented in Figure 4-Figure 7.

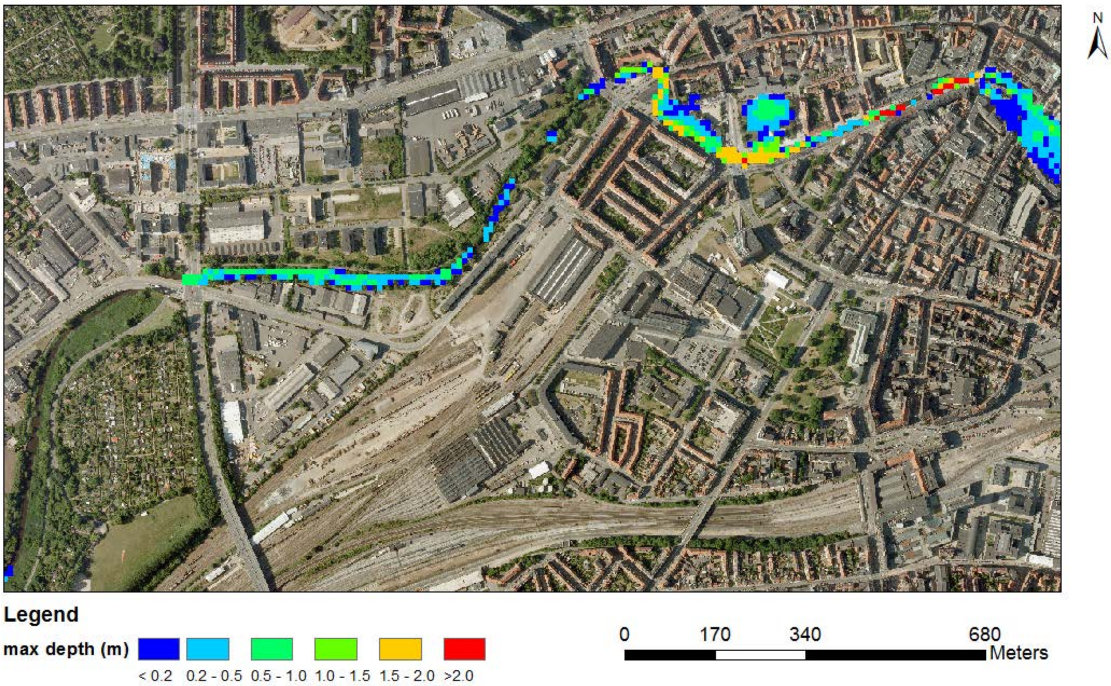

Figure 4: 50th Percentile of the maximum flood depths obtained from the rainfall ensembles (LHS approach)

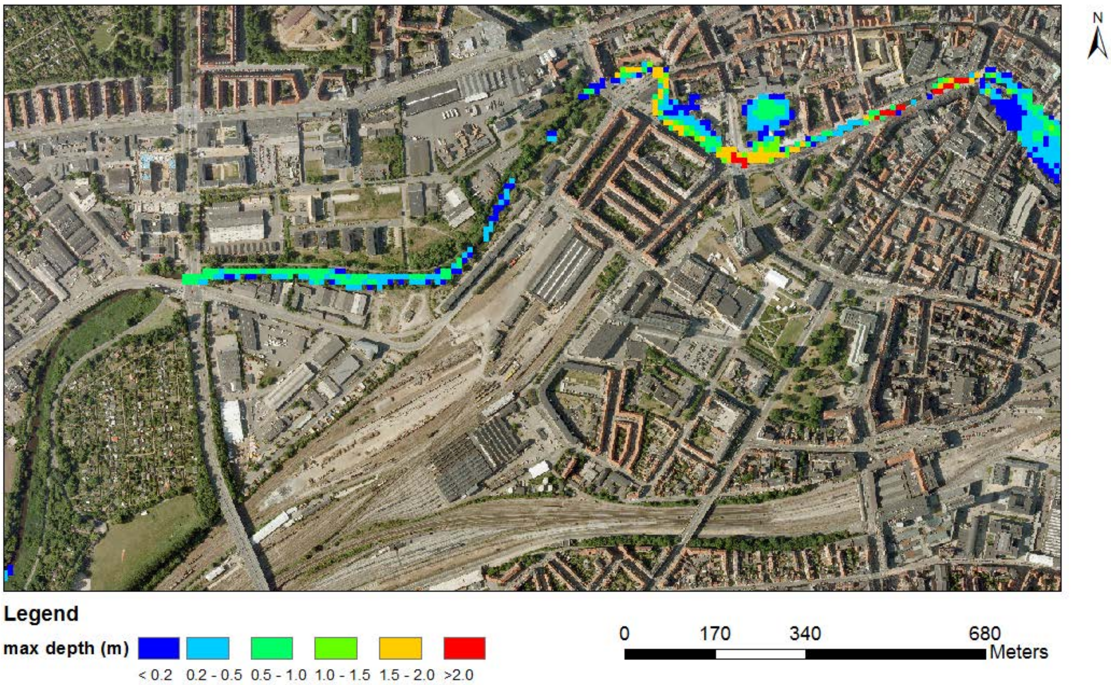

Figure 5: Maximum depth obtained when using the 50th percentile of the rainfall probability distribution (direct quantile approach)

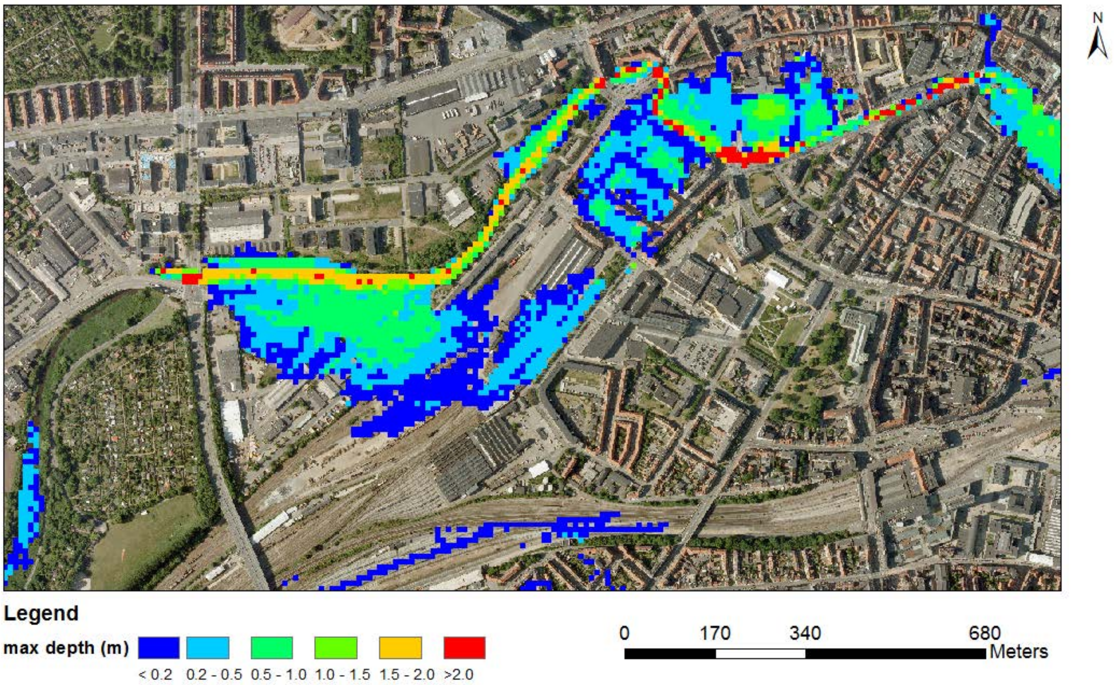

Figure 6: 95th Percentile of the maximum flood depths obtained from the rainfall ensembles (LHS approach)

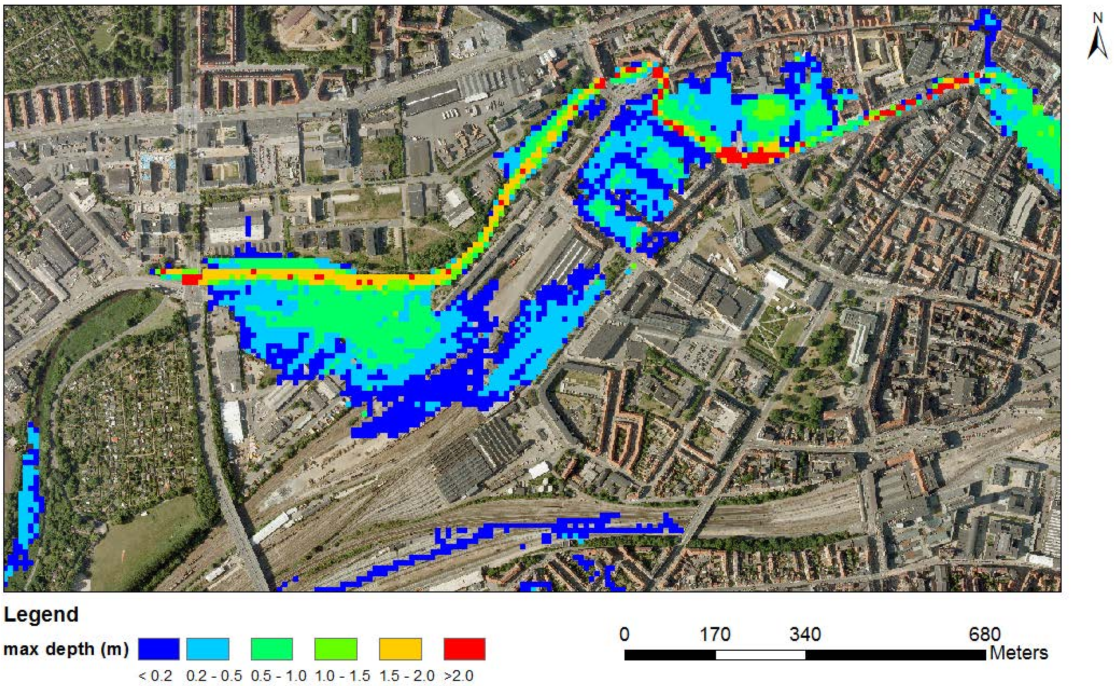

Figure 7: Maximum depth obtained when using the 95th percentile of the rainfall probability distribution (direct quantile approach)

Comparison of 50^{th} and 95^{th} percentiles of maximum flood depths over each computational grid cell over the entire model domain for the LHS approach and the direct quantile approach shows that the methods give approximately the same results as shown in the histogram plot in Figure 8.

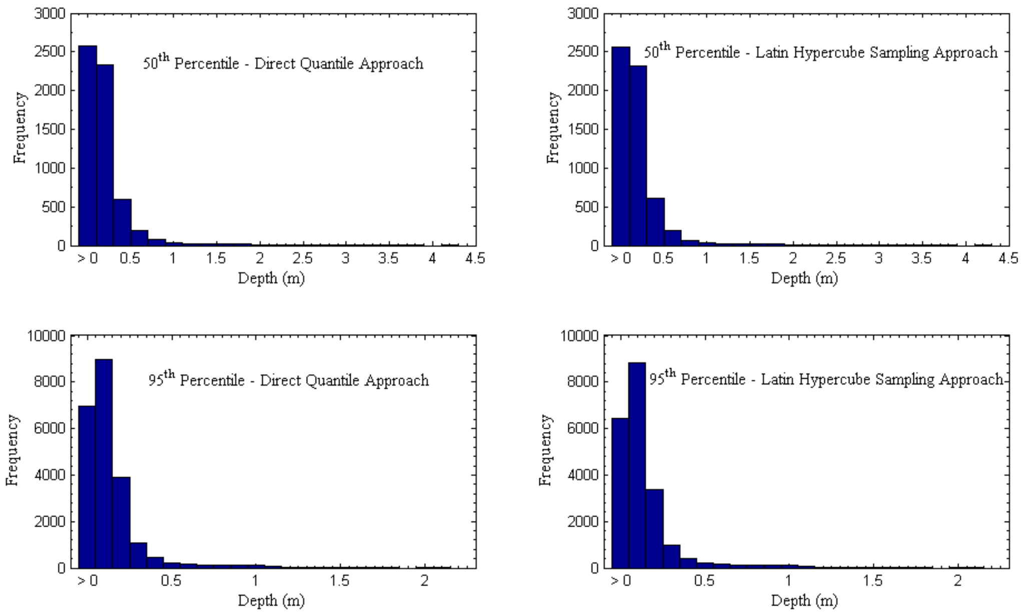

Figure 8: Comparison of the 50^{th} and 95^{th} percentile of the maximum flood depth obtained using the direct quantile approach and the Latin hypercube sampling approach. The data represents the maximum flood depth obtained for each grid cell over the model domain

From the figure it can be observed that both methods generally show similarity in producing maximum flood depth distribution. Although the total number of wet cells is not the same as shown in Table 1, the difference is very small.

Table 1: Number of wet cells obtained for each percentile using LHS and direct quantile approach

	Number of wet cells	
Percentile	Direct Quantile Approach	Latin Hypercube Sampling
50^{th}	5928	5905
95^{th}	22323	20971

The difference in maximum water levels between the methods can be visualised using a scatter plot as presented in Figure 9.

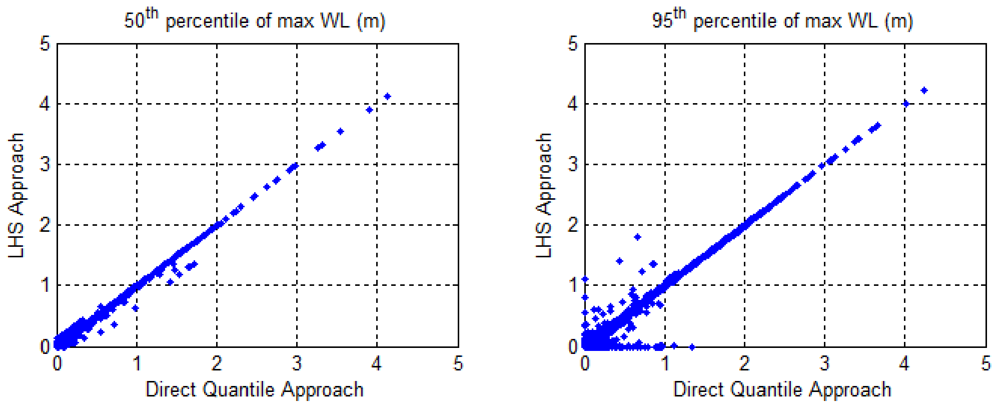

Figure 9: Maximum flood depth obtained for each grid cell over the model domain

It is observed that in some instances the direct quantile approach computes water on the surface whereas the LHS doesn't and vice versa, but more so for the direct quantile approach. This is observed for only a small quantity of grid cells.

The methods are also compared by viewing time series plots for a selected grid cell i.e. a specific location in the model (see Figure 10). These results also confirm similarity in the methods for simulation of the temporal development of the flooding.

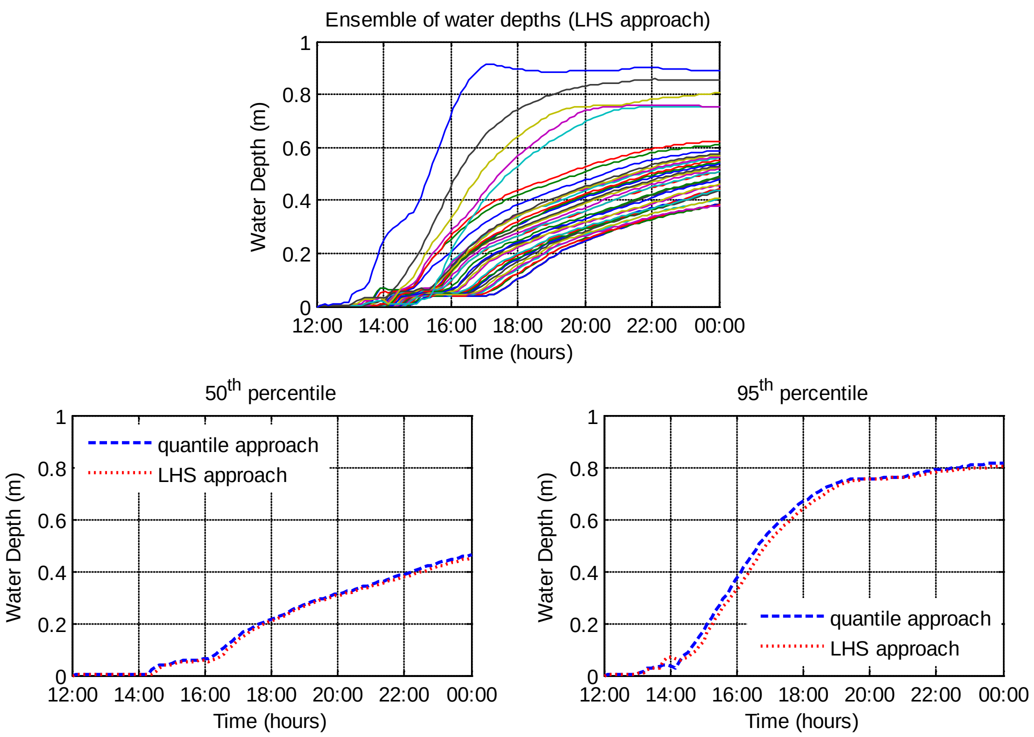

Figure 10: Comparison of time series of water level at a selected location in the urban flood model

The results show that the direct quantile approach provides an accurate estimate of the probabilistic flood maps. The direct quantile approach only requires a few model simulations compared to a series of simulations when using the LHS approach. The efficiency in computational time makes the direct quantile approach feasible for real-time application.

The large difference between flood extents maps between the two percentiles in Figure 4- Figure 7 highlight the relevance of making probabilistic forecasts. Because the actual observed event is not known, then conclusions concerning accuracy of flood depths and extents cannot be made. However, the results reiterate the relevance of probabilistic flood forecasting and the potential in using the direct quantile approach.

4 CONCLUSION

The applicability of an approach for probabilistic urban water management has been presented in this paper. Two methods for generating probabilistic rainfall forecasts from probability distribution functions are compared for the selection of an efficient approach which can be applied in real-time. The results from the LHS approach and the direct quantile approach show similarities in many ways and as a result the direct quantile approach proved to be the most attractive for real-time application.

The results obtained from the case study looks very promising for use operationally for 1D/2D models in conjunction with deterministic quantitative rainfall forecasts. The approach provides the opportunity for decision makers to make better-informed decisions by providing them with confidence levels in the flood forecast. This has been found to be an important feature for risk assessment, warning and evacuation. It is clear that the rainfall forecast has a large contribution to the uncertainty in the flood forecast. Moreover, in order for this approach to be successful, it requires diligent collection of both observed and forecasted rainfall data.

5 ACKNOWLEDGEMENT

This study was supported with great support from Municipality of Aarhus, Denmark.

6 REFERENCES

DHI (2011) MIKE URBAN, 2011 edn, Copenhagen.

Krzysztofowicz R (2001) The case for probabilistic forecasting in hydrology. Journal of Hydrology 249: 2-9 DOI 10.1016/s0022-1694(01)00420-6

Liguori S, Rico-Ramirez MA, Schellart ANA, Saul AJ (2012) Using probabilistic radar rainfall nowcasts and NWP forecasts for flow prediction in urban catchments. Atmospheric Research 103: 80-95 DOI 10.1016/j.atmosres.2011.05.004

Pappenberger F, Beven KJ (2006) Ignornace is Bliss:Or seven reasons not to use uncertainty analysis. Water Resources Research 42 DOI 10.1029/2005WR004820

René J-R, Madsen H, Mark O (2012) A Methodology for Probabilistic Real -Time Forecasting – An urban Case Study. Journal of Hydroinformatics submitted for review

Rico-Ramirez MA, Schellart ANA, Liguori S, Saul AJ (2009) Quantitative precipitation forecasting for a small urban area: Use of a high-resolution numerical weather prediction model. Paper presented at the 8th International workshop on precipitation in urban areas, St.Moritz, Switzerland2009

Schellart A, Ochoa S, Simões N, Wang L-P, Rico-Ramirez M, Liguori S, Duncan A, S. Chen A, Keedwell E, Djordjević S, Savić DA, Saul A, Maksimović Č (2011) Urban pluvial flood modelling with real time rainfall information – UK case studies. Paper presented at the 12nd International Conference on Urban Drainage, Porto Alegre/Brazil, 10-15th September 2011

Schellart ANA, Rico-Ramirez MA, Liguori S, Saul AJ (2009) Quantitive precipitation forecasting for a small urban area: use of radar nowcasting. Paper presented at the 8th International workshop on precipitation in urban areas, St.Moritz, Switzerland2009

Simões N, Wang L, Ochoa S, Leitão JP, Pina R, Onof C, Sá Marques A, Maksimovic C, Carvalho R, David L (2011a) A coupled SSA-SVM technique for stochastic short-term rainfall forecasting. Paper presented at the 12th International Conference on Urban Drainage, Porto Alegre/Brazil, 11-16th September 2011a

Simões N, Wang L, Ochoa S, Leitão JP, Pina R, Sá Marques A, Maksimovic C, Carvalho RF, David L (2011b) Urban flood forecast based on raingauge networks. Paper presented at the 12th International Conference on Urban Drainage, Porto Alegre/Brazil, 10-16th September 2011 2011b

(2011) http://www.stormgeo.com/