MUFFIN Technical Report

1. Introduction

This Technical Report focuses on the work performed during the "second half" of the MUFFIN project, complemented with some summaries of the work performed during the "first half". For further details on the first half, see Section 2a in the Mid-term Progress Report (Appendix IIf).

In the following, this report is divided into three sections:

- *The Joint Experiments*. Three experiments pertinent to different key aspects urban flood forecasting, where the joint expertise of the different partners produce an added value.
- *Partner-specific Development*. Descriptions of local obervations, analysis and modelling by the partners related to the project objectives.
- *End-user Feedback*. A summary of the feedback received from end-users associated with the project and an overview of how that feedback was considered in the project activities.

2. The Joint Experiments

Options for the Joint Experiments were conceptually outlined in the Mid-term Progress Report (p. 11). After discussions between all partners, three experiments were designed:

- *Hydrodynamic vs. high-resolution hydrological modelling*
- *High-intensity rainfall in European operational radar observations*
- *Development of a multi-scale flood forecasting system*

The experiments were designed with the dual aim to fulfill the different objectives of the project (see Section 2a in the Final Progress Report) as well as to take end-user requirements into account (see Section 4 below).

2.1 Joint Experiment 1: Hydrodynamic vs. high-resolution hydrological modelling

The purpose of this experiment was to assess to which degree the output from a large-scale (but high-resolution) hydrological model can identify and quantify urban flooding, as estimated from a local hydrodynamic model. Models were set up as follows:

- *High-resolution hydrological model*. Set-ups of the 1D rainfall-runoff model HYPE on two local urbanized basins in DK (Kærby, in Aalborg) and FI (Länsi-Pakila, in Helsinki), customized to operate in high resolution. In MUFFIN, the HYPE model is updated with a higher resolution and detail compared to the present national Swedish setup (S-HYPE) and the pan-European set-up (E-HYPE) (see also Section 3.1 below). This means a 1-h temporal resolution, a higher level of detail with regards to surface elevation based on high-resolution digital elevation models (DEM), as well as a detailed description of imperviousness in urban areas in order to improve the prediction of runoff in urban areas.
- *Hydrodynamic model*. These setups are customized hydrodynamic distributed (1D) local flood models, which include the relevant hydrological processes required to predict urban flooding in a high level of detail. The models thus account for the urban drainage system, preferential water

ways (e.g. roads and channels), infiltration/runoff processes, etc., to a larger degree than in the high-resolution hydrological model. In the two basins, different commercial models were applied (Länsi-Pakila: SWMM; Kærby: MIKE URBAN).

From the rain gauge network in the greater Copenhagen area, a common long-term precipitation forcing dataset has been produced. The dataset consists of gridded precipitation data for each year between 1979-2017 with spatial resolution 1x1 km and temporal resolution 1 hour. It has been used in all four models in this joint experiment.

This experiment thus explores the flood risk correspondence between a high-resolution large-scale hydrological model and a local flood model (MIKE, SWMM). Despite the temporal and spatial differences, runoff dynamics and volume in HYPE performs well compared to the local flood models. Runoff peaks are however not always captured in HYPE and as such HYPE lacks the ability to predict flood events caused by short high-intensity rainfall. For flood events caused by medium to long rainfall events, HYPE has a similar performance in terms of estimating whether a flood occurred compared to the local flood models. However, the magnitude of the flood events differs between HYPE and the local flood models. HYPE does provide benefits beyond the possibilities of the local scale flood model in Aalborg in case of fluvial flood events.

The required resolution of the rainfall input depends on the study case. Flood risk in the local flood models is best correlated to 10-30 minutes aggregated rainfall but even 60 minutes rainfall aggregates seem to be correlated to flood risk, see Figure 1. This suggests rainfall forecasts of 10 – 30 minute temporal resolution are optimal for the joint case studies.

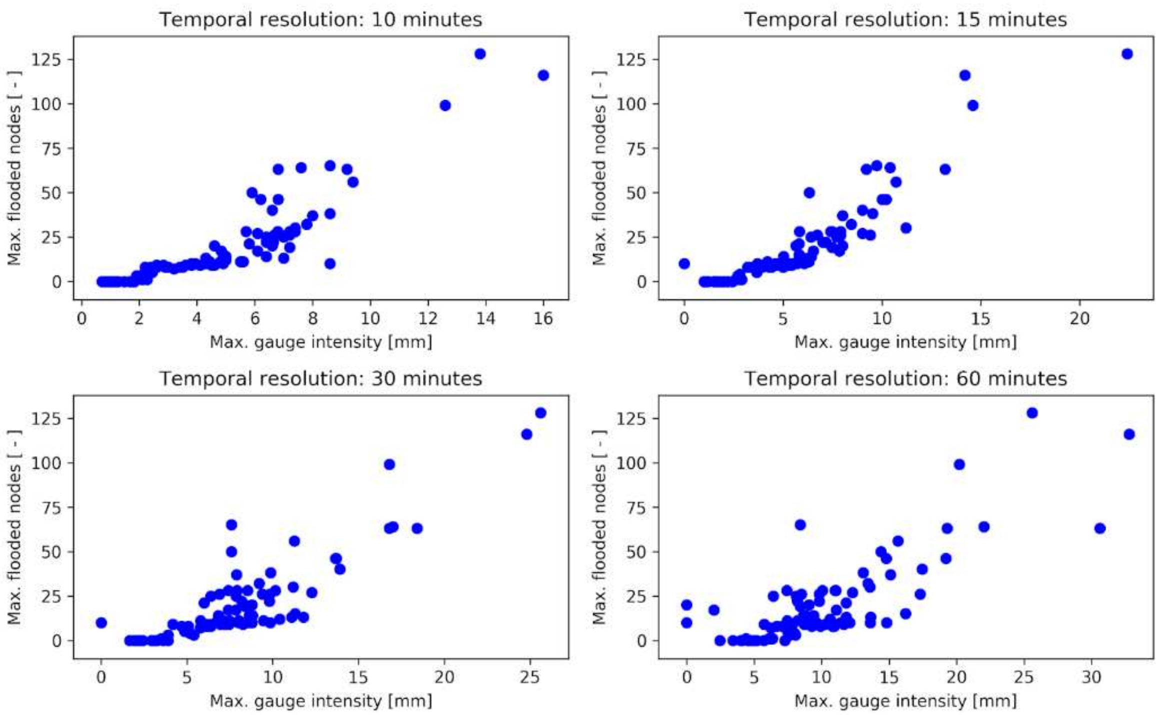

Figure 1. Example of aggregated rain gauge measurements into 10, 15, 30, and 60 minutes compared to the number of flooded nodes (flood risk) simulated in the Kærby catchment.

In terms of available forecasts with the optimal temporal scale, radar nowcasting with a lead time of a few hours at most is the go-to solution. Numerical weather forecasts (1 hour resolution and several days lead times) could be feasible to predict the risk of an incoming flood but with less certainty when it comes to the flood magnitude.

The results of Joint Experiment 1 will be published in Nielsen et al. (2019a).

2.2 Joint Experiment 2: High-intensity rainfall in European operational radar observations

A multinational assessment of radar's ability to capture heavy rain events at scales of 5 min up to 2 hours was performed. Until now, only few comparisons between countries were available, especially when it comes to extremes. In total, 6 different radar products in Denmark, the Netherlands, Finland and Sweden (representing different scales/resolutions and accuracies) were analysed. The 50 most intense events for each country were used to quantify the overall agreement between radar and gauges as well as the errors affecting the peaks. Results showed that the overall agreement between radar and gauges is fair, with radar underestimating rainfall rates by 29-40% compared with gauges. However, the bias increased with intensity to reach 46%-66% during the peaks. Only part of the bias (i.e., 13%-30%) could be explained by differences in measurement areas between gauges and radar, see Figure 2.

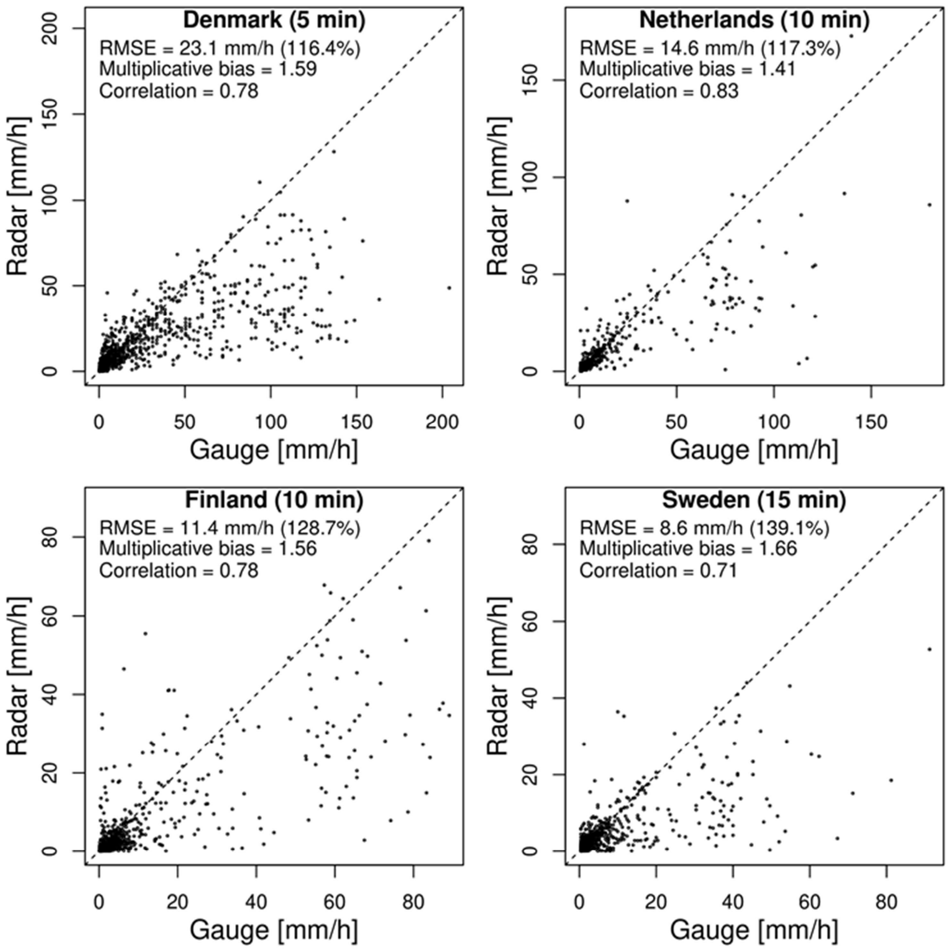

Figure 2. Agreement between radar and rain gauges for the 50 most intense rain events in each country.

Overall, radar products with higher spatial and temporal resolution agreed better with each other, highlighting the importance of high-resolution radar for urban hydrology. The X-band data for Denmark showed very promising results, outperforming all other products in terms of accuracy and correlation. However, for catching the rainfall peaks responsible for flooding, the ability to combine measurements from multiple overlapping radars to help mitigate attenuation and reduce bias seemed to play a more important role than resolution. The use of dual-polarization and phase information (e.g., Kdp) in the Finnish product also seemed to provide a slight advantage in heavy rain. But improvements were hard to quantify and similarly good results were achieved in the Netherlands by applying a simple Z-R relation together with a mean field bias-correction. Another important finding of this joint experiment was that the largest bias between radar and gauges in terms of peak intensities does not necessarily occur at the highest temporal sampling resolution. Depending on the autocorrelation structure of the errors, multiplicative biases may amplify over time instead of averaging out. This mostly happens at the sub-hourly time scales and roughly affects 40-50% of all events in single-radar products and 15-30% in composite products. Most of these problematic cases were characterized by a succession of multiple rainfall peaks or one very intense peak of 15-30 min during which radar strongly underestimated the intensity during 2 or more consecutive time steps. The strong dependence of the error structure on the underlying aggregation time scale still represents a major challenge in terms of how to correctly represent rainfall extremes and rainfall measurement uncertainties in hydrological models.

The results of Joint Experiment 2 will be published in Schleiss et al. (2019).

2.3 Joint Experiment 3: Development of a multi-scale flood forecasting system

The purpose of this experiment was to explore the possibility of coupling a large-scale high-resolution hydrological model (HYPE) to a local hydrodynamic model (MIKE URBAN) into a multi-scale urban flood forecasting system. The experiment was performed in the Kaerby basin (Section 2.1), Aalborg. The model set-ups were similar to the ones in Joint Experiment 1 (Section 2.1), but instead of a 1D hydrodynamic model a 1D-2D local flood model was set up for the Kærby catchment. This model expands the local catchment setup by simulating the flood extent and water depth on the terrain in 2D.

This experiment thus investigates the potential gain in flood forecasting performance by utilizing the fast-computational time of HYPE to forecast river discharge as a boundary condition for the 1D-2D local model setup in Aalborg. Known flood prone areas in the catchment along the river system are not observed in flood simulations using the current available urban drainage models. This is mainly due to the description of the river discharge from a flow and water level relation (Q-h relation) being incorrect during high-discharge scenarios. Therefore, substituting the river discharge boundary condition from a Q-h relationship to a simulated river discharge from a calibrated HYPE model might improve performance.

To investigate this two numerical weather forecasts and a radar nowcast (representing e.g. different scale/resolution or different observational sensors or different forecasting methodology) was used as forcing data for the coupled HYPE and 1D-2D urban drainage flood model. See Table 1 for an overview of the forecasts used in this experiment.

Table1. Forecast details (R: Radar Nowcast, N: Numerical Weather Forecast).

Forecast	**Time step**	**Spatial resolution**	**Lead time**
AAU nowcast (R)	1 min	100 x 100 m2	10, 15, 30, and 60 min
MEPS (N)	1 h	2.5 by 2.5 km2	0-6 h, 6-12 h, 12-18 h, 18-24 h
ECMWF (N)	1 h	11 x 11 km2	0-12 h and 12-24 h

Forcing the coupled flood model with numerical weather forecasts yields better performance when predicting fluvial related flood events. However, using high-resolution nowcasts as forcing data captures pluvial flood events. Thus, combining both forecasts (i.e., forcing the 1D-2D flood model with radar nowcast and HYPE with the numerical weather forecast in the coupled model yields the best overall performance. See Figure 3 for a comparison of using the combined forecasts vs. the radar nowcast only.

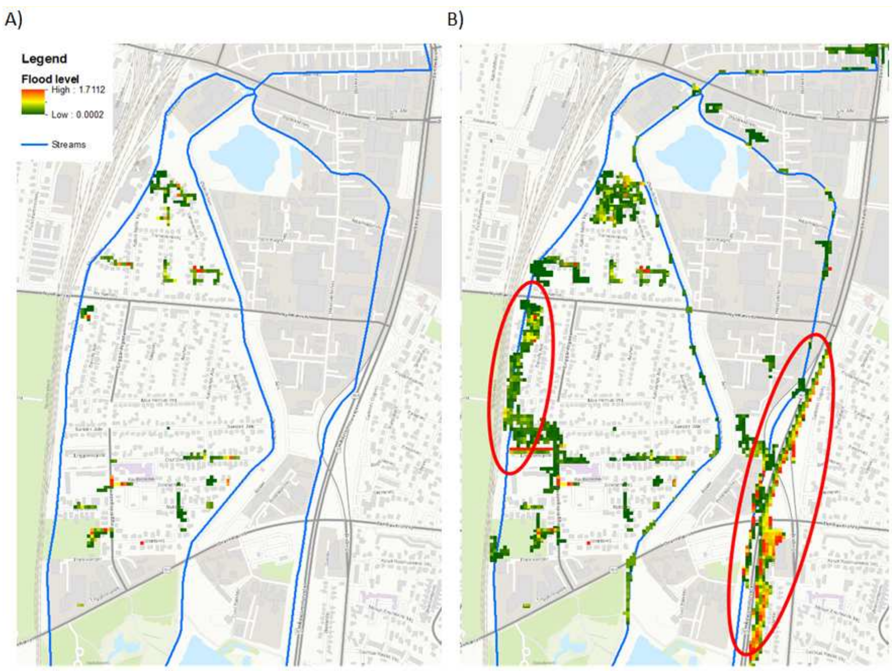

Figure 3. Predicted flood depth and extent from an event simulated using (a) the radar nowcast for the urban flood model only and (b) the radar nowcast for the local flood model and the numerical weather forecast for the HYPE model. This event had a high river discharge and medium rainfall intensity, which generally would cause flooding in the marked areas in (b) but did not show in (a).

This coupled model with the combined forecasts expands the current limitations in flood forecasting in the Aalborg catchment, where previous attempts were unable to map the flood prone areas around the river system correctly. For more information and possibility to test different settings concerning rain events and forecasts, see the web based interface on the MUFFIN webpage (http://www.muffin-project.eu/).

The results of Joint Experiment 3 will be published in Nielsen et al. (2019b).

3. Partner-specific Development

In this section, work performed in the project by each partner besides the Joint Experiments (although often with clear connection to them) is reported.

3.1 Swedish Meteorological and Hydrological Institute (SMHI)

The work during the first half of the project was focused on collecting and quality-assuring observed precipitation data and start setting up the HYPE model in the study basins in Aalborg and Helsinki (see below). The data represent different scales; local, national and pan-european. Also, storage of high-resolution precipitation nowcasts and ensemble forecasts has been automatized, for re-forecast experiments later in the project. A key activity was to develop the HIPRAD product (Berg et al. 2016), which merges station and radar observations into a high-resolution (2×2 km2, 15-min) and long-term (since 2000) precipitation data base. SMHI also built the website (http://www.muffin-project.eu/) and led the first advisory board meeting in November 2016 at Aalborg University.

Most of the previous large scale applications of the hydrological model HYPE were based on simulations at a daily time step. This is mainly because the forcing data at such scales are commonly available at a daily time step. However, with the development of new and improved estimation techniques of precipitation at high temporal resolution from sources such as weather radar, forcing data for hydrological simulation at sub-daily time scale are becoming available. One of the activities within the project was development of 1h-HYPE for high resolution flood forecasting for Sweden.

The same model setup used operationally for Sweden at a daily resolution (S-HYPE) was employed for the development with different parametrization and forcing data. The HIPRAD data set (Berg et al. 2016), which is hourly precipitation data derived from weather radars by constraining the monthly sum against gridded daily data used in the operational daily model was used together with hourly temperature from the MESAN reanalysis system (Häggmark et al. 2000). Hourly discharge data at several stations were also used for the development.

Much of the development work has been focused on parametrization of the model. The first attempt was done by rescaling time dependent model parameters estimated for the daily model to the appropriate time scale and retaining time invariant parameters from the daily model. Although this resulted in a model that performs acceptably, further work was done to improve estimation of parameters by identifying sensitive parameters and account for variability of model parameters with respect to catchment characteristics. A set of catchments were selected with different catchment characteristics from different regions so that the different climate regimes are represented. A global sensitivity analysis was performed to identify sensitive parameters by performing a large number of model simulations with quasi-random parameter samples (Saltelli, et al. 2006). As expected, the sensitive parameters identified for different climate regions were different as the dominant processes are different. Values of the identified sensitive parameters for each catchment were then estimated as average values of the corresponding parameters in the 10% best performing model simulations. These parameter values are transferred to other catchments based on similarity of the catchment characteristics, such as soil type and landuse. The parameters estimated in this way also resulted in a model that performs comparably well across the whole model domain as the model parametrized by rescaling the time dependent parameters and retaining the time invariant parameters of the daily model. Further development is underway to establish a relationship between parameters of the daily and hourly models so that new developments in the daily model can be incorporated to the hourly model.

A 1D rainfall-runoff model HYPE is setup for the two case study catchments in Helsinki and Aalborg as a larger scale flood forecasting tool in the multi-scale forecasting experiment. This setup is customized to operate at higher resolution and detail compared to the present setup of the pan-European E-HYPE and the Swedish S-HYPE models. The models are setup to run at 1h temporal resolution and a higher level of detail with regards to surface elevation based on high resolution digital elevation models (DEM), as well as detailed description of imperviousness in urban areas are incorporated in order to improve the prediction of runoff in urban areas.

The Østero catchment in Aalborg, which is 125km² in size, is subdivided into 2343 adaptive polygon subcatchments, where the highly urbanized areas are subdivided into fine subcatchments as small as 1000m² and the rural settings are divided into much coarser sizes (Figure 4a). Hourly precipitation data from seven stations within and around the catchment were used to run the model by using data from the nearest station to each subcatchment. Hourly temperature data from Aalborg airport was also used, which was uniformly applied throughout the catchment. Hourly discharge data from a station that drains 115.7km² of the catchment for the summer winter seasons over period Oct 2013 to March 2016 was used for model calibration.

The 10km² Haaganpuro catchment in Helsinki was subdivided into 62 regular 500m grid based subcatchments, with cells at the boundaries cut into irregular shapes to match the catchment outline (Figure 4b). Precipitation measured at a station within the catchment at 1min resolution for the period May 18 to Oct 30, 2017 was aggregated to hourly data and used together with 10min temperature from the nearest FMI station Kumula averaged over each hour to run the model. Hourly discharge data at stations Pakila (0.42km²) and Hagganpuro (10.8km²) were used to calibrate the model.

a (Aalborg) b (Helsinki)

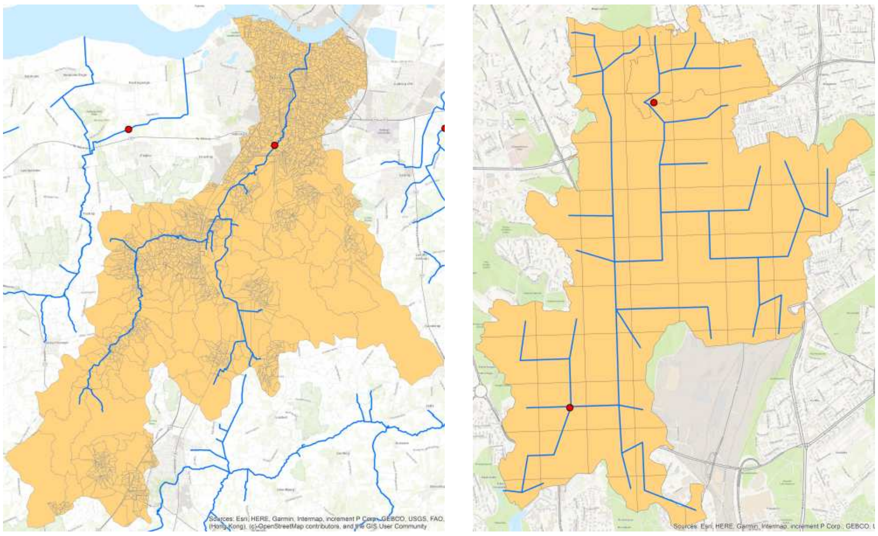

Figure 4. Set-ups of high-resolution catchment models for the case study catchments

While the calibrated model in Helsinki performs remarkably well, given the size and time resolution of the model, the model in Aalborg performs not that well (Figure 5). The discharge data for the Aalborg catchment is not that reliable since the employed rating curve was not established correctly and this might be the reason for the poor model performance in Aalborg.

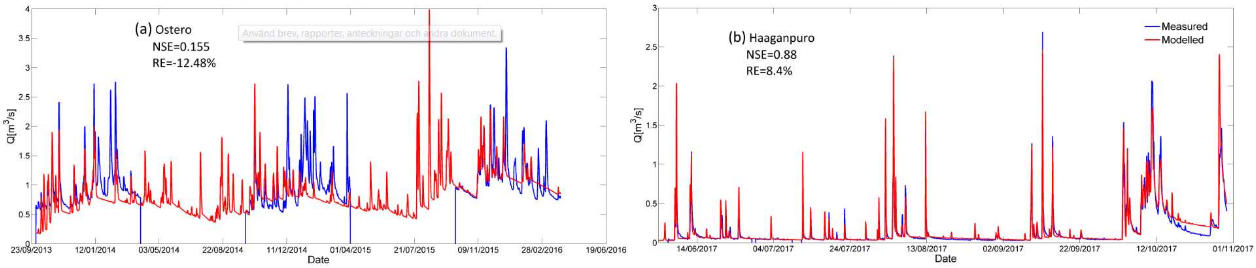

Figure 5. Comparisons of model simulated and observed hydrographs at stations Østero (Aalborg) (a) and Haaganpuro (Helsinki) (b).

3.2 Swedish Geotechnical Institute (SGI)

Rainvis is a real-time high-resolution high-intensity rainfall visualization prototype from Sweden (Figure 6). It has been developed within the MUFFIN project in order to provide the user the best possible information and decision support both *before* a flood event (forecasts – for early warning), *during* the event (observation and forecast – for situation awareness) and *after* the event (observations – for post-event analysis). There are many visualization products today to assess the risk of high-intensity rainfall and subsequent pluvial flooding. However, there are hydrological limitations with these which Rainvis has attempted to overcome, for example the following. (1) The radar-based rainfall estimates are more closely adjusted towards gauge observations, represented by daily gridded fields, which ensures accurate long-term accumulations (Berg et al. 2016). (2) Concerning spatial resolution, the radar rainfall is averaged over hydrological basins (~40 000 sub-basins) covering Sweden with a median size of ~7 km². (3) Concerning temporal resolution, besides the highest available (1 hour), rainfall may be averaged over durations of 2, 3, 6 or 12 hours. (4) Observations from the recent hour(s) may be combined with forecasts for the coming hour(s).

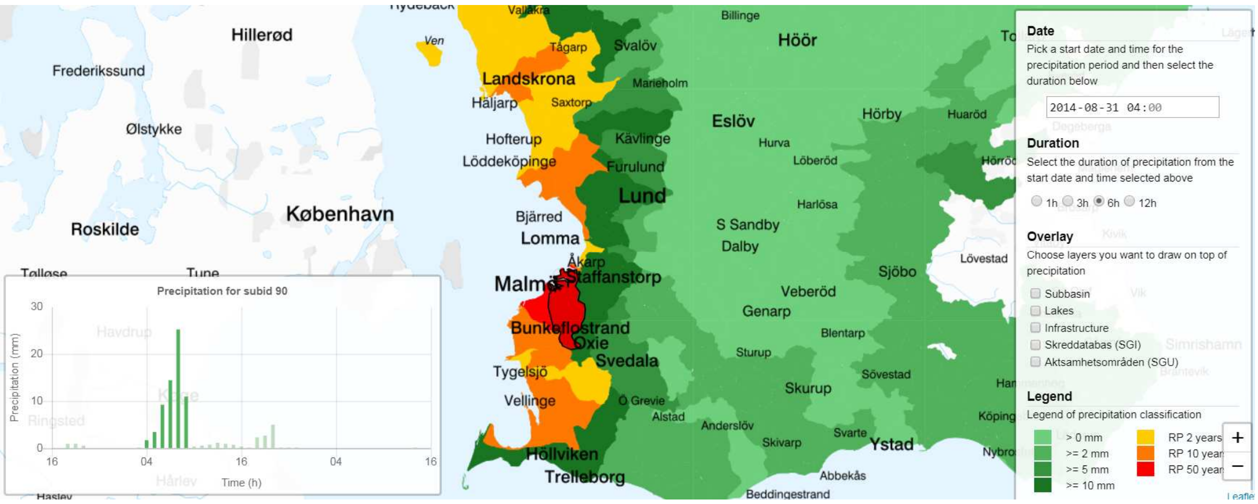

Figure 6. Screenshot from the Rainvis prototype.

3.3 Technical University of Delft (TUD)

During the first half of the project, TUD set up a state-of-the-art nested numerical weather prediction model (WRF 3.9) with a multi-scale configuration of three two-way nested domains of BeNeLux region, the Netherlands and Rotterdam respectively. Furthermore, data from professional weather sensors (10 Campbell weather stations) in Rotterdam were automatically transferred to the TUD web servers. From July 2017, six Davis weather stations together with the GNSS receivers for water vapour retrievals were operational.

Detailed investigations into the relation between urban pluvial flooding and rainfall were performed for the Rotterdam area thanks to 70'000 citizen flood reports over a 10 year period from 2008 to 2017. Reports were collected by means of phone, email, mobile app, and webpage and stored in a database along with their date, location and a short description of the event. In Figure 7a the most frequent keywords associated with the citizen reports are shown. Using machine learning, several critical rainfall thresholds beyond which urban pluvial flooding is likely to occur were derived.

Our results show that 37%–52% of all flood occurrences and 95%–97% of all non-flood occurrences can be predicted, which is a fair performance given the uncertainties associated with citizen data. More importantly, all models agreed on which rainfall features are the most important for predicting flooding, reaching optimal performance whenever short- and long-duration rainfall peak intensities were combined together to make a prediction. The encouraging results suggest that citizen observatories, although prone to larger errors and uncertainties, constitute a valuable alternative source of information for gaining insight into urban pluvial flooding. The study also highlighted a number of issues related to citizen flood observations which need to be improved (i.e., the lack of information about the type and origin of the flooding and the limited accuracy of the time stamps given to citizen observations). And although the derived rainfall thresholds are specific to Rotterdam, the approach and methods are sufficiently general to be transferable to other urban catchments as well (Tian et al., 2019).

In addition to citizen flood reports, TU Delft also collected a large number of rainfall measurements from 750+ citizen weather stations (CWS) in the greater Rotterdam area. The latter provide higher-density spatial measurements compared with professional networks, albeit with more frequent and larger errors. The lower quality and poor maintenance of CWSs mean that rigorous quality control mechanisms had to be implemented. Therefore, an automated filter for detecting and removing incorrect zero rainfall values from CWS was designed. Simulation experiments showed that 74% of the faulty zeros can be identified with this technique, with a false negative rate of only 1.1%. Furthermore, most mistakes occur in periods of low rainfall intensities which are irrelevant for flooding. In a subsequent study, the new filter was applied to a subset of 74 CWSs in the city center of Rotterdam to create high-resolution 1x1 km rainfall maps see Figure 7b. The latter were compared to a state-of-the-art gauge-adjusted radar rainfall composites and to measurements from a single professional weather station at Rotterdam airport. Results show that for the city of Rotterdam, whenever a professional station is located within a distance of 4 km or less from a CWS, the rain rate measured by the professional station should be used, as it will be more accurate and reliable. At distances of 4-8 km from a professional station, CWS stations generally have similar representativity. Beyond 8 km from a professional station, it is better to rely on CWS data. Also, multiple instances were identified in which the CWS-interpolated rainfall maps were more plausible than radar estimates, whose quality over urban areas tends to be lower. This is very encouraging and clearly highlights the potential of crowdsourcing for urban rainfall estimation. The results of this study are currently being compiled into a paper to be submitted to the journal “Frontiers”.

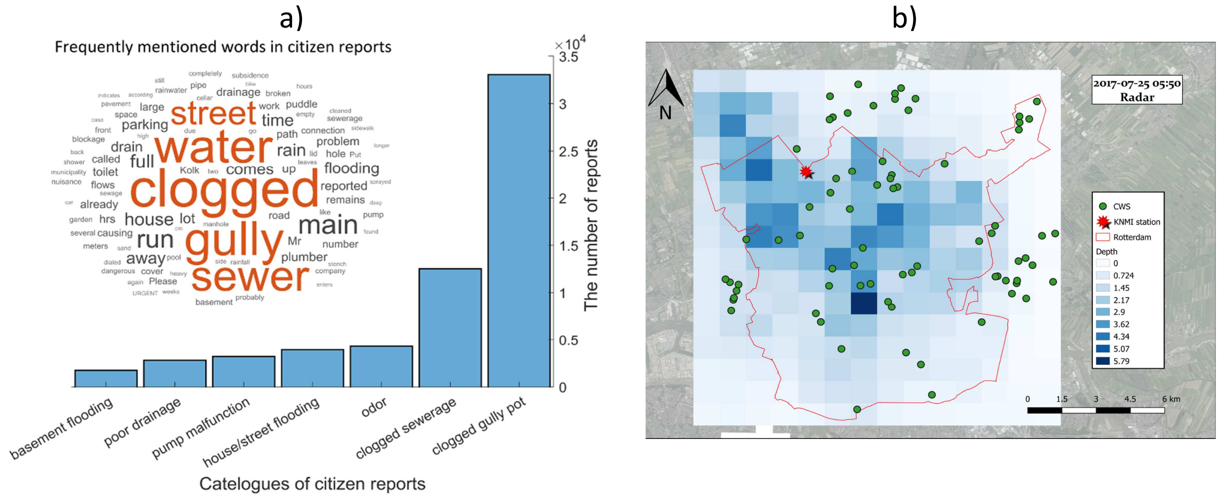

Figure 7. Word cloud of the most frequent keywords associated with the citizen reports (a) and estimated rainfall map from citizen weather stations for the city centre of Rotterdam for one event in 2017.

The critical element of the project was the Rijnmond X-band radar in the city center of Rotterdam. The radar briefly came to live in 2017-2018. But very few good data could be collected. Then, after a few months, the radar got severely damaged during a storm. After the incident, the whole mechanical part had to be redesigned, which took time and meant that this radar could not be used during MUFFIN. On the bright side, lots of data were collected using a smaller, vertically profiling micro rain radar (MRR), see Figure 8. The later was installed between Nov 2018 and April 2019 near the Cabauw site for atmospheric research, between the X-band research radar IDRA and the C-band operational weather radar In Herwijnen. In total, 19 events were sampled. The data were used to retrieve vertical profiles of raindrop size distributions to help calibrate the Herwijnen radar and improve its accuracy. The results are currently being compiled into a publication to be submitted to “Atmospheric Measurement Techniques”. At a later stage, once the Rijnmond X-band radar will be operational, the MRR acquired during the MUFFIN project will be moved to Rotterdam, to help improve the urban rainfall estimates.

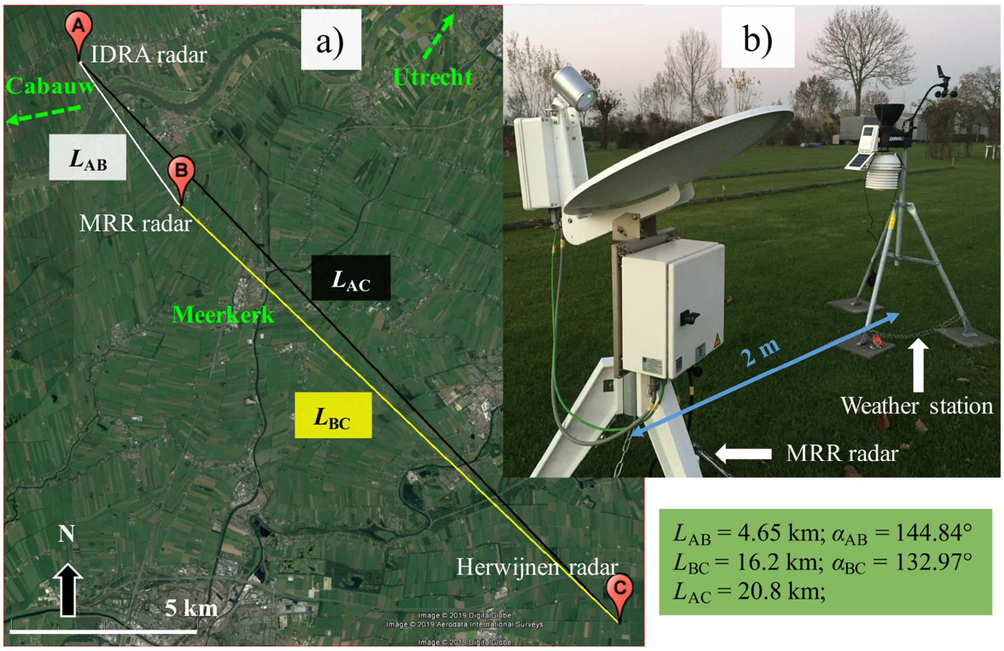

Figure 8. Setup of the micro-rain radar near the Cabauw site for atmospheric research.

Freely available operational KNMI national radar composites at 1km and 5min and radar nowcasts up to 2h lead times are also available from May 2017 onwards. They will be used as a benchmark to test and validate the high-resolution X-band weather radar data and the numerical weather forecasts.

3.4 Aalto University (AALTO)

The Länsi-Pakila (~ 1 km2) catchment was instrumented for measuring storm water flow and on-site precipitation. The area has experienced problems associated with excess storm water and is expected to become more densely built in the future, which can lead to increasing challenges with excess water unless sound storm water management practices are implemented. An acoustic storm water flow gauge was installed to a 800 mm pipe and it recorded the discharge at one minute resolution for a five month period from May 23, 2017, to Nov 2, 2017. The on-site rain measurement comprised three tipping bucket rain gauges attached to a tripod that was placed on the roof of a nearby low rise building. The temporal resolution of the on-site rainfall measurement was one minute.

A Storm Water Management Model (SWMM) spatial parametrization for the Länsi-Pakila catchment was constructed. The high-resolution model was generated with an automatic tool utilizing land-cover, topographical and stormwater drainage network information (Niemi et al., 2019).

This SWMM model, and Itzï (Courty et al. 2017) flooding extent model, was linked with pySTEPS rainfall nowcasting model (Pulkkinen et al. 2019) to produce a nowcast for urban runoff and flooding extent. An ensemble of urban runoff nowcasts was compared against observed stormflow for an extreme rainfall event in Länsi-Pakila, which indicated the need for further improvement in predicting the time of the storm arrival.

In addition to the Länsi-Pakila catchment, a flow gauge was also installed to the Haaganpuro brook at a location that collects runoff from a ca. 11 km2 catchment. The Länsi-Pakila catchment is a sub-catchment of the larger, more natural Haaganpuro catchment and this nested experimental setup provides data that can help e.g. in exploring the role of pervious areas in urban runoff generation.

The radar meteorology group of the University of Helsinki has constructed a composite rainfall product for the capital area utilizing three dual polarized C-band radars. This product takes advantage of the dual polarization measurements and it is available for the MUFFIN project. Weather radar data were also available from a nationwide OSAPOL research product (Rimpiläinen, L., 2017), which was utilized in comparing the performance of national scale weather radar data in describing high intensity precipitation events (Schleiss et al. 2019).

In the MUFFIN project earlier work on comparing the usability of different scale land-cover information in urban runoff modelling has been continued. Spatial datasets depicting urban surfaces was compiled from several sources with varying coverage and spatial detail. The sources include 1) manually interpreted information from city maps, aerial photographs and site visits; 2) regional high resolution land-cover data based on e.g. interpretation of colour infrared orthoimages LiDAR data, and road data; and 3) the Urban Atlas dataset available for a large number of major cities in Europe, see Figure 9. The results were published in Kokkonen et al. (2019) and they conclude 1) that openly available, metropolitan scale land-cover description is on par in performance with laboriously collated tailored land-cover description, and 2) that mixed, aggregated land-cover types typically present in continental scale datasets are problematic particularly in low density urban

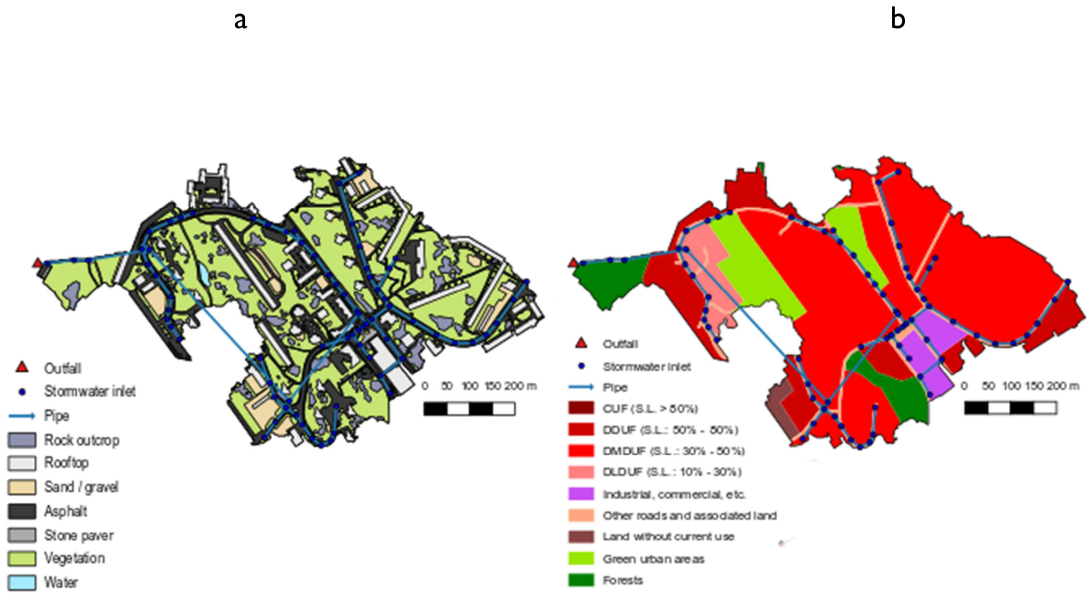

Figure 9. Land-cover in the same catchment as obtained from manually interpreted information (a) and the EEA Urban Atlas (b).

3.5 Aalborg University (AAU)

During the first half of the project, the work of AAU selected the Kaerby catchment as case study area, developed an urban drainage model of the urban drainage system, streams and ground water and planned a detailed measurement and monitoring campaign of different hydrological processes.

As AAU were most heavily involved in the Joint Experiments, there was little resources left for other work. However, a prototype real-time model was developed for the Kærby catchment using a neural network to relate rain intensity to the flood extent and depth. In Figure 10, a comparison of the predicted flood extent and depth from the neural network is compared to the simulation of the urban drainage system. The neural network predicts the extent of a flood event similar to an urban drainage model in less than 15 seconds.

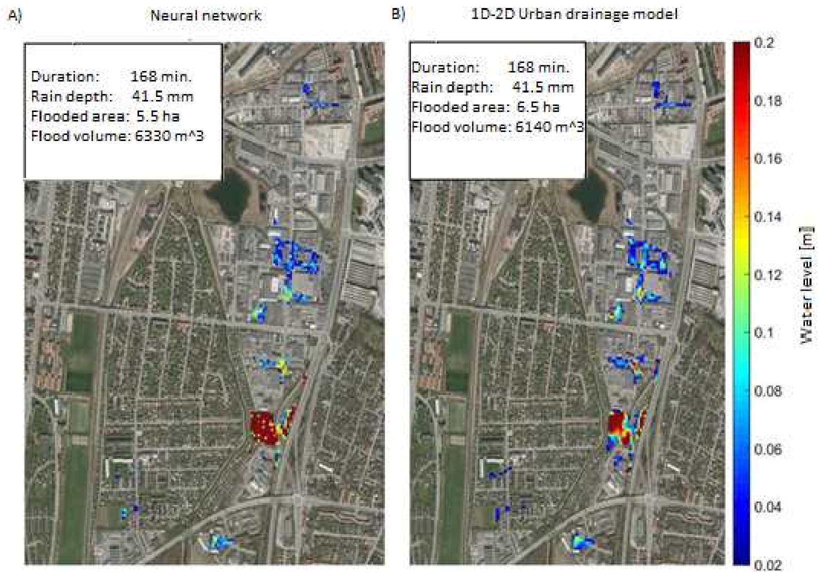

Figure 10. Illustration of the estimated flood depth and extent by a neural network (a) and the urban drainage model (b). The neural network performs best on flood extent.

4. End-user Feedback

To specify the end-user value, SGI used a three-prong or triangulation method to understand the needs and requirements of the MUFFIN end-users. In the MUFFIN case these three methods consisted of 1) an international workshop in Feb 2017, 2) an end-user survey administered in Dec 2017 and 3) in-depth telephone interviews with end-users in Nov 2017- Feb 2018. In summary, end-users had differing needs and conditions and thus MUFFIN cannot provide one-size-fits all solution to the problems of urban flooding. Some (important) input was not related to MUFFIN, and in some cases we referred to the INXCES project.

In Table 2, a synthesis of end-user requirements are listed (see Section 4.7 in Appendix IIb), together with a description of how these requirements were taken into account in the project work.

Finally, the final conference CITIES, RAIN and RISK was overall highly appreciated by the almost 150 participants, out of which 72% were end-users and 28% researchers. On a scale from 0 (very bad) to 10 (very good), the average rating was 8.1.

Table 2. End-user requirements and actions taken in MUFFIN.

End-user requirement	**Action taken**
MUFFIN can continue to explore the added value of resolution and lead time in terms of accuracy and uncertainty, both nationally (SMHI) and in the case studies.	The impact of resolution on accuracy and uncertainty is a main aspect in all three Joint Experiments, covering rainfall obsevations and forecasts as well as flood models.
MUFFIN can consider how forecasts can be used as input to authorities to enable them to optimize when and how often warnings are given.	The identification of critical rainfall thresholds associated with urban flooding, performed in MUFFIN, is highly relevant for issuing and optimizing flood warnings.
MUFFIN can provide short user-friendly guidance papers about how the project's forecasts and monitoring can be utilized by the various types of actors and how the uncertainty associated with the forecasts can be interpreted.	The two main prototypes developed in the project (for (*i*) rainfall visualization and (*ii*) multi-scale flood forecasting) will be accompanied by documentation, guidelines and manuals that are intended to fulfill this user-request.
MUFFIN can consider more merging of data not only from rain gauges and from radar, but also real-time observations from the general public, including photos and web-cams at critical locations.	The work at TUD with exploring private weather stations as well as citizen reports are efforts in this direction, that will continue in follow-up projects (e.g. using tele-communication links for rainfall estimation).
MUFFIN can consider a framework for comparing and contrasting the processes in each country of how cases work to bridging the gap between urban and large scale hydrological models and specific tools for reducing the impacts of precipitation.	Joint Experiment 1, comparing hydrological and hydro-dynamic tools under different conditions (but in a common framework), can be viewed as at least a preliminary attempt in this direction, to be further developed in follow-up projects.
MUFFIN can consider, in a complementary project at the end of the project or shortly thereafter, an evaluation on how results were actually utilized by end-users.	The prototypes and other output from the project will be widely disseminated from now on and we will monitor the user uptake and feedback.

References

Berg, P., Noring, L., and Olsson, J. 2016. Creation of a high resolution precipitation data set by merging gridded gauge data and radar observations for Sweden. Journal of Hydrology, 541(A), 6-13.

Courty, L.G., Pedrozo-Acuña, A., Bates, P.D. 2017. Itzï (version 17.1): an open-source, distributed GIS model for dynamic flood simulation. Geoscientific Model Development 10, 1835–1847. https://doi.org/10.5194/gmd-10-1835-2017

Häggmark, L., Ivarsson, K.-I., Gollvik, S. and Olofsson, P.-O. 2000. Mesan, an operational mesoscale analysis system. Tellus 52A, 2–20.

Kokkonen T, Warsta L, Niemi TJ, Taka M, Sillanpää N, Pusa M et al. 2019. Impact of alternative land cover descriptions on urban hydrological model simulations. Urban Water Journal. 2019 Feb 7;16(2):103-113. https://doi.org/10.1080/1573062X.2019.1634742

Nielsen, R., Niemi, T., Thorndahl, S., Hundecha, Y., Schleiss, M., Kokkonen, T., Olsson, J. 2019a. Hydrological vs. Hydrodynamic Modelling – A Comparative Study in Aalborg and Helsinki, manuscript in preparation.

Nielsen, R. Thorndahl, S., Hundecha, Y., Olsson, J. 2019b. Large-scale and local flood modelling – Overcoming drawbacks of both domains by combining strengths, manuscript in preparation.

Niemi T, Kokkonen T, Sillanpää N, Setälä H, Koivusalo H. 2019. Automated Urban Rainfall–Runoff Model Generation with Detailed Land Cover and Flow Routing. Journal of Hydrologic Engineering. 2019 May 1;24(5). 04019011. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001784

Pulkkinen, S., Nerini, D., Pérez Hortal, A., Velasco-Forero, C., Germann, U., Seed, A., Foresti, L. 2019. Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0). Geoscientific Model Development Discussion. https://doi.org/10.5194/gmd-2019-94

Rimpiläinen, L. 2017. Tutkapohjainen mitoitussateiden muotokirjasto (Selection of radar based design storms) (Master's Thesis). Aalto University, Espoo, Finland.

Saltelli, A., Ratto, M., Tarantola, S., Campolongo, F. 2006. Sensitivity analysis practices: Strategies for model-based inference. Reliab. Eng. Syst. Saf. 91, 1109–1125. doi.org/10.1016/j.ress.2005.11.014.

Schleiss, M., Olsson, J., Berg, P., Niemi, T., Kokkonen, T., Thorndahl, S., Nielsen, R., Ellerbæk, J., Bozhinova, D. and Pulkkinen, S. 2019. The accuracy of weather radar in heavy rain: a comparative study for Denmark, the Netherlands, Finland and Sweden, Hydrology and Earth System Sciences, submitted.

Tian, X., ten Veldhuis, M-C., Schleiss, M., Bouwens, C., van de Giesen, N. 2019. Critical rainfall thresholds for urban pluvial flooding inferred from citizen observations. Science of The Total Environment; 689: 258-268.