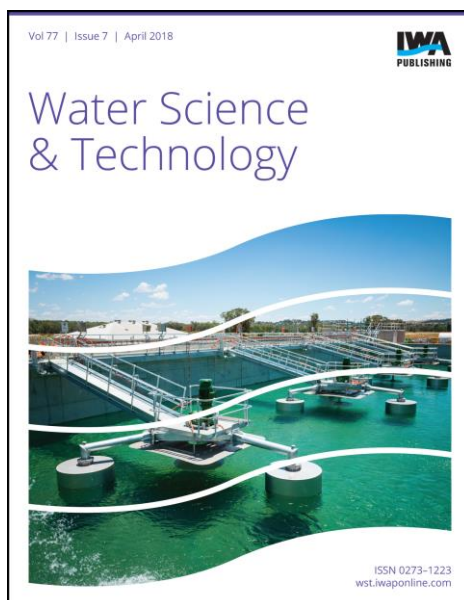


## ELECTRONIC OFFPRINT

Use of this pdf is subject to the terms described below



This paper was originally published by IWA Publishing. The author's right to reuse and post their work published by IWA Publishing is defined by IWA Publishing's copyright policy.

If the copyright has been transferred to IWA Publishing, the publisher recognizes the retention of the right by the author(s) to photocopy or make single electronic copies of the paper for their own personal use, including for their own classroom use, or the personal use of colleagues, provided the copies are not offered for sale and are not distributed in a systematic way outside of their employing institution. **Please note that you are not permitted to post the IWA Publishing PDF version of your paper on your own website or your institution's website or repository.**

If the paper has been published "Open Access", the terms of its use and distribution are defined by the Creative Commons licence selected by the author.

Full details can be found here: <http://iwaponline.com/content/rights-permissions>

Please direct any queries regarding use or permissions to [wst@iwap.co.uk](mailto:wst@iwap.co.uk)

## Ensembles of radar nowcasts and COSMO-DE-EPS for urban flood management

Alrun Jasper-Tönnies, Sandra Hellmers, Thomas Einfalt, Alexander Strehz and Peter Fröhle

### ABSTRACT

Sophisticated strategies are required for flood warning in urban areas regarding convective heavy rainfall events. An approach is presented to improve short-term precipitation forecasts by combining ensembles of radar nowcasts with the high-resolution numerical weather predictions COSMO-DE-EPS of the German Weather Service. The combined ensemble forecasts are evaluated and compared to deterministic precipitation forecasts of COSMO-DE. The results show a significantly improved quality of the short-term precipitation forecasts and great potential to improve flood warnings for urban catchments. The combined ensemble forecasts are produced operationally every 5 min. Applications involve the Flood Warning Service Hamburg (WaBiHa) and real-time hydrological simulations with the model KalypsoHydrology.

**Key words** | COSMO-DE-EPS, ensemble runoff forecast, operational flood warning, radar nowcast ensembles, real-time runoff simulations, urban flood management

**Alrun Jasper-Tönnies** (corresponding author)  
**Thomas Einfalt**  
**Alexander Strehz**  
hydro & meteo GmbH & Co. KG,  
Breite Str. 6-8, Lübeck D-23552,  
Germany  
E-mail: [jasper-toennies@hydrometeo.de](mailto:jasper-toennies@hydrometeo.de)

**Sandra Hellmers**  
**Peter Fröhle**  
Institute of River and Coastal Engineering,  
Hamburg University of Technology,  
Hamburg,  
Germany

### INTRODUCTION

The high variability of local intense rainfall events and the short response time of flow in urban catchments are a challenging task for short-term rainfall forecasts and flood warnings (Germann *et al.* 2006; Achleitner *et al.* 2009; Schellart *et al.* 2014). Recent studies describe the forecast quality of different nowcast methods, numerical weather prediction (NWP), and combined forecasts (Atencia *et al.* 2010; Kober *et al.* 2012; Liguori *et al.* 2012). A number of studies illustrate the benefit of using NWP ensembles for run-off simulations and flood warnings in small to medium size catchments (e.g. Addor *et al.* 2011; Alfieri *et al.* 2012). Smith *et al.* (2014) use radar based ensemble nowcasts for run-off forecasts in the Alpine Verzasca catchment and evaluate the forecasts with regard to flash flood warnings. A joint evaluation using nowcasts, NWP and combined forecasts for run-off simulations in an urban catchment in England was performed by Schellart *et al.* (2014). Two urban flash floods in a small city in Denmark are analysed by Thorndahl *et al.* (2016) using radar nowcasts and numerical weather predictions for hydrological simulations. In the presented study, we concentrate on short-term forecasts of convective precipitation and on the usefulness of the forecasts for an urban flood warning system.

A key objective of this study is to improve short-term forecasts of heavy rainfall by combining ensembles of radar nowcasts with NWP ensembles. The combined forecasts are aimed at improving warnings of the Flood Warning Service Hamburg (WaBiHa). The presented work is part of the German research project StucK 'Long term drainage management of tide-influenced coastal urban areas with consideration of climate change'. While the research project addresses the issue of waters influenced by both tides and rainfall, in this study we focus on warnings for urban rivers not influenced by tides.

We present the results of an initial study evaluating three months of precipitation forecasts and measurements from summer 2016 and a case study of runoff simulations using ensemble precipitation forecasts as input to a hydrological model. The results of one convective precipitation event are presented in more detail as an example. The results of the initial study are used to design the build-up of the combined forecasts. They are also used to investigate which information of the forecast ensemble is most suitable as an input for the flood warning system WaBiHa. The combined precipitation forecasts are tested based on 3 months of forecasts,

radar measurements, and water gauge measurements from 2017.

## METHODS AND MATERIALS

### COSMO-DE-EPS

Numerical weather predictions of precipitation are produced by the German Weather Service (DWD). COSMO-DE/COSMO-DE-EPS are non-hydrostatic and convection-permitting models with a horizontal resolution of  $0.025^\circ$  (2.8 km). The deterministic forecasts COSMO-DE (Baldauf *et al.* 2011) and the ensemble forecasts COSMO-DE-EPS (Kühnlein *et al.* 2014) with 20 ensemble members are provided every 3 h with a lead time of 27 h.

### Ensemble precipitation nowcasts

The nowcasts are based on current measurements:

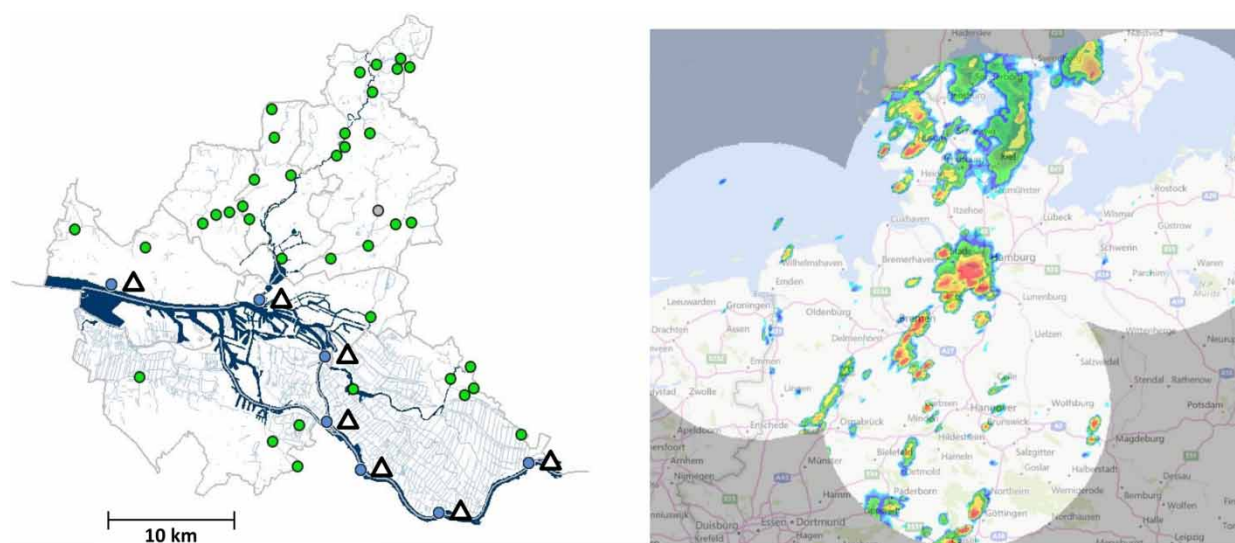
- Radar data from four DWD radar stations: Boostedt, Rostock, Hannover and Emden. The radar product (DX) is a PPI with a resolution of  $1 \text{ km} \times 1^\circ$  and 5 min.
- Rain gauge measurements of 400 stations (Hamburg/Northern Germany, time step 5 min–1 h).

The radar data are processed and corrected with the software SCOUT (hydro & meteo 2009), using a number of filter and correction methods specified in the scope of

the EU-project VOLTAIRE (Golz *et al.* 2006). The rain gauge measurements are used for adjusting the radar measured rainfall sums at every radar time step, based on data of the past 3 h. A composite with a resolution of  $1 \text{ km} \times 1 \text{ km}$  is produced every 5 min from the data of the four radar stations. An example of the adjusted composite is shown in Figure 1.

The radar composites of the last 30 min serve as input for calculating nowcasts with a lead time of 1–3 h. The nowcast method is based on a cell-tracking algorithm, tracking rain cells with a minimum size of 20 grid points above a reflectivity threshold. Displacement vectors are calculated for cells which are tracked based on their position, size and shape. These vectors are used to approximate the 2D advection field. A Semi-Lagrange scheme is used to forecast the advection of the rainfall field. Additionally, growth and decay of rain cells are extrapolated. The method builds on the nowcast described by Tessendorf & Einfalt (2012).

Ensemble nowcasts are generated by varying the initial forecast variables (displacement vectors and growth rates) depending on the variance of the observed variables during the past 30 min. The variables are randomly varied assuming normal probability distributions in order to obtain individual values for each ensemble run (Tessendorf & Einfalt 2012). The effects of additional sources of uncertainty, such as measurement errors and changing weather conditions, are approximated by two empirical parameters increasing the width of the probability distributions.



**Figure 1** | Left: Location of the water gauges included in the warning system WaBiHa, Hamburg. The gauges are denoted by circles. Tide-influenced water gauges which are processed by the Hamburg Port Authority are also plotted (triangles). Right: Operational North German composite of the four radars Boostedt, Emden, Hannover, Rostock in dBZ during a heavy rain event at 28.08.2016 14:00 UTC.

Combined precipitation forecasts are produced from the numerical weather predictions and the ensemble nowcasts. These are constructed in dependence of the forecast lead time:

0–2 h: Nowcast ensembles: 10 ensemble runs with a time step of 5 min.

2–4 h: Blending: Nowcasts and numerical weather predictions are merged. The nowcasts are aggregated to hourly precipitation sums and the COSMO-DE-EPS forecasts are mapped to the  $1 \text{ km} \times 1 \text{ km}$  grid of the radar data in order to obtain data with the same temporal and spatial resolution. The blended forecasts are determined as a simple weighted mean of the nowcasts and the COSMO-DE-EPS forecasts as described e.g. by Golding (1998):

$$R_{\text{blend}}(x_i) = w_{nc} \cdot R_{nc}(x_i) + w_{NWP} \cdot R_{NWP}(x_i)$$

where  $R_{\text{blend}}$  is the blended precipitation forecast,  $R_{nc}$  the nowcast and  $R_{NWP}$  the COSMO-DE-EPS forecast at the grid position  $x_i$ . The weighting factors  $w_{nc}$  and  $w_{NWP}$  are derived as a function of the lead time from an evaluation of the forecast quality based on 3 months of data: for the 3rd hour  $w_{nc} = 1/3$  and  $w_{NWP} = 2/3$ , and for the 4th hour:  $w_{nc} = 1/4$  and  $w_{NWP} = 3/4$ , if both forecasts are available at the position  $x_i$ . For points which are outside of the radar range at the start time of the nowcast but shifted into the investigation area, the factors change to  $w_{nc} = 0$  and  $w_{NWP} = 1$ , respectively.

4–20 h: NWP: The numerical weather predictions COSMO-DE-EPS are mapped on the radar grid.

The number of nowcast ensemble members is restricted to 10 as this number of runs is realised in almost real-time. The 10 nowcasts are used twice and combined with the 20 COSMO-DE-EPS runs, resulting in 20 different ensemble runs for lead times longer than 2 h. The forecasts are updated every 5 min with new radar nowcasts and every 3 h with new numerical weather predictions. The results of all ensemble members are provided as time series with a time step of 5 min in the first two hours and 1 h afterwards, and as aggregated forecast sums.

## FORECAST APPLICATIONS

### Warning system WaBiHa

The warning system WaBiHa (<http://www.wabiha.de>), run by the State Agency of Roads, Bridges and Waters (LSBG), is an operational warning system in the city

Hamburg. The system indicates the warning level for a number of urban rivers at the location of the water level gauges (see Figure 1). At every gauge, two warning thresholds are specified by the LSBG: the ‘moderate warning’ level, corresponding to some overflow, and the ‘high warning’ level, corresponding to larger flooding. Site-specific flood warnings are produced when the gauge measured water levels exceed the defined thresholds. In addition, numerical weather predictions of COSMO-DE serve as an input into the warning system. Warnings based on the precipitation forecasts indicate a risk for increased water gauge levels and apply to the entire city. They are produced evaluating standardized thresholds of accumulated precipitation over 1, 6, and 12 h corresponding to first exceedances of warning levels at the water gauge sites.

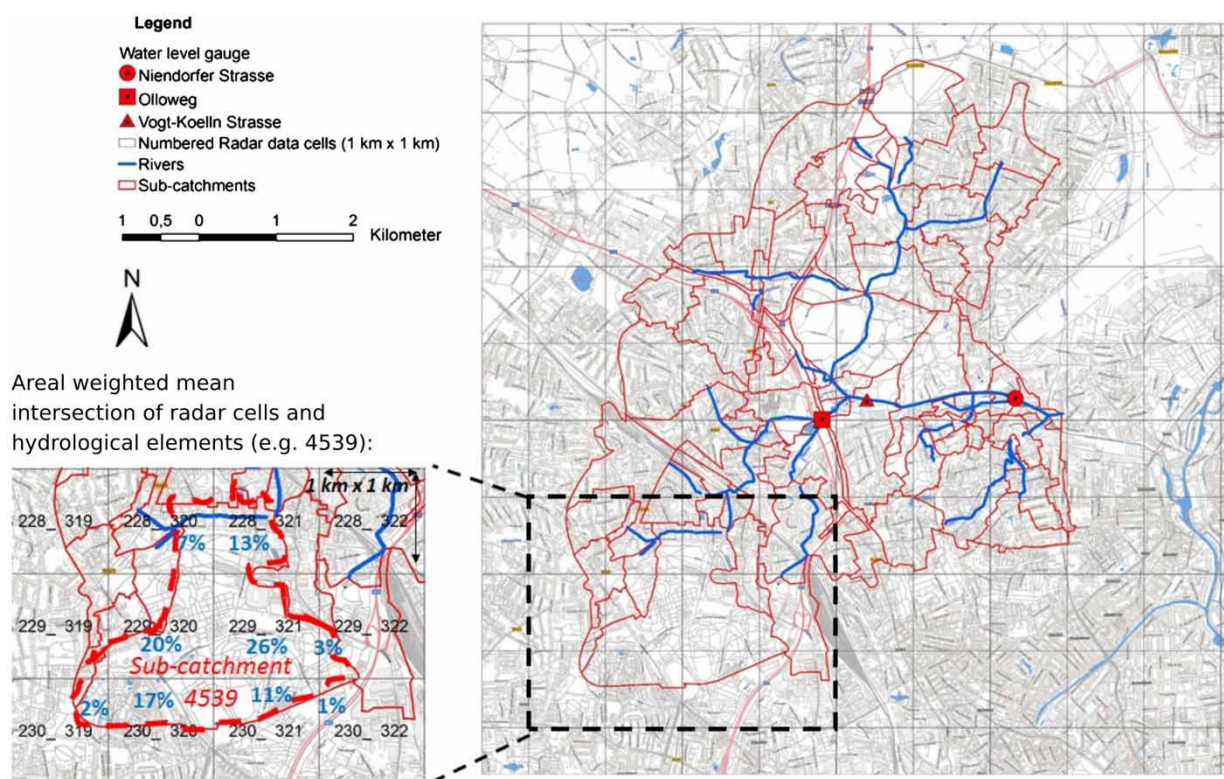
Within the project StucK, the use of combined ensemble forecasts for the warning system WaBiHa is evaluated. The replacement of the COSMO-DE forecasts by combined ensemble forecasts is one of several envisaged steps in order to improve the warning system.

### Ensemble flood hydrographs with KalypsoHydrology

The ensemble forecasts are also used as an input for hydrological models. For the eight main river catchments in Hamburg hydrological models are implemented using the software application KalypsoHydrology. Results of the runoff simulations are intended to supplement the information of the warning system WaBiHa at the water level gauges. The precipitation forecasts are processed in the semi-distributed rainfall runoff model KalypsoHydrology. The hydrological model supports the simulation of hydrological processes on the surface and the main subsurface processes: snow, evapotranspiration, evaporation from water surfaces in retention ponds, soil moisture, interflow, baseflow, and groundwater flow processes (TUHH 2013; Hellmers & Fröhle 2017). The first hydrological model extended to an operational application is the model for the Kollau catchment (see Figure 2) using the calculation core KalypsoNA (Version 3.1.1). The catchment with a high imperviousness due to industrial spaces, large infrastructures and dense residential areas has a size of  $33.6 \text{ km}^2$  and a lag time of approximately 1 h.

In order to integrate the spatially distributed precipitation data into the model an intersection of the radar grid ( $1 \text{ km} \times 1 \text{ km}$ ) with the spatial elements of the semi-distributed hydrological model is performed. For each subcatchment, the precipitation is calculated as an area-weighted mean of the overlaying radar grid cells (see the example in Figure 2).





**Figure 2** | Map of the Kollau catchment (33.6 km<sup>2</sup>), location of water level gauges and illustration of the intersection of precipitation data with hydrological elements.

Further input parameters are temperature, wind, sunshine duration, and relative humidity (weather station measurements).

In order to obtain the initial soil moisture and ground water conditions the past 3 years are simulated with a daily time step. In the operational model, the initial soil moisture and water balance computations are performed each day at 3:00 UTC. The operational forecast model starts 4 days before the current time, using radar measured precipitation data as well as ensemble forecasts for the future time steps as input. For each forecast ensemble member the flood hydrograph is computed at all nodes of the river network. The flood hydrographs are transferred to water levels to enable a comparison with the data of the water level gauges.

In the period November 2015 until September 2016, four precipitation events were analysed in detail (November 2015, February 2016, May 2016 and August 2016). These four precipitation events led to an increase of the water level above the moderate warning level at the downstream gauge Niendorfer Str. For these four events, a parameter study was carried out which showed good agreement between the hydrological model results and the water level gauge measurements (Mikkelsen 2017).

The data of three water level gauges in the Kollau catchment are available online for urban flood management with a temporal resolution of 5 min. The operational model for the Kollau catchment is running in a test phase since mid 2017. Simulations using the current ensemble precipitation forecasts with 10 ensemble members as input are conducted and updated every 15 min. The simulated hydrographs at the three water level gauges are presented to the StucK project partners on a webpage.

## RESULTS AND DISCUSSION

### Evaluation of the forecast quality of nowcasts and NWP forecasts

Convective heavy precipitation is difficult to predict because of the high variability of the rainfall field. With life times of convective cells often shorter than 30 min, it is challenging to predict the distribution and amount of rainfall even for nowcasts with lead times up to 2 h. On the other hand, for flood warning systems like the system WaBiHa in Hamburg, it is not absolutely essential that the place and the time of

the maximum precipitation are predicted precisely. The crucial criterion for initiating a rainfall warning is a threshold exceedance at at least one place in the city area. The requirements of the warning system WaBiHa are considered in the subsequent evaluation.

A joint evaluation of the forecast quality of the ensemble nowcasts (10 ensemble members), the deterministic COSMO-DE and the ensemble forecasts COSMO-DE-EPS (20 ensemble members) was performed. The rainfall forecasts were compared to corrected and adjusted radar measurements. The period of the evaluation is 1 June–31 August 2016. This summer period was characterised by a high number of heavy precipitation events in Northern Germany.

The evaluation is done using two thresholds for mean precipitation over an area of  $5\text{ km} \times 5\text{ km}$ , accumulated over 1 h. The 7 mm threshold is relevant as precipitation above this value can lead to first exceedances of warning levels at the Hamburg water gauges. The 3 mm threshold is evaluated in addition to account for a larger number of precipitation events. In the evaluation period, the 3 mm threshold was exceeded in the radar measurements in 108 cases and the 7 mm threshold in 39 cases, corresponding to 4.9% and 1.8% of the evaluated time steps. All 39 exceedances of the 7 mm threshold were due to convective precipitation events, thereof 17 cases were part of a larger-scale convective pattern caused by a weather front or convergence zone.

The forecast quality indices for rare events *hit rate* (also: *probability of detection*) and *false alarm ratio* (Donaldson et al. 1975) are used. Evaluations are done for the Hamburg city area approximated by a square of  $25\text{ km} \times 25\text{ km}$  divided into 25 grid points of  $5\text{ km} \times 5\text{ km}$ . Two spatial scales are investigated:

- (1) City area: A forecast is counted as *hit* if an hourly precipitation sum above the threshold is predicted and measured for at least one  $5\text{ km} \times 5\text{ km}$  grid point within the city area.
- (2)  $5\text{ km} \times 5\text{ km}$ : A forecast is counted as *hit* if an hourly precipitation sum above the threshold is predicted and measured at the same  $5\text{ km} \times 5\text{ km}$  grid point. (See next section.)

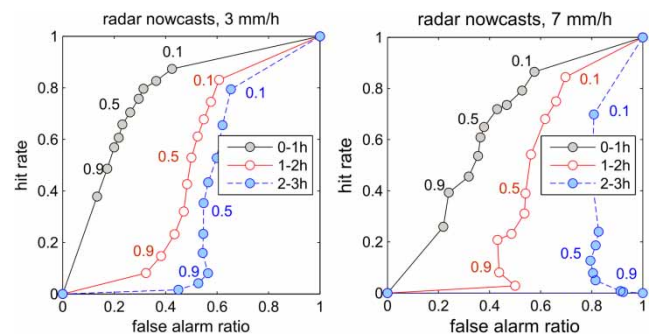
In each case, a forecast above the threshold without a corresponding measurement is counted as a *false alarm* and a measurement which was not forecasted as a *miss*. The *hit rate* is defined as  $\text{hits}/(\text{hits} + \text{misses})$  and the *false alarm ratio* as  $\text{false alarms}/(\text{hits} + \text{false alarms})$ . Except for the first hour of the nowcast, a temporal shift of the forecast by  $\pm 1\text{ h}$  is tolerated, i.e. a forecast above the threshold is still

counted as *hit* if the actual event occurs 1 h earlier or later than predicted.

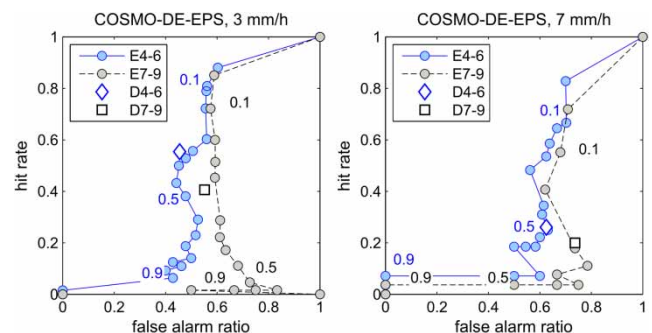
In Figures 3–5, the *hit rate* is plotted against the *false alarm ratio*, hence the perfect forecast would be found in the upper left corner (*hit rate* = 1, *false alarm ratio* = 0) and a random forecast at the lower right. The ensemble quantiles – which correspond to exceedance probabilities predicted by the ensemble forecast – are indicated.

Figure 3 illustrates the forecast quality of the ensemble nowcasts with lead time 3 h, for the 1st, 2nd, and the 3rd hour of the forecast period separately. The first hour of the nowcast evaluated with a threshold of 3 mm shows the best forecast quality, with *hit rates* above 0.8 and *false alarm ratios* below 0.4 for low ensemble quantiles. The forecast quality rapidly decreases for the 2nd and the 3rd hour. Regarding the 7 mm threshold, the forecast quality is lower for all ensemble quantiles. However, a high *hit rate* is still achieved for low ensemble quantiles.

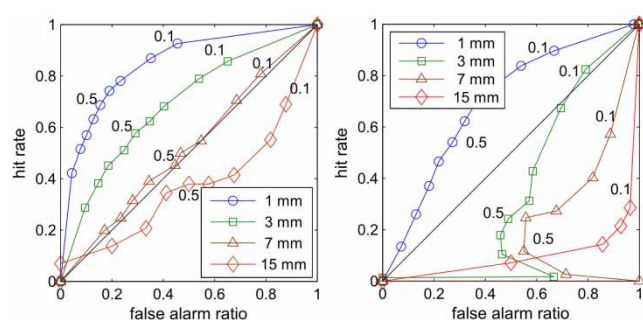
Figure 4 shows the results of the COSMO-DE and COSMO-DE-EPS forecasts for the lead times 4–6 h and 7–9 h. The deterministic COSMO-DE forecasts perform better than the median of the COSMO-DE-EPS forecasts;



**Figure 3** | Hit rate against false alarm ratio of radar nowcast ensembles for the 1st, 2nd, and the 3rd hour of the nowcast evaluating the forecast skill to predict hourly precipitation above 3 mm (left) and 7 mm (right). The investigated spatial scale is the city area. Circles indicate quantiles of the ensemble forecast.



**Figure 4** | Same as Figure 3 but for COSMO-DE-EPS (E4-6 and E7-9) and the deterministic COSMO-DE (D4-6, D7-9) for the lead times 4–6 h and 7–9 h.



**Figure 5** | Hit rate against false alarm ratio of radar nowcasts for the 1st hour (left) and the 2nd hour (right) of the nowcast evaluating the forecast skill to predict precipitation above different thresholds between 1 mm and 15 mm accumulated over 1 h. The investigated spatial scale is 5 km × 5 km. Markers indicate quantiles of the ensemble forecast.

however, the best results are obtained by the lowest ensemble quantiles of COSMO-DE-EPS. In contrast to the nowcasts, the forecast quality of COSMO-DE-EPS decreases only slightly with the lead time. Since the computing and delivery time of the COSMO-DE-EPS is approximately 3 h, it is reasonable to compare the COSMO-DE-EPS with lead time 4–6 h to the radar nowcasts with lead time 1–2 h when rating the value of the respective forecasts for a warning system. When regarding all ensemble quantiles and both rain thresholds, the forecast quality of the nowcasts is better compared to COSMO-DE-EPS for the first two hours. For lead times of 3 h and longer, COSMO-DE-EPS shows the better results. A common result which is found, especially for the COSMO-DE-EPS but also for the nowcast ensembles, is that the *hit rate* is best for low ensemble quantiles and decreases significantly for higher quantiles. This is partially accompanied by a decrease of the *false alarm ratio*, but to a minor degree. As a conclusion for the warning system WaBiHa, on average, the lowest ensemble quantile seems to provide the best information for issuing warnings.

### Precipitation forecasts at the catchment scale

The ensemble nowcasts are evaluated on the spatial scale 5 km × 5 km which can be considered as representative for the Kollau catchment. Different precipitation thresholds between 1 mm and 15 mm accumulated over 1 h are investigated (period 1 June–31 August 2016). The results of the 1st and 2nd hour of the nowcast are shown in Figure 5. The results state a generally lower hit rate and higher false alarm ratio compared to the city scale (Figure 3). The forecast quality strongly decreases with increasing precipitation thresholds depending on the lead time. From the 1st to the 2nd hour the forecast quality remains relatively high for the

1 mm threshold but strongly declines for the high thresholds 7 mm and 15 mm. These relations illustrate the particular difficulties of flood warnings for small urban catchments which are especially sensitive to convective precipitation with high rain intensities. However, a forecast quality is achieved which may still be useful for precipitation thresholds of 7 mm/h or higher if the lowest ensemble quantile is regarded.

### Simulation of a convective precipitation event

The runoff simulations of the convective precipitation event on 28 August 2016 in the Kollau catchment are evaluated in more detail as an example for using the ensemble nowcasts as model input. The precipitation came from two intensive rain cells which moved over the catchment between 14:00 and 16:00 UTC. The rest of the day was nearly dry with accumulated precipitation below 1 mm. Figure 6 shows the accumulated precipitation measured by the radar composite and the 2 h ensemble nowcasts (ENS 1 – ENS 10) initiated at 14:00 UTC.

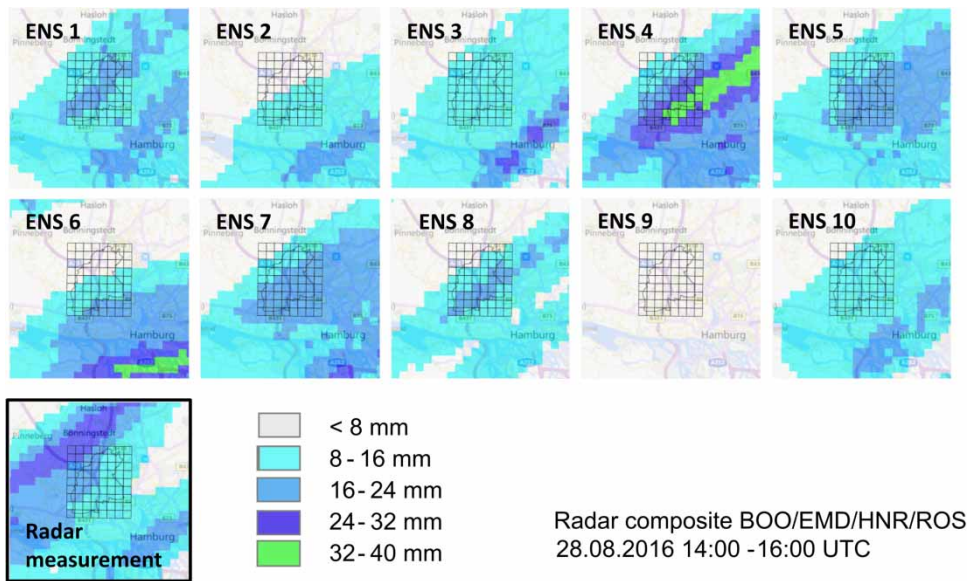
The simulated flood hydrographs and gauge measurements of the precipitation event are presented in Figure 7. The ensemble simulations are based on the radar measurements before 14:00 UTC and on ensemble nowcasts of 14:00 UTC for 14:00–16:00 UTC (dashed lines/quantiles). After 16:00 UTC, the precipitation input to the model was set to zero. The hydrograph based on radar measurements for the full event is also shown ('Water level Sim'). The moderate warning level of 620 cm above mean sea level (NHN + cm) is exceeded at the gauge station at 19:45 UTC. The exceedance of the warning level is simulated by 5 of the 10 ensemble members.

In this example, the ensemble precipitation forecasts enable an early warning and help to quantify the uncertainty of the predicted runoff. However, for the small Kollau catchment the spatial uncertainty of the precipitation forecasts has a stronger effect on the forecast quality compared to larger catchments and compared to the system WaBiHa where precipitation warnings are evaluated over the whole city area. The results of the evaluation suggest that, at this scale, only nowcasts with a short lead time of 1–2 h may lead to useful results in case of heavy convective events. They also illustrate the specific gain of the ensemble compared to a single nowcast for warning purposes.

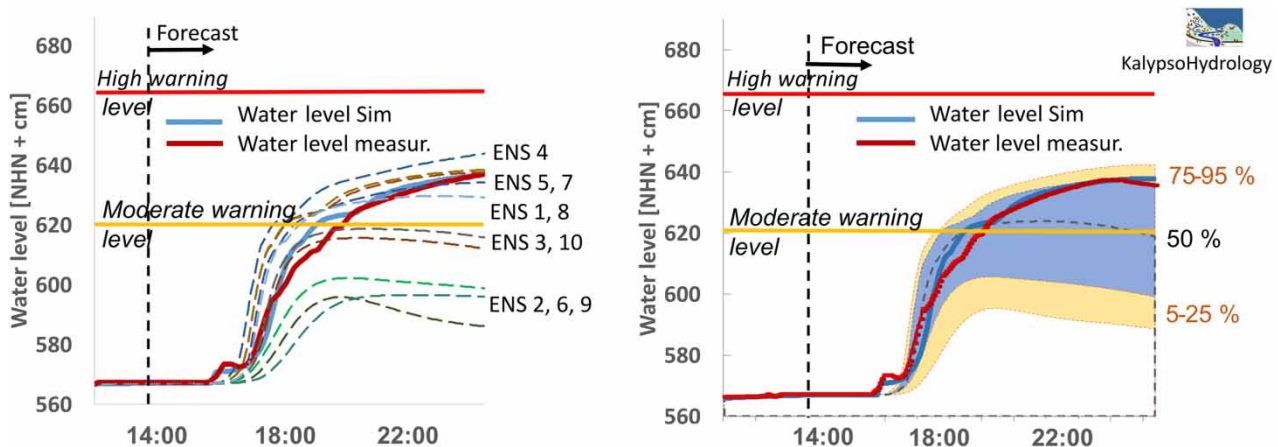
### Evaluation of the combined precipitation forecasts and WaBiHa warnings

The combined forecasts and the resulting gauge warnings were evaluated in an independent test period over 3





**Figure 6** | Ensemble nowcasts of 28.08.2016 14:00 UTC (10 runs) over 2 h and the accumulated precipitation measured by the radar composite from 14:00–16:00 UTC.



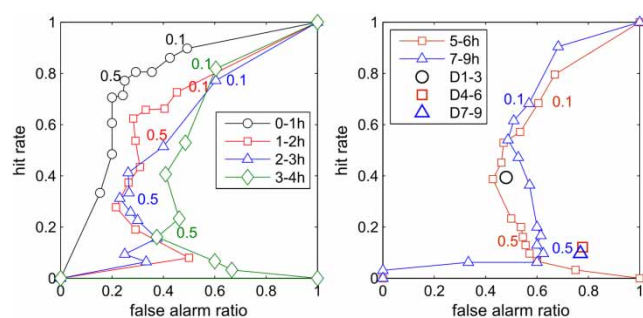
**Figure 7** | Flood hydrographs in cm NHN of the rainfall event 28 August 2016 in the Kollau catchment. The results based on 10 ensemble nowcasts initiated at 14:00 UTC are shown as scatter plot (left) and as shaded quantile plot (right). In addition, flood hydrographs based on radar measurements ('Water level Sim') and gauge measurements at the station Niendorfer Str. ('Water level measur.') are plotted.

months (22 March–22 June 2017). A comparison to radar measurements and water gauge measurements was conducted using the same thresholds as in the first evaluation. In the investigated period the 3 mm threshold was exceeded in 32 cases and the 7 mm threshold in 11 cases. In comparison to the summer period 2016, more large-scale precipitation events are included in the sample. From the 11 events with exceedances of the 7 mm threshold, one event was a small-scale convective event, five events were part of a larger-scale convective pattern and five events were large-scale precipitation events related to weather fronts.

### Combined forecasts

The forecast results of *hit rate* against *false alarm ratio* of the combined forecasts are displayed in Figure 8 for the 3 mm threshold. The left figure shows the results for lead times of 1–4 h (nowcasts and blended forecasts) and the right figure for lead times of 5–9 h (from COSMO-DE-EPS). Also plotted are the results of the COSMO-DE forecast for different lead times (1–3 h, 4–6 h, 7–9 h). The *hit rate* of the combined forecasts is relatively high for all lead times, with values above 0.7 for the lowest ensemble quantile. The expected decrease of the forecast quality with the lead



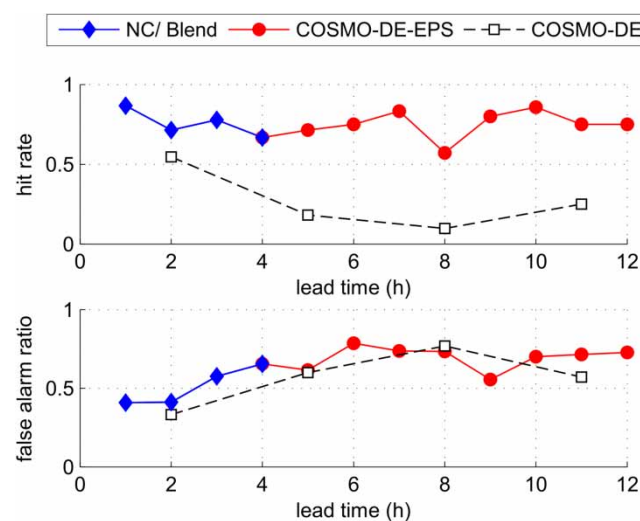


**Figure 8** | Hit rate against false alarm ratio of the combined forecasts for the 1st–4th hour (left) and for the 5th–9th hour of the forecast (right) for the threshold 3 mm/h. In addition, the results of the COSMO-DE forecast ('D1-3', 'D4-6', 'D7-9') are plotted in the right figure for lead times of 1–9 h.

time is mainly found in an increase of the *false alarm ratio*. The forecast quality of the blended forecasts (3rd and 4th hour) is comparable to the COSMO-DE-EPS forecasts with a lead time of 5–6 h but with slightly lower *hit rates* and *false alarm ratios*. The combined forecasts clearly outperform the COSMO-DE forecast for every lead time.

### Wabiha warnings

All 11 events with exceedances of the 7 mm precipitation threshold for at least one 5 km × 5 km area within the city area coincided with exceedances of the moderate warning level at one or several water level gauges in the warning system WaBiHa. Exceedances of the warning levels at the water gauges occurred in four more cases which were not due to convective precipitation and which are not further addressed here. The warnings the combined ensemble forecasts would have produced in the warning system WaBiHa during the three month period are analysed. A warning is generated if the maximum of all ensemble members (i.e. the lowest ensemble quantile) exceeds the 7 mm threshold within the city area. The results are displayed in Figure 9 in terms of *hit rate* and *false alarm ratio* and compared to the warnings based on COSMO-DE. The *hit rate* of the combined forecasts is considerably higher than the *hit rate* of COSMO-DE for almost all lead times and the *false alarm ratio* is similar to the *false alarm ratio* of COSMO-DE. The relatively small number of 11 threshold exceedances during the investigated period does not allow for robust quantitative results. However, the large difference between the forecast quality of the combined forecasts and COSMO-DE confirms the results of the first evaluation period and suggests that the warning system WaBiHa will be significantly improved by using the combined ensemble forecasts.



**Figure 9** | Hit rate and false alarm ratio of the lowest ensemble quantile of the combined forecasts and of the COSMO-DE forecast for lead times of 0–12 h, based on exceedances of the threshold 7 mm/h within the city area in the period 22 March–22 June 2017.

### CONCLUSIONS

Combined ensemble precipitation forecasts are produced based on ensemble nowcasts and COSMO-DE-EPS forecasts. In the investigation periods the single products as well as the combined forecasts clearly outperform the COSMO-DE forecast with regard to short-term convective precipitation events. This is due to the effects of using nowcasts instead of numerical weather predictions in the first two hours and of using the ensembles instead of a single forecast which has a large effect on the *hit rate* at most of the investigated lead times. However, the best results are obtained analysing the lowest ensemble quantile and little potential is found to reduce the *false alarm ratio* by using a higher quantile of the ensemble forecasts instead of a single forecast. This finding may be relevant as well for other warning purposes with a focus on short-term convective events. The results suggest that the combined forecasts will markedly improve short-term rainfall forecasts and flood warnings for the Flood Warning Service Hamburg. At present, the precipitation thresholds in the warning system are relatively low. This way, the automated warnings indicate an increased flood risk and show a specific need for a more intensive and frequent monitoring of the weather situation. The warning information is supplemented by the latest radar measurements and water gauge measurements which the person in charge can consult in order to assess the actual flood risk. Individual precipitation thresholds

for different catchments and water gauges corresponding to moderate and high warning levels are planned to be added to the system in the future. The next steps will be to integrate the combined precipitation forecasts into the operational flood warning system.

A real-time implementation of runoff simulations with the software KalypsoHydrology for the Kollau catchment using the ensemble precipitation forecasts was realised in 2017. After a test phase with the operational system the results of the runoff simulations are envisaged to supplement the warning information in the system WaBiHa and to further improve warnings at the simulated gauges. Further on, the developed methodology can be extended to the other seven existing KalypsoHydrology models for the main river catchments in Hamburg.

## ACKNOWLEDGEMENTS

StucK (2015–2018; [www.stuck-hh.de](http://www.stuck-hh.de)) is a joint project in the funding measure ‘Regional Water Resources Management for Sustainable Protection of Waters in Germany’ (ReWaM) sponsored by the German Federal Ministry of Education and Research (BMBF). The authors gratefully acknowledge this support.

## REFERENCES

- Achleitner, S., Fach, S., Einfalt, T. & Rauch, W. 2009 [Nowcasting of rainfall and of combined sewage flow in urban drainage systems](#). *Water Science and Technology* **59** (6), 1145–1151.
- Addor, N., Jaun, S., Fundel, F. & Zappa, M. 2011 [An operational hydrological ensemble prediction system for the city of Zurich \(Switzerland\): skill, case studies and scenarios](#). *Hydrology and Earth System Sciences* **15** (7), 2327.
- Alfieri, L., Thielen, J. & Pappenberger, F. 2012 [Ensemble hydro-meteorological simulation for flash flood early detection in Southern Switzerland](#). *Journal of Hydrology* **424**, 143–153.
- Atencia, A., Rigo, T., Sairouni, A., More, J., Bech, J., Vilaclara, E., Cunillera, J., Llasat, M. C. & Garrote, L. 2010 [Improving QPF by blending techniques at the meteorological service of Catalonia](#). *Natural Hazards and Earth System Sciences* **10**, 1443–1455.
- Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M. & Reinhardt, T. 2011 [Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities](#). *Monthly Weather Review* **139** (12), 3887–3905.
- Donaldson, R. J., Dyer, R. M. & Krauss, M. J. 1975 An objective evaluator of techniques for predicting severe weather events. In: *Preprints: 9th Conf. Severe Local Storms, Norman, Oklahoma*. American Meteorological Society, pp. 321–326.
- Germann, U., Zawadzki, I. & Turner, B. 2006 [Predictability of precipitation from continental radar images. Part IV: limits to prediction](#). *Journal of the Atmospheric Sciences* **63** (8), 2092–2108.
- Golding, B. 1998 [Nimrod: a system for generating automated very short range forecasts](#). *Meteorological Applications* **5**, 1–16.
- Golz, C., Einfalt, T. & Galli, G. 2006 [Radar data quality control methods in VOLTAIRE](#). *Meteorologische Zeitschrift* **15** (5), 497–504.
- Hellmers, S. & Fröhle, P. 2017 [Integrating local scale drainage measures in meso scale catchment modelling](#). *Water* **9** (2), 71.
- hydro & meteo 2009 *The SCOUT Documentation, Version 3.30*. Lübeck, Germany.
- Kober, K., Craig, G. C., Keil, C. & Dörnbrack, A. 2012 [Blending a probabilistic nowcasting method with a high-resolution numerical weather prediction ensemble for convective precipitation forecasts](#). *Quarterly Journal of the Royal Meteorological Society* **138** (664), 755–768.
- Kühnlein, C., Keil, C., Craig, G. C. & Gebhardt, C. 2014 [The impact of downscaled initial condition perturbations on convective-scale ensemble forecasts of precipitation](#). *Quarterly Journal of the Royal Meteorological Society* **140** (682), 1552–1562.
- Liguori, S., Rico-Ramirez, M. A., Schellart, A. N. A. & Saul, A. J. 2012 [Using probabilistic radar rainfall nowcasts and NWP forecasts for flow prediction in urban catchments](#). *Atmospheric Research* **103**, 80–95.
- Mikkelsen, A. K. L. 2017 *Analysis of Hydrological Parameters and Modelling the Impacts of Land use Change, Climate Change and SUDS Implementations in an Urban Catchment*. Master Thesis, Institute of River and Coastal Engineering, Hamburg University of Technology, Hamburg, Germany.
- Schellart, A., Liguori, S., Krämer, S., Saul, A. & Rico-Ramirez, M. 2014 [Comparing quantitative precipitation forecast methods for prediction of sewer flows in a small urban area](#). *Hydrological Sciences Journal* **59** (7), 1418–1436.
- Smith, P. J., Panziera, L. & Beven, K. J. 2014 [Forecasting flash floods using data-based mechanistic models and NORA radar rainfall forecasts](#). *Hydrological Sciences Journal* **59** (7), 1403–1417.
- Tessendorf, A. & Einfalt, T. 2012 Ensemble radar nowcasts – a multi-method approach. *International Association of Hydrological Sciences Publications* **351**, 305–310.
- Thorndahl, S., Nielsen, J. E. & Jensen, D. G. 2016 [Urban pluvial flood prediction: a case study evaluating radar rainfall nowcasts and numerical weather prediction models as model inputs](#). *Water Science and Technology* **74** (11), 2599–2610.
- TUHH 2013 *Software-Entwicklung Niederschlag-Abfluss-Modell KalypsoNA (Software Development Rainfall Runoff Model KalypsoNA)*. <http://www.tuhh.de/wb/forschung/software-entwicklung/kalypso/kalypso-na.html> (accessed 20 September 2017).