RESEARCH ARTICLE

A percentile-based approach to rainfall scenario construction for surface-water flood forecasts

Steven J. Böing[1] 🔵 | Cathryn E. Birch[1] | Benjamin L. Rabb[1] | Kay L. Shelton[2]

[1]School of Earth and Environment, University of Leeds, Leeds, UK

[2]JBA Consulting, Skipton, UK

Correspondence
Steven J. Böing, School of Earth and Environment, University of Leeds, LS2 9JT Leeds, UK.
Email: s.boeing@leeds.ac.uk

Funding information
Natural Environment Research Council, Grant/Award Number: NE/P011160/1

Abstract

A novel technique to produce reasonable worst-case rainfall scenarios from ensemble forecasts is presented. This type of scenario is relevant for predicting the risk of localized, intense rainfall events with a duration between 15 min and several hours. Such rainfall events can cause surface-water (pluvial) flooding. Producing useful forecasts of these events at lead times of more than a few hours is challenging due to the precision and accuracy in rainfall intensity, duration and location that is required. The technique described here addresses these challenges by constructing appropriate scenarios using a neighbourhood technique in combination with ensemble forecasting. It is similar to the distance-dependent depth–duration analysis described in earlier studies, but it introduces an additional post-processing step based on probability distribution functions of rainfall accumulation near a location of interest. This additional step makes the reasonable worst-case scenarios less dependent on grid-scale behaviour, and helps to generate scenarios with a consistent interpretation. The method is used to compare forecasts with a lead time of 6–36 hr to radar data for several case studies that occurred in Yorkshire. These comparisons also introduce new techniques to present maps of the reasonable worst-case rainfall accumulation at each location.

KEYWORDS

flood forecasting, neighbourhood processing method, rainfall prediction

1 | INTRODUCTION

In the UK alone, it has been estimated that 3.2 million properties are at risk of surface-water flooding (Department for Environment, Food and Rural Affairs, 2018). This type of flooding often occurs as a result of intense, localized and short-lived rainfall events, particularly in summer. Although surface-water flooding is more localized than fluvial floods, it is likely to affect predominantly urban areas. It often takes place away from river and coastal flood plains, but sometimes occurs in addition to other types of flooding. Estimated annual losses in the UK as of 2012 were £320 million (Committee on Climate Change, 2012), and the intensity of UK heavy rainfall during summer and autumn is expected to increase in the latest climate change scenarios (Met Office, 2019). Surface-water flooding can present a serious hazard due to high peak flow rates, which can develop over a short time interval (Archer and Fowler, 2018).

Since the advent of convection-permitting weather forecasts, which explicitly model the dynamics of convective clouds, numerical weather prediction (NWP) models have had a much more realistic representation of extreme rainfall events which may lead to pluvial flooding. Compared to earlier NWP models where convection is parametrized (modelled in terms of larger-scale variables), convection-permitting models are better at capturing the location and timing of the systems in which individual convective cells are embedded, the characteristics of storms within such systems, and the influence of the land surface and topography on these cells (e.g., Prein *et al.*, 2015; Clark *et al.*, 2016). These models also better represent the diurnal cycle of convective rainfall, have more realistic frequency statistics of extreme precipitation, and differ in their predictions for the sensitivity of heavy rainfall to climate change (e.g., Hohenegger *et al.*, 2008; Kendon *et al.*, 2014; Chan *et al.*, 2018).

Uncertainty about forecast precipitation remains the key challenge for accurate predictions of flood risk (Hapuarachchi *et al.*, 2011). Despite recent advances, the location of individual storm cells continues to be largely unpredictable in convection-permitting forecasts (Collier, 2007; Hohenegger and Schär, 2007). Only when events are detected on radar images can a much improved prediction of where the most intense rainfall might actually occur be given, although even then the development of the intensity of a cell is hard to predict.

To represent the uncertainty in both the behaviour of storm systems and the location of convective cells that individual deterministic forecasts provide, an ensemble forecast can be used (e.g., Murphy and Palmer, 1986). Ensemble forecasts improve the consistency between forecasts with different lead times and provide information about the probability of occurrence of events (Buizza, 2008). They have also been widely used to drive hydrological models; in this case, a key challenge is to represent and assess the combined uncertainty in rainfall and hydrology (Demeritt *et al.*, 2007; Cloke and Pappenberger, 2009). Members of an ensemble forecast can vary in their initial conditions and the choices made for physical parameters. Additional perturbations can be made to the atmospheric boundary-layer state or the output from physical parametrizations (Palmer, 2001; Berner *et al.*, 2011; Bouttier *et al.*, 2012).

Over the last decade, it has become possible to perform convection-permitting ensemble simulations for regional NWP. Although these ensembles include a realistic representation of storm dynamics, they also present new challenges. One of these challenges is that the assumptions that have traditionally been used in data assimilation to obtain initial forecast conditions are no longer valid (e.g., Clark *et al.*, 2016). In particular,

assimilating radar observations for short-term forecasting requires careful processing of the radar retrievals (Dance *et al.*, 2019). Moreover, the evaluation of ensemble forecasts requires an approach that takes into account that it is not possible to predict the exact location of individual storm cells accurately.

Traditional forecast verification methods penalize a forecast twice if a rain event is forecast correctly but with an offset in its location, a problem known as the double-penalty (Mass *et al.*, 2002). This issue can be overcome with a set of techniques that are collectively referred to as neighbourhood processing methods (Theis *et al.*, 2005; Ebert, 2008; 2009; Roberts and Lean, 2008; Gilleland *et al.*, 2009), which include smoothing of output fields and taking a maximum in a neighbouring region. Schwartz and Sobash (2017) look in detail at neighbourhood methods and show that seemingly small variations on an existing method such as using the same operators (e.g., smoothing, taking a maximum over a given area) in a different order can lead to results which require a different interpretation. Neighbourhood methods can also be used in combination with ensemble forecasting (Schwartz *et al.*, 2010).

Here, we explore an approach that has its roots in neighbourhood-based approaches but with a focus on producing plausible rainfall scenarios, rather than evaluating a forecast. We will describe a novel method for constructing and visualizing heavy rainfall scenarios. The aim is to explore a way of presenting localized scenarios that are coherent and plausible, which can be used as input for hydrological modelling (Birch *et al.*, in preparation). The work is motivated by a recent study (Ochoa-Rodríguez *et al.*, 2018) which found that users of flood guidance are interested in an accurate characterization of areas where heavy rainfall is likely to occur.

The scenarios are formed by taking time-series of precipitation from selected grid boxes in the neighbourhood around the location of interest from an ensemble of NWP forecasts. A 6–36 hr lead time is used, and the scenarios correspond to relatively high rainfall accumulations over a short interval (compared to the full neighbourhood in all ensemble members).

Neighbourhood-processed convection-permitting ensemble forecasts have previously been used in the context of flood forecasting by, for example, Vincendon *et al.* (2011), Golding *et al.* (2016), Hardy *et al.* (2016) and Olsson *et al.* (2017). Processed ensemble forecast information is often presented in terms of probability of exceedance for a given threshold, such as the probability of 1 hr mean rainfall over 30 mm·hr^{-1}. A somewhat different approach is taken in this study, where a given probability of exceedance is considered (e.g., the 95th percentile of accumulated rainfall, i.e., 5% probability of exceedance), the corresponding

threshold is derived, and scenarios that correspond to this threshold are reconstructed by using neighbourhood information. In other words, the method aims to find a local reasonable worst-case scenario with a consistent interpretation. This is not the absolute worst-case scenario that could occur, both because of the use of the percentile-based approach and because the forecast ensemble may not include the full range of possible scenarios. The approach is similar to that of Olsson *et al.* (2017), although there are two major differences to that study:

1. The novel visualization methods explored here make it possible to present forecast data for a wider region and identify areas within this wider region that are at increased risk of flooding. The percentile-based approach generates a single map summarizing rainfall totals corresponding to scenarios with a consistent interpretation for each ensemble.

2. There is more emphasis on short-duration (1 hr and below, using a running time window) events in the current work, both in terms of the analysis of the ensemble forecast as well as in comparisons to radar rainfall. Two of the case studies (August 22, 2015; August 13, 2018) in particular are characterized by highly localized convective cells that generate rainfall at these sub-hourly scales. The shorter time scales are relevant for local surface-water flooding, but less so for basin-scale average rainfall. Because of this difference in application, the data used also consider the local scale (a single grid point in the model or radar data), rather than the scale of an urban basin as in Olsson *et al.* (2017) (which is already small, in their case 38 km^2). For the local scale, the use of a percentile-based approach makes it possible to focus on events that correspond to a reasonable worst-case scenario, whereas an approach that takes the maximum as in Olsson *et al.* would be sensitive to rainfall accumulations in single grid points which may not be realistic.

We explore the novel technique in detail using an intense rainfall event that occurred in Garforth (West Yorskhire, UK) on August 22, 2015. The case study is described in more detail in Birch *et al.* (in preparation), which briefly mentions the methodology that is described in full in the current article. Birch *et al.* also explore hydrological modelling of the corresponding surface-water flooding event.

The forecast and radar datasets, as well as the neighbourhood processing method, are introduced in Section 2. In Section 3, we apply the method to the case study, illustrate how the information can be visualized over a larger area and explore sensitivities to parameters of the post-processing algorithm. A number of other cases that occurred in Yorkshire are also discussed. In Section 4, we discuss how the method could be further developed and applied to test the ensemble forecasts systematically.

2 | METHODS

2.1 | MOGREPS-UK

The Met Office Global and Regional Ensemble Prediction System—UK system (MOGREPS-UK) is described in Hagelin *et al.* (2017). The model has a grid spacing of 2.2 km in the central part of the domain, which covers Ireland and the United Kingdom. The physical parametrizations in MOGREPS-UK are largely based on those of the deterministic UK variable resolution model (Lean *et al.*, 2008), which has a grid spacing of 1.5 km. MOGREPS-UK does not use a subgrid parametrization of either shallow or deep moist convection.

Several technical improvements to MOGREPS-UK have taken place over the past years. Of particular importance is the upgrade to the ENDGame (Even Newer Dynamics for General atmospheric modelling of the environment, Wood *et al.*, 2014) dynamical core which was introduced in February 2015 and is used in the current work. ENDGame solves the Euler equations using a semi-implicit, semi-Lagrangian technique, and ensures better conservation of mass and more accurate dynamics compared to the previous dynamical core.

This work uses the MOGREPS-UK forecasts that were produced at the time of each case study. An overview of recent changes to MOGREPS-UK is given by Hagelin *et al.* (2017). At the time of the Garforth event, the ensemble took its initial and boundary conditions from the Met Office Global Ensemble Prediction system (Bowler *et al.*, 2008). For the more recent case studies, the forecasts were initialized with convective-scale data assimilation (Tennant, 2015) and use stochastic physics (Lock and Boutle, 2016; McCabe *et al.*, 2016). Convective-scale data assimilation reduces the time needed for convection to develop during the spin-up phase, whereas the stochastic physics helps both to spin up convection and to create differences in cell development between members (Leoncini *et al.*, 2013).

For all the case studies, a 12-member ensemble was generated every 6 hr. In other applications, often two of these ensembles are used to create a 24-member time-lagged ensemble (e.g., Golding *et al.*, 2014), but for simplicity this has not been done here. MOGREPS-UK ensemble forecasts are already used as input to the UK Flood Forecasting Centre's grid-to-grid model (Price

et al., 2012), which is used to provide regional flood risk guidance information (Pilling, 2016).

2.2 | Radar observations

The radar observations used in this study are described by Harrison *et al.* (2000; 2009). They are part of the Met Office Nimrod/Radarnet system, which is also used for very short range forecasting (Golding, 1998). The 5 min rainfall product has a resolution of 1 km and is generated directly from radar data in polar coordinates. The data were retrieved from the National Centre for Atmospheric Science British Atmospheric Data Centre archives (Met Office, 2003).

Rainfall rates are derived from a power-law relationship to radar reflectivity. The data are corrected and quality-controlled (e.g., the mean noise is removed, the data are filtered, and radar beam attenuation and topography are corrected for) and regularly calibrated against rain gauges. Nevertheless, radar data are known to have errors which can be up to a factor 2 even after calibration (Joss *et al.*, 1990). However, when it comes to evaluation of extreme events, radar data are often needed to provide the necessary spatial resolution to capture these events (Mittermaier, 2008), as gauge data do not provide a dense enough spatial sampling.

For Nimrod/Radarnet, some comparisons between time-series of accumulation with gauge measurements are given by Harrison *et al.* (2009; 2012). Although the timing of rainfall events is consistent between rainfall and gauge measurements, there tend to be substantial quantitative errors in the radar precipitation. Harrison *et al.* (2012) found errors of about 25% in 24 hr accumulations for locations where gauge measurements were over 30 mm. The radar mostly underestimated precipitation compared to the gauges (although this may be because gauges with a high accumulation were selected for the comparison). Further calibration of the radar data for the specific purpose

of generating extreme rainfall scenarios would help to improve the quantitative validation of the results shown here but is beyond the scope of the current study. The fact that good calibration is hard for the very highest rainfall intensities is one of the reasons we focus on rainfall accumulation that corresponds to the 95th and 99th intensity percentiles within a given neighbourhood, rather than the maximum rainfall intensity, in the current study.

2.3 | Case studies

The case studies were chosen to correspond to observed surface-water flooding events. The main case study concerns August 22, 2015. The Met Office daily weather summaries mention that 62.5 mm of rainfall was measured in Bramham, 10 km to the north of Garforth, over the course of the day (Met Office, 2015). For comparison, the 1981–2010 mean monthly rainfall for August at the Church Fenton observing site 12 km to the east of Garforth was 57.9 mm (Met Office, 2020). Garforth was affected by several storm cells that passed over it in the northward direction from 1500 UTC, as can be seen in the radar images in Figure 1. Each of these cells resulted in significant rainfall rates. The most intense rainfall events occurred around 1515 and 1730 UTC (see also Figure S1, which contains half-hourly radar images).

The next case study concerns August 23, 2017. A cold front with north–south orientation passed Wyke Beck in Leeds (West Yorkshire) around 0830 UTC and moved towards the northeast, reaching Scarborough (on the North Sea coast) around 1100 UTC. Flash flooding occurred in both of these locations. The Met Office daily weather summaries mention that 40.6 mm of rainfall was recorded over the course of the day by the gauge station at Bramham, which is about 10 km from Leeds.

August 13, 2018, was selected as a third case study. On this day, a convergence line with a northwest/southeast orientation, which stretched across much of the
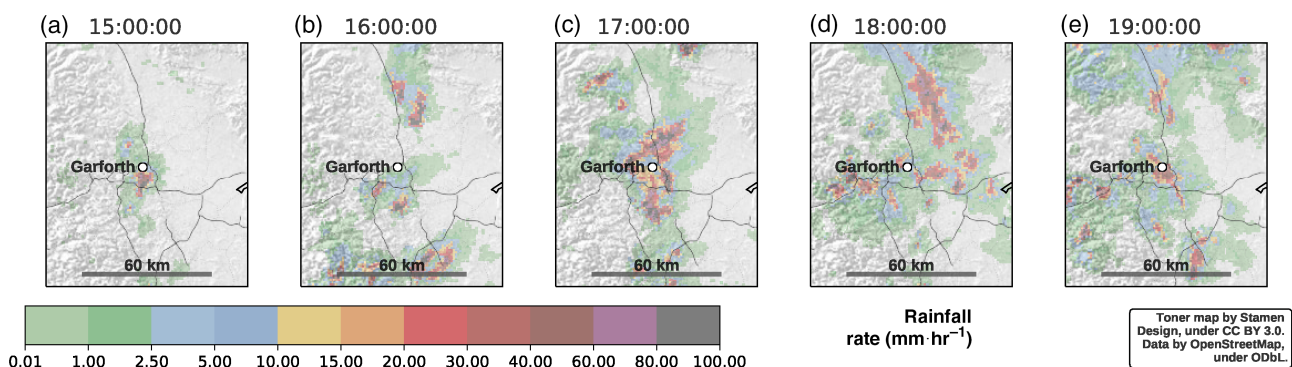

FIGURE 1 Radar rainfall rates for the Garforth event on August 22, 2015

north of the UK, remained stationary in the area surrounding York between 1300 and 1800 UTC. Individual cells within this system moved towards the southeast. Rainfall amounts of around 40 mm over 15 min were recorded at a number of locations in the city of York (City of York Council, 2019).

Radar images for the August 23, 2017, and August 13, 2018, case studies are given in the supplementary material (Figures S2 and S3).

2.4 | Maximum rainfall accumulation maps

Flooding is sensitive to both the duration and intensity of rainfall, as well as to antecedent ground conditions. In constructing rainfall scenarios, we will focus on events that correspond to a relatively large rainfall accumulation over a short interval.

We first determine an optimal interval of interest of duration T, for example 60 min. For each grid box, we find the maximum rainfall accumulation (a_{max}) over T using a rolling window (starting every 5 min) over a search period of interest. In this study, the first 36 hr after the initialization time of each forecast is used as the period of interest for all plots. This means that, when we compare forecasts with different initialization (and hence lead) times in the remainder of this article, the period of interest differs. However, we have made sure that the period of interest always contains the case study event.

For some applications, it may be useful to split each forecast into shorter search periods of, for example, 12 hr so that there is a smaller risk of the method picking up events that occur further than 12 hr apart in the same analysis (preferably with overlap between the periods, so that events are always fully contained within at least one analysis period). t_{max} is the time of maximum accumulation, that is, the time of the start of the hour in which a_{max} occurred. A rolling window starting every 5 min was chosen because this is the temporal resolution of both the radar and MOGREPS-UK data.

A map of $T = 60$ min a_{max} values computed from the radar observations, corresponding to the August 22, 2015 event, is shown in Figure 2a. The reader should bear in mind that the map corresponds to rainfall accumulations with different values of t_{max} at different grid box locations; such a map can be called asynchronous. An emphasis on short-duration events is retained with this strategy. Some of the results shown by Golding *et al.* (2016) are also presented as asynchronous maps (e.g., the probability of hourly rainfall accumulation above a certain threshold at any time of day). The figure shows the rainfall footprint of a storm system which passed over Garforth and its surrounding area. The footprint of the storm accumulation is oriented along the direction of storm movement.

In Figure 2b–d we show the $T = 60$ min a_{max} from three of the MOGREPS-UK ensemble members. The first of these generally agrees well with the radar, even if the location of the storm is not exactly correct. The second
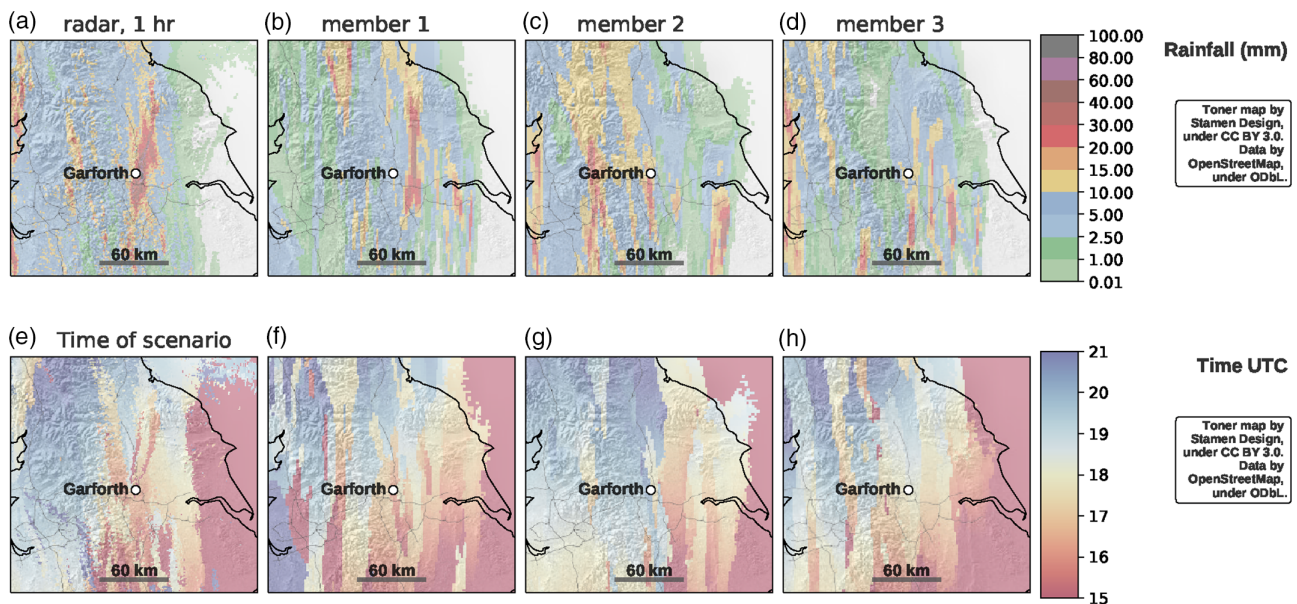

FIGURE 2 Maximum $T = 60$ min accumulations a_{max} from (a) radar and (b)–(d) three ensemble members. Time of maximum accumulation, t_{max}, for (e) radar and (f)–(h) three ensemble members. Computed using a 36 hr search period starting August 22, 2015 at 0300 UTC

6 of 16

Meteorological Applications
Science and Technology for Weather and Climate
Open Access
RMetS

BÖING ET AL.

and third members, on the other hand, fail to predict intense rainfall in the vicinity of Garforth. The maps in Figure 2a–d indicate the total area affected by short-duration events throughout the period of interest. The associated time of maximum accumulation, t_{max}, is shown in Figure 2e–h. This shows that t_{max} tends to be a patchy field, which displays the character of individual storm cells.

2.5 | Neighbourhood processing method

The following steps are used to construct rainfall scenarios using the neighbourhood method.

- a_{max} is computed and mapped as described in Section 2.4 for the area of interest. We define the neighbourhood of a grid box as a circle with a radius of, for example, 30 km (r30), centred at that particular grid box (although any reasonable radius can be used). For each grid box within the area of interest, the values of a_{max} in its neighbourhood are ranked from lowest to highest and percentiles are computed (95th percentile, p95, is used here; see Figure 3). The rainfall accumulation (e.g., r30, p95, T60; the sensitivity to the relevant parameters is explored in Section 3) values can then be plotted on a map.
- For a given grid box of interest (e.g., over Garforth), the rainfall scenario can be extracted from the corresponding target grid box, considering the full 36 hr search period, and used as input for a small-scale hydrological model. This input would typically be provided as a hyetograph, although in this article we will mostly show accumulation curves. We refer to these

scenarios using the same notation as above, for example, r30, p95, T60.

This approach is followed for the radar, individual ensemble members and the ensemble as a whole (by ranking all the values of a_{max} corresponding to the neighbourhood in the full ensemble), which makes it possible to summarize the information in a way that shows how well forecast members agree and which areas the forecasts consider to be at risk of heavy rainfall events.

Rainfall probability distributions within a given area have been used by, for example, Rezacova *et al.* (2007) in the context of verification of a convection-permitting model. Another approach that has been successful in precipitation forecasting is the use of a so-called probability-matched mean (Ebert, 2001). In this approach, the structure of the ensemble mean precipitation field over the domain is retained, but the corresponding rainfall values are adjusted in such a way that the probability distribution function of rainfall matches that of the full ensemble (either over the full domain or locally, see Clark, 2017). However, this approach is less suitable for identifying the potential for smaller and less predictable features in the rainfall field (Surcel *et al.*, 2014).

Olsson *et al.* (2017) go even further by merging different durations (T) of interest into a single metric. We do not employ a single metric for different durations here, although this would be easy to implement as an extension. The advantage of using a specified time scale is that the value of the accumulation has an easy-to-interpret physical meaning, and that it is possible to consider different time scales that are relevant to different types of event. Generally, the same regions are identified when different short time scales are used (e.g., 15 min or an hour).

For a small search radius, the ranking procedure is computationally affordable, as only a single field of accumulation (a_{max}) is used in the neighbourhood analysis around each grid box, rather than information from many different time steps. Nevertheless, the computational effort associated with calculating the probability distributions increases rapidly with the resolution of the simulations and the search radius (it is also much more expensive for the full ensemble than for the individual members). (The expected computational cost of a single sorting operation involving n grid points scales as n log (n) for the quick-sort algorithm used here. This implies that if we consider a domain of constant size and a constant search radius, the total cost of these operations scales as $\rho_t^2 \rho_s^2$ log (ρ_s^2). Here, ρ is the density of grid points (the inverse grid spacing) in a single direction, and

a_{max}: accumulation over 60 minute period (within longer interval) as shading

direction of storm movement

neighbourhood: circle with 30km radius

target grid box r30,p95,T60

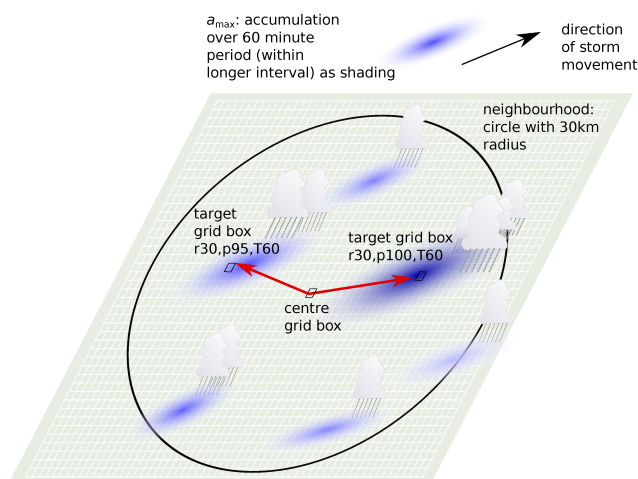target grid box r30,p100,T60

centre grid box

FIGURE 3 Schematic illustration of the post-processing procedure used

the subscripts t and s refer to the target (output) and sampling (input) points, respectively. This means that for the results in this work, the computational cost increases by a factor of about 20 when the resolution of both the target points and sampled points is doubled.)

3 | RESULTS

The accumulation curves corresponding to raw forecast/radar data and the (r30, p95, T60) and (r30, p99, T60) scenarios at Garforth for the August 22, 2015, case study are shown in Figure 4. They are based on the forecast

FIGURE 4　Radar and ensemble accumulation curves using the August 22, 2015, 0300 UTC forecast for Garforth with (a) unprocessed data and (b), (c) processed data corresponding to (b) the (r30, p95, T60) and (c) the (r30, p99, T60) scenarios. In all panels, the thick black line corresponds to the unprocessed radar data

ensemble that was initialized on August 22 at 0300 UTC and are compared to the unprocessed ensemble data, which are shown in Figure 4a. In the unprocessed ensemble data, none of the members captures the high intensity event observed by the radar (thick black line in Figure 4a), and the rainfall tends to occur later in the day. When we consider the post-processed ensemble data, at least some of the members produce similar amounts of rainfall to observations, although even for the (r30, p99, T60) scenarios all the forecasts still have lower rainfall. There is a wide spread in rainfall rates between the post-processed ensemble members, which indicates that some members fail to produce heavy rainfall in a 30 km radius. The timing of the rainfall is also closer to that in observations for the post-processed data.

From the post-processed radar data in Figure 4b,c, it appears that Garforth was one of the worst affected locations in a 30 km radius, as the unprocessed radar accumulation (black line) corresponds to about the 95th percentile in the post-processed radar data (grey line in Figure 4b).

Figure 5 shows the map of post-processed ensemble forecast (r30, p95, T60) accumulations, using the same forecast initialization time as in Figure 4. The ensemble shows that the highest risk of heavy rainfall is located in the immediate vicinity of Garforth (Figure 5n), which is consistent with the post-processed radar data. Individual members, however, vary widely in the approximate location where they predict events and their forecasted intensity. About half of the forecast members predict a single event to occur with its centre near Garforth, whereas some of the other ensemble members predict multiple areas of heavy precipitation.

The maps may look similar to guidance maps, but care should be taken in interpreting them. Besides not giving information about impacts, the maps do not account for model bias, ensemble underspread and the limitations of the method. For example, previous work has shown that the location where convection is triggered is influenced by topography and boundary-layer forcings in the UK (Bennett et al., 2006). We are currently not making any adjustments to take into account preferential triggering of convection at certain locations, which means that the representativeness of the scenarios is likely to degrade near coastlines and near topography. However, this issue affects both the radar data and the forecasts in a similar way, and the exact location of rainfall maxima downwind of a mountain range or near coastlines will still have some inherent uncertainty. In principle, the search neighbourhood could incorporate additional criteria: for example, points that have a large vertical separation from the target location or are over sea could be excluded.

We also remark that, unlike the maps, the extracted accumulation curves are not coherent on scales larger than the grid scale (i.e., between neighbouring grid boxes, this spatial coherency is not preserved in the ranking procedure). It would be possible to address this issue by performing the ranking on spatially averaged data at a
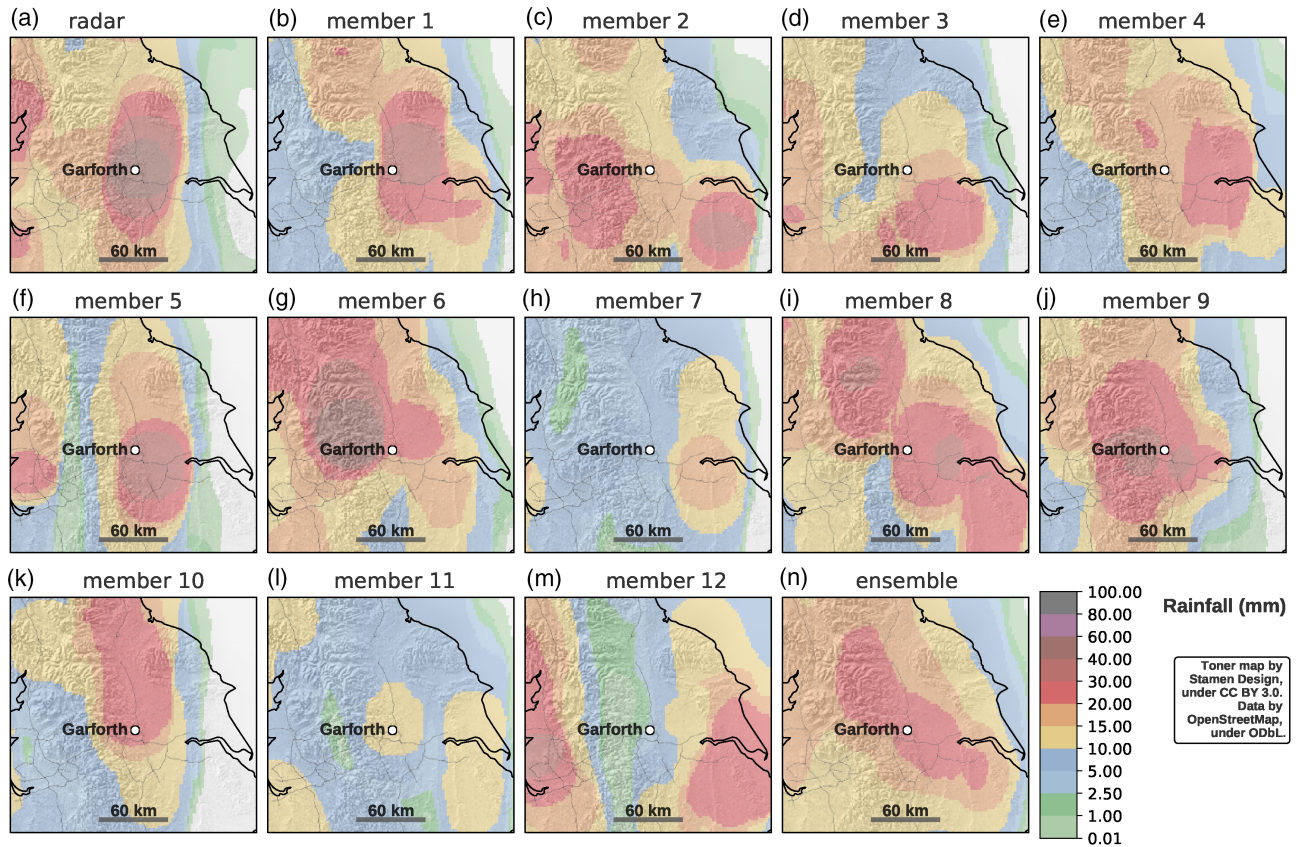

FIGURE 5   Map of (r30, p95, T60) ensemble forecast accumulations, using the August 22, 2015, 0300 UTC forecast
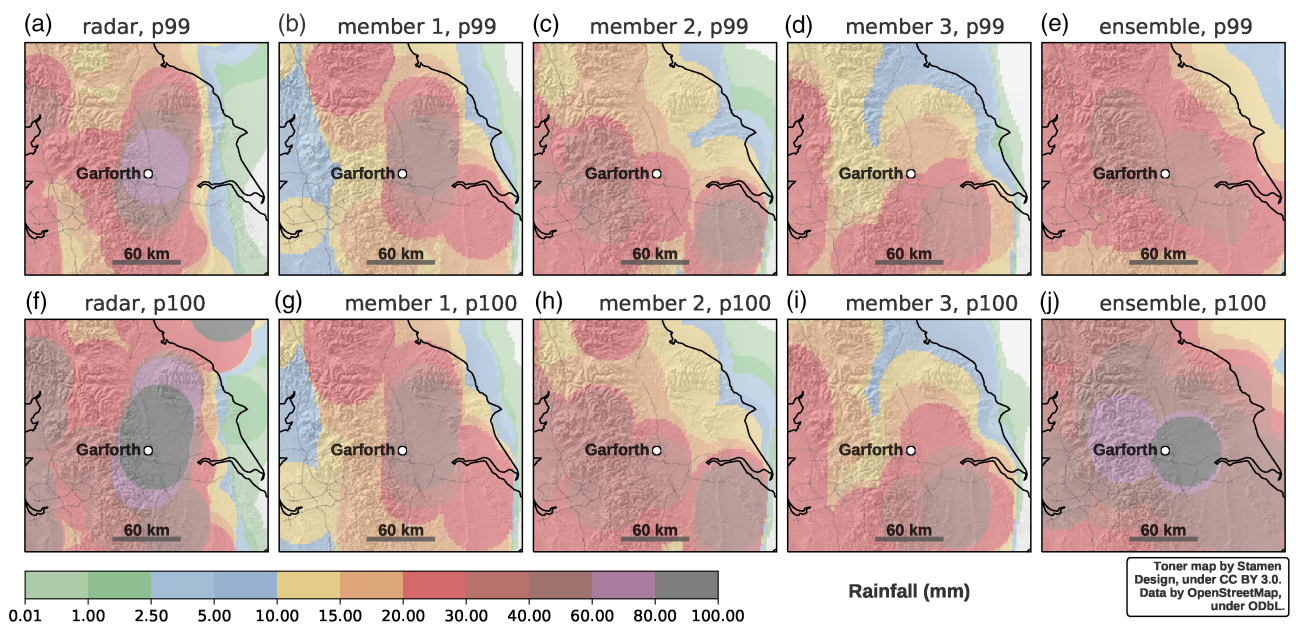

FIGURE 6   Maps of (a)–(e) the post-processed (r30, p99, T60) rainfall accumulation from the radar, three ensemble members and the full ensemble; (f)–(j) the same using the (r30, p100, T60) scenario. Based on the August 22, 2015, 0300 UTC forecast

scale of interest. In the current work, the scale of interest is assumed to be similar to the grid scale (or at least below the scale at which accumulation curves vary substantially).
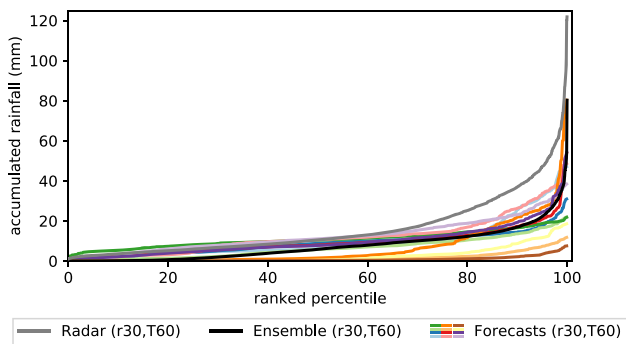

FIGURE 7 Probability distribution of maximum ($r30$, $T60$) accumulation for a circular area around Garforth. The grey line represents a probability distribution derived from radar, the black line the probability distribution over the entire ensemble, and the coloured lines the probability distribution corresponding to individual ensemble members

When a higher percentile is used (Figure 6), the maps show the imprint of very local high accumulations clearly: they appear as circles on the maps. In the ($r30$, $p100$, $T60$) maps (Figure 6f–j), the single grid box with the highest accumulation will become the target grid box for all its neighbours in a 30 km radius. For the individual members in the ($r30$, $p99$, $T60$) maps (Figure 6g–i) the same behaviour is also found. The corresponding ensemble map in Figure 6e shows a smoother field. There are large differences between the $p95$, $p99$ and $p100$ accumulations (it should be kept in mind that the colour scale is very nonlinear). Both the radar data and the forecasts may contain considerable errors for the more extreme accumulations, as the radar may not be as well calibrated for such events and the microphysics and numerical mixing in the forecast will also play a role in determining the behaviour of resolved storm cells.

When the full probability distributions in the 30 km search radius around Garforth are considered (Figure 7), for each of the 12 forecast members (coloured lines) and the radar (thick black line), it can be seen that all of the ensemble members have lower rainfall within the search
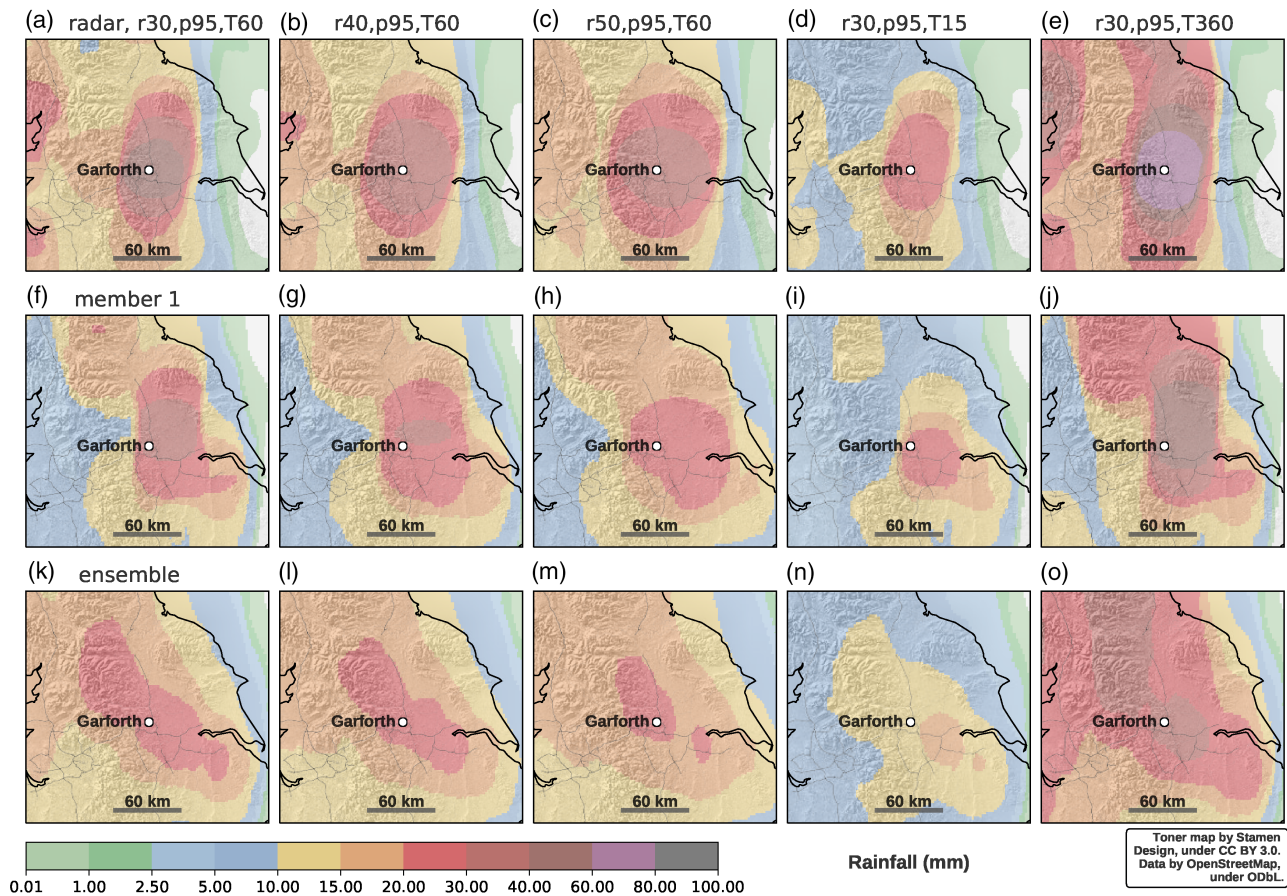

FIGURE 8 For radar (a)–(e), one individual member (f), (j) and the ensemble (k)–(o). Columns correspond to the ($r30$, $p95$, $T60$), ($r40$, $p95$, $T60$), ($r50$, $p95$, $T60$), ($r30$, $p95$, $T15$) and ($r30$, $p95$, $T360$) scenarios (i.e., the search radius varies in the second and third columns and the duration of accumulation in the fourth and fifth columns)

radius, particularly for the higher percentiles. Although some ensemble members produce heavy rainfall near Garforth, none of them produces heavy rainfall over as large a part of the search neighbourhood as observed. This could be due to a systematic issue with the initial conditions of the ensemble or the model physics, or it might be that there would be value to using a larger

ensemble or search radius. The heavy rainfall events in some of the forecasts and the radar data correspond to a rapid increase in accumulations at the higher percentiles. This again illustrates that the choice of percentile is crucial for the accumulations obtained.

Increasing the search radius to 40 km or 50 km results in a smoother field, particularly for radar data and
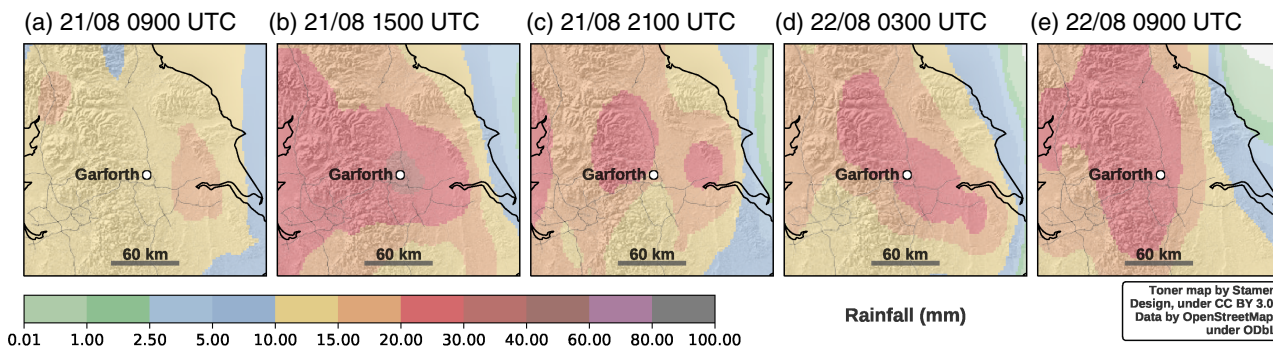

FIGURE 9 Maps of ensemble accumulations based on ($r30$, $p95$, $T60$) scenarios for different forecast initialization times. (d) The initialization time used in Figures 4–8
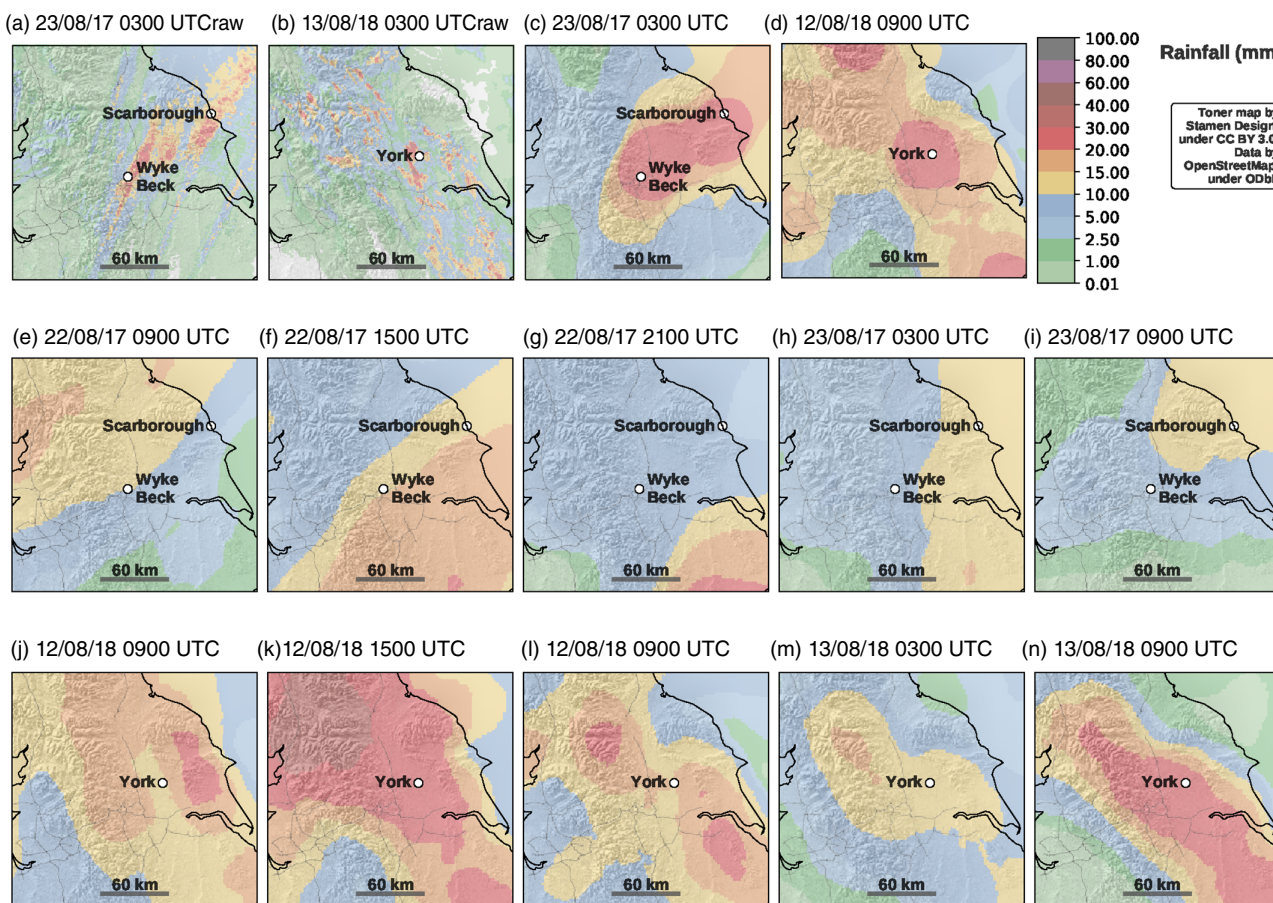

FIGURE 10 (a), (b) Raw radar $T = 60$ min a_{max} for the additional case studies, (c), (d) radar ($r30$, $p95$, $T60$) accumulation maps, (e)–(i) ($r30$, $p95$, $T60$) ensemble accumulation maps for the August 23, 2017 Scarborough case study with forecast initialization times shown above the panels, (j–n) the same, for the August 13, 2018 York case study

individual ensemble members (Figure 8a–c, 8f–h, 8k–m). Additional work would be needed to determine an optimal radius which reflects the uncertainty in the ensemble. This will be case-dependent as well, as it relates to the size and spacing of storms as well as terrain heterogeneity, orography and land–sea contrast. The time interval can also be varied to correspond to a representative storm duration and/or flooding response time, which will depend on the catchment/drainage characteristics of the location of interest. Figure 8d–f, 8i–j and 8n–o show the results for $T = 15$ min and $T = 6$ hr. The 15 min interval largely detects the same events (although with smaller accumulations), whereas the 6 hr interval of accumulation accentuates the steady rainfall over the Pennine hills, in the west of the domain, in the forecasts.

Maps of the post-processed ensemble for different forecast initialization times are shown in Figure 9, where the initialization time is shown at the top of the figure panels. The post-processing uses the first 36 hr of forecast data for each initialization time. These maps show that the possible location of heavy rainfall events is quite consistently predicted in forecasts with different initialization times up to 2 days ahead.

Further analysis found that the August 21, 1500 UTC, ensemble contains events with a higher intensity than the ensemble analysed above (which was initialized on August 22 at 0300 UTC and is shown in Figure 9d). The use of forecasts initialized at multiple lead times, as is the current operational practice, is useful to obtain a wider range of possible scenarios.

Figure 10 displays summary maps for the other two case studies. The radar data in Figure 10a,b show that the events of interest correspond to high observed rainfall accumulations. For the August 23, 2017 event, the accumulations show two maxima near Scarborough and Wyke Beck. For the August 13, 2018 event the precipitation occurs over smaller areas embedded in a band from the northwest to the southeast. Both the two maxima and the band are visible in the ($r30$, $p95$, $T60$) post-processed data in Figure 10c,d (using a start time of August 12, 2018, 0900 UTC in Figure 10c as some of the radar data for August 14 are unavailable).

For the August 23, 2017 Wyke Beck/Scarborough event the ensembles (Figure 10e–i, corresponding to different forecast initialization times) did not predict the heavy rainfall that occurred, although a risk of high rainfall was predicted for Scarborough. The storm system in which these events occurred was not well represented in the forecasts, the reasons for which will require further analysis. For the August 13, 2018 York case study (Figure 10j–n), on the other hand, the ensemble forecasts accurately predicted the convergence line, and the prediction of a chance of high rainfall is particularly accurate in

the forecast ensemble with the shortest lead time. The ensemble accumulation curves from the 0300 UTC forecasts on the day of each event (Figure 11) show that, for both events, the accumulation itself is uncertain, and for the August 13, 2018 York case study the timing of the maximum accumulation is earlier in the ensemble members than in the radar observations.

One of the challenges in providing pluvial flood guidance is that a system can easily indicate the risk of events that fail to materialize (false alarms). In order to obtain a first impression of whether the post-processed forecasts are different on days without the same intense rainfall events, we consider the maps of the ($r30$, $p95$, $T60$) accumulation for all forecasts initialized at 0300 UTC on all days in August 2018 in Figure 12. This includes the days of the York case study: it appears that the rainfall amounts in the post-processed forecasts around August 13 (the York case study, cf. Figure 10j–n) stand out compared to any other days during that month. Accumulation maps for June–August 2018, and the corresponding radar-derived maps, are provided in the supplementary material (Figures S4–S9). These figures show, at least qualitatively, that the approach is generally capable of identifying those dates and areas where significant rainfall occurs.
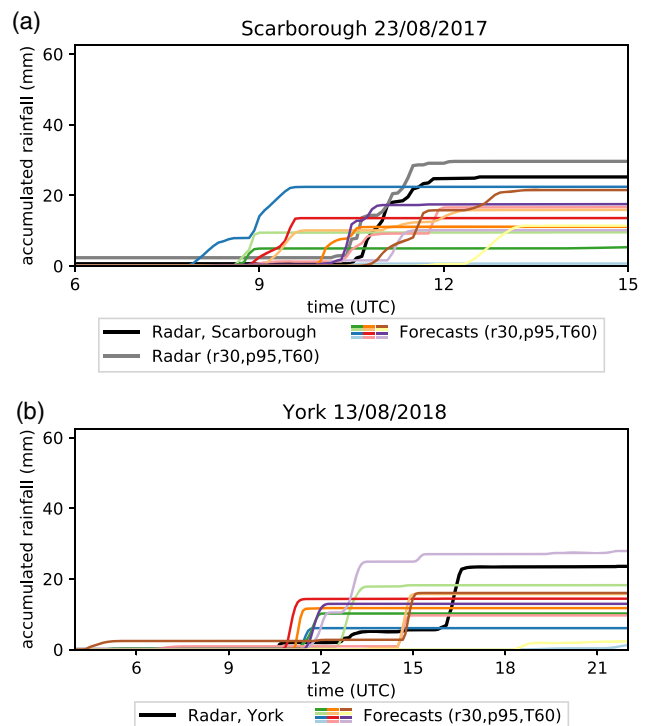

FIGURE 11 Ensemble and radar accumulation curves for the other case studies, using the 0300 UTC forecasts on the day of the event. The thick black line corresponds to unprocessed radar data, the grey line for the Scarborough case study to post-processed radar data (missing for the York case)
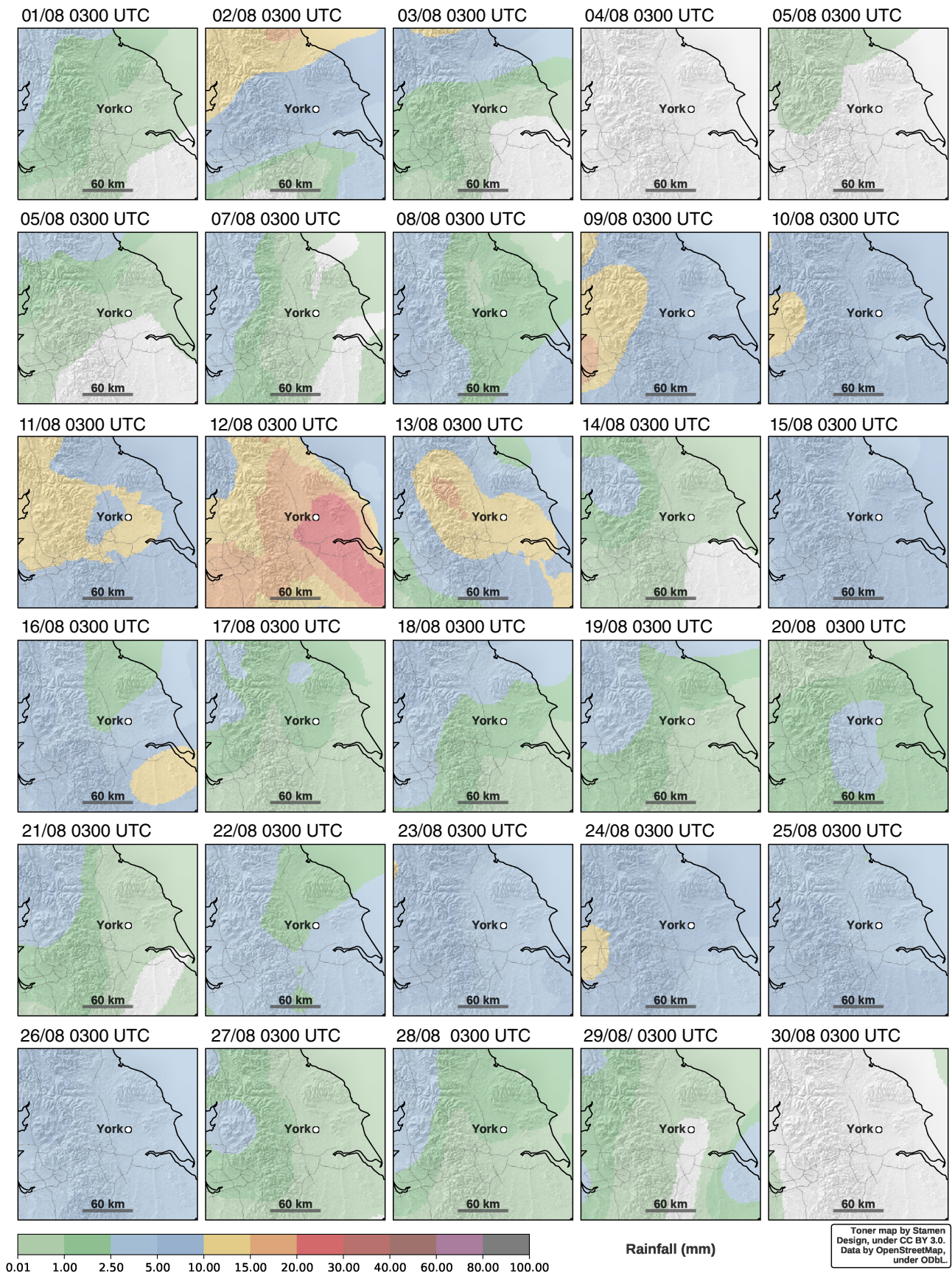
FIGURE 12 Maps of the (r30, p95, T60) post-processed rainfall accumulations for 0300 UTC ensemble forecasts throughout August 2018

4 | CONCLUSIONS AND DISCUSSION

Short-lived but intense rainfall events play an important role in surface-water flooding. These events remain difficult to forecast, despite the availability of ensemble forecasts and neighbourhood processing methods. The combination of these methods can be very effective, but it is a challenge to develop techniques that have a clear-cut interpretation and can be used for hydrological forecasting.

We have demonstrated a method for generating scenarios that are representative for heavy rainfall events from an ensemble forecast and can be used to evaluate these forecasts against radar observations. The approach also makes it possible to produce maps that show variability within a larger region.

The results are encouraging: the forecast maps effectively point to areas of increased probability of rainfall in two of the three case studies considered. For the third case study, the possibility of an event in the Scarborough region is detected in the forecasts with a shorter lead time, but the possibility of an event at Wyke Beck is less clear.

Like the distance-dependent depth–duration analysis approach taken by Olsson et al. (2017), the method provides an ensemble reasonable worst-case approach. However, our approach considers events that correspond to a high percentile within a search radius, rather than the very worst-case scenario within the search radius, in order to prevent emphasis on results that may not be physical and reduce sensitivity to single grid box accumulations. By performing an efficient neighbourhood search, the method distinguishes between the differences in the mesoscale meteorological conditions that develop between ensemble members and the fact that the location of triggering for individual storm cells is largely unpredictable. Although using a high percentile in the analysis will unavoidably lead to the occurrence of false alarms on a local scale, we did not find the possibility of heavy rainfall events being detected on a large number of days.

In future work, the maps could be modified to display event return times or expected impacts by combining the rainfall accumulation curves with other data. The scenarios have already been used to produce flood forecasts and both the rainfall scenarios and flood forecasts have been tested in a user workshop (see Birch et al., in preparation, for more details).

A more systematic evaluation of the technique on several years of data would be needed in order to quantify how useful it is over a longer time interval. Such an evaluation could further improve the method by considering an optimal search radius, and it could include further bias correction of the forecast data and radar data against quality-controlled ground-based observations (see, for example, Blenkinsop et al., 2017; Lewis et al., 2018). Calibration is particularly important for the more rare events, where the rain rates provided by either forecast or radar may have systematic biases. In future, a newly developed integrated gauge–radar–satellite merged dataset (UKGrsHP, Yu et al., 2020) could also provide an improved observational reference.

The choice of length scale can probably be optimized on a case-by-case basis, depending on ensemble spread and, for example, the size and spacing of storms. Methods for determining the scales over which ensemble members agree are discussed by Dey et al. (2016a; 2016b). Automating this would require some effort: one idea would be to use machine learning techniques (see, for example, Gagne et al., 2014).

The current serial algorithm can present an issue for operational application, especially for the larger search radii (where running the serial code takes several hours). This issue could be addressed by a combination of parallelizing the code, reducing the number of target and sampling points, optimizing the sorting algorithm to exploit overlap in search neighbourhoods between adjacent grid points, and restricting the sort algorithm to accumulations over a given threshold.

For improving the maps, methods from extreme value theory could be applied to address the fact that the ensemble forecasts represent only a sub-sampling of possible scenarios. Fitting a distribution to the results (e.g., Alfieri et al., 2012) may be beneficial, even though the spatial data in the forecasts cannot be treated as independent data. Together, such improvements could help to develop the current prototype analysis into a useful tool to assist flood guidance providers and hydrological modelling.

14 of 16

Meteorological Applications
Science and Technology for Weather and Climate
Open Access
RMetS

BÖING ET AL.

ORCID

Steven J. Böing https://orcid.org/0000-0003-3794-2563

REFERENCES

Alfieri, L., Thielen, J. and Pappenberger, F. (2012) Ensemble hydro-meteorological simulation for flash flood early detection in southern Switzerland. *Journal of Hydrology*, 424, 143–153.

Archer, D. and Fowler, H. (2018) Characterising flash flood response to intense rainfall and impacts using historical information and gauged data in Britain. *Journal of Flood Risk Management*, 11, S121–S133.

Bennett, L.J., Browning, K.A., Blyth, A.M., Parker, D.J. and Clark, P.A. (2006) A review of the initiation of precipitating convection in the United Kingdom. *Quarterly Journal of the Royal Meteorological Society*, 132, 1001–1020.

Berner, J., Ha, S.-Y., Hacker, J., Fournier, A. and Snyder, C. (2011) Model uncertainty in a mesoscale ensemble prediction system: stochastic versus multiphysics representations. *Monthly Weather Review*, 139, 1972–1995.

Birch, C., Rabb, B., Böing, S., Shelton, K., Lamb, R., Hunter, N., Trigg, M., Hines, A., Taylor, A. and Dale, M. (in preparation) *Enhanced surface water flood forecasts for England and Wales: user-led development and testing.* The final report of the corresponding project. Available at: https://icasp.org.uk/enhanced-surface-water-flood-forecasting-final-report/.

Blenkinsop, S., Lewis, E., Chan, S.C. and Fowler, H.J. (2017) Quality-control of an hourly rainfall dataset and climatology of extremes for the UK. *International Journal of Climatology*, 37, 722–740.

Bouttier, F., Vié, B., Nuissier, O. and Raynaud, L. (2012) Impact of stochastic physics in a convection-permitting ensemble. *Monthly Weather Review*, 140, 3706–3721.

Bowler, N.E., Arribas, A., Mylne, K.R., Robertson, K.B. and Beare, S.E. (2008) The MOGREPS short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 134, 703–722.

Buizza, R. (2008) The value of probabilistic prediction. *Atmospheric Science Letters*, 9, 36–42.

Chan, S.C., Kahana, R., Kendon, E.J. and Fowler, H.J. (2018) Projected changes in extreme precipitation over Scotland and Northern England using a high-resolution regional climate model. *Climate Dynamics*, 51, 3559–3577.

City of York Council (2019) *York 13th August 2018 section 19 surface water flood investigation.* Available at: https://democracy.york.gov.uk/ieDecisionDetails.aspx?AIId=51613. Accessed November 21, 2019.

Clark, A.J. (2017) Generation of ensemble mean precipitation forecasts from convection-allowing ensembles. *Weather and Forecasting*, 32, 1569–1583.

Clark, P., Roberts, N., Lean, H., Ballard, S.P. and Charlton-Perez, C. (2016) Convection-permitting models: a step-change in rainfall forecasting. *Meteorological Applications*, 23, 165–181.

Cloke, H. and Pappenberger, F. (2009) Ensemble flood forecasting: a review. *Journal of Hydrology*, 375, 613–626.

Collier, C. (2007) Flash flood forecasting: what are the limits of predictability? *Quarterly Journal of the Royal Meteorological Society*, 133, 3–23.

Committee on Climate Change (2012) *Climate change – is the UK preparing for flooding and water scarcity? (adaptation subcommittee progress report 2012).* Available at: https://www.theccc.org.uk/publication/climate-change-is-the-uk-preparing-for-flooding-and-water-scarcity-3rd-progress-report-2012/. Accessed April 22, 2020.

Dance, S.L., Ballard, S.P., Bannister, R.N., Clark, P., Cloke, H.L., Darlington, T., Flack, D.L., Gray, S.L., Hawkness-Smith, L., Husnoo, N., et al. (2019) Improvements in forecasting intense rainfall: results from the FRANC (forecasting rainfall exploiting new data assimilation techniques and novel observations of convection) project. *Atmosphere*, 10, 125.

Demeritt, D., Cloke, H., Pappenberger, F., Thielen, J., Bartholmes, J. and Ramos, M.-H. (2007) Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting. *Environmental Hazards*, 7, 115–127.

Department for Environment, Food and Rural Affairs (2018) *Surface water management—an action plan.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/725664/surface-water-management-action-plan-july-2018.pdf. Accessed February 18, 2020.

Dey, S.R., Plant, R.S., Roberts, N.M. and Migliorini, S. (2016a) Assessing spatial precipitation uncertainties in a convective-scale ensemble. *Quarterly Journal of the Royal Meteorological Society*, 142, 2935–2948.

Dey, S.R., Roberts, N.M., Plant, R.S. and Migliorini, S. (2016b) A new method for the characterization and verification of local spatial predictability for convective-scale ensembles. *Quarterly Journal of the Royal Meteorological Society*, 142, 1982–1996.

Ebert, E.E. (2001) Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Monthly Weather Review*, 129, 2461–2480.

Ebert, E.E. (2008) Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. *Meteorological Applications*, 15, 51–64.

Ebert, E.E. (2009) Neighborhood verification: a strategy for rewarding close forecasts. *Weather and Forecasting*, 24, 1498–1510.

Gagne, D.J., McGovern, A. and Xue, M. (2014) Machine learning enhancement of storm-scale ensemble probabilistic quantitative precipitation forecasts. *Weather and Forecasting*, 29, 1024–1043.

Gilleland, E., Ahijevych, D., Brown, B.G., Casati, B. and Ebert, E.E. (2009) Intercomparison of spatial forecast verification methods. *Weather and Forecasting*, 24, 1416–1430.

Golding, B. (1998) Nimrod: a system for generating automated very short range forecasts. *Meteorological Applications*, 5, 1–16.

Golding, B., Ballard, S., Mylne, K., Roberts, N., Saulter, A., Wilson, C., Agnew, P., Davis, L., Trice, J., Jones, C., et al. (2014) Forecasting capabilities for the London 2012 Olympics. *Bulletin of the American Meteorological Society*, 95, 883–896.

Golding, B., Roberts, N., Leoncini, G., Mylne, K. and Swinbank, R. (2016) MOGREPS-UK convection-permitting ensemble products for surface water flood forecasting: rationale and first results. *Journal of Hydrology*, 17, 1383–1406.

Hagelin, S., Son, J., Swinbank, R., McCabe, A., Roberts, N. and Tennant, W. (2017) The Met Office convective-scale ensemble, MOGREPS-UK. *Quarterly Journal of the Royal Meteorological Society*, 143, 2846–2861.

Hapuarachchi, H., Wang, Q. and Pagano, T. (2011) A review of advances in flash flood forecasting. *Hydrological Processes*, 25, 2771–2784.

Hardy, J., Gourley, J.J., Kirstetter, P.-E., Hong, Y., Kong, F. and Flamig, Z.L. (2016) A method for probabilistic flash flood forecasting. *Journal of Hydrology*, 541, 480–494.

Harrison, D., Driscoll, S. and Kitchen, M. (2000) Improving precipitation estimates from weather radar using quality control and correction techniques. *Meteorological Applications*, 7, 135–144.

Harrison, D. L., Norman, K., Pierce, C. and Gaussiat, N. (2012) Radar products for hydrological applications in the UK. In: *Proceedings of the Institution of Civil Engineers—Water Management*, vol. 165, pp. 89–103.

Harrison, D. L., Scovell, R. W. and Kitchen, M. (2009) High-resolution precipitation estimates for hydrological uses. In: *Proceedings of the Institution of Civil Engineers—Water Management*, vol. 162, pp. 125–135.

Hohenegger, C., Brockhaus, P. and Schär, C. (2008) Towards climate simulations at cloud-resolving scales. *Meteorologische Zeitschrift*, 17, 383–394.

Hohenegger, C. and Schär, C. (2007) Atmospheric predictability at synoptic versus cloud-resolving scales. *Bulletin of the American Meteorological Society*, 88, 1783–1794.

Joss, J., Waldvogel, A. and Collier, C. (1990) Precipitation measurement and hydrology. In: *Radar in Meteorology*. Boston, MA: American Meteorological Society, pp. 577–606.

Kendon, E.J., Roberts, N.M., Fowler, H.J., Roberts, M.J., Chan, S.C. and Senior, C.A. (2014) Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change*, 4, 570.

Lean, H.W., Clark, P.A., Dixon, M., Roberts, N.M., Fitch, A., Forbes, R. and Halliwell, C. (2008) Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Monthly Weather Review*, 136, 3408–3424.

Leoncini, G., Plant, R., Gray, S. and Clark, P.A. (2013) Ensemble forecasts of a flood-producing storm: comparison of the influence of model-state perturbations and parameter modifications. *Quarterly Journal of the Royal Meteorological Society*, 139, 198–211.

Lewis, E., Quinn, N., Blenkinsop, S., Fowler, H.J., Freer, J., Tanguy, M., Hitt, O., Coxon, G., Bates, P. and Woods, R. (2018) A rule based quality control method for hourly rainfall data and a 1 km resolution gridded hourly rainfall dataset for Great Britain: CEH-GEAR1hr. *Journal of Hydrology*, 564, 930–943.

Lock, A.P. and Boutle, I.A. (2016) *Stochastic Physics Code in the UM*. Exeter, UK: Met Office, p. 81. Unified Model Documentation Paper. Tech. rep.

Mass, C.F., Ovens, D., Westrick, K. and Colle, B.A. (2002) Does increasing horizontal resolution produce more skillful forecasts? The results of two years of real-time numerical weather prediction over the Pacific Northwest. *Bulletin of the American Meteorological Society*, 83, 407–430.

McCabe, A., Swinbank, R., Tennant, W. and Lock, A. (2016) Representing model uncertainty in the Met Office convection-permitting ensemble prediction system and its impact on fog forecasting. *Quarterly Journal of the Royal Meteorological Society*, 142, 2897–2910.

Met Office (2003) Met Office rain radar data from the NIMROD system (through the NCAS British Atmospheric Data Centre). Available at: http://catalogue.ceda.ac.uk/uuid/82adec1f896af6169112d09cc1174499. Accessed November 25, 2019.

Met Office (2015) Met Office daily weather summary August 2015. Available at: https://digital.nmla.metoffice.gov.uk/IO_34ed588a-7798-4ed0-97d0-19c4f42140ab/. Accessed November 20, 2019.

Met Office (2019) UK climate projections: headline findings. Available at: https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/research/ukcp/ukcp-headline-findings-v2.pdf. Accessed February 10, 2020.

Met Office (2020) Church Fenton UK climate averages. Available at: https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-climate-averages/gcx4kb837. Accessed February 20, 2020.

Mittermaier, M. (2008) Introducing uncertainty of radar-rainfall estimates to the verification of mesoscale model precipitation forecasts. *Natural Hazards and Earth System Sciences*, 8, 445–460.

Murphy, J. and Palmer, T. (1986) Experimental monthly long-range forecasts for the United Kingdom. 2. A real-time long-range forecast by an ensemble of numerical integrations. *Meteorological Magazine*, 115, 337–348.

Ochoa-Rodríguez, S., Wang, L.-P., Thraves, L., Johnston, A. and Onof, C. (2018) Surface water flood warnings in England: overview, assessment and recommendations based on survey responses and workshops. *Journal of Flood Risk Management*, 11, S211–S221.

Olsson, J., Pers, B.C., Bengtsson, L., Pechlivanidis, I., Berg, P. and Körnich, H. (2017) Distance-dependent depth–duration analysis in high-resolution hydro-meteorological ensemble forecasting: a case study in Malmö City, Sweden. *Environmental Modelling and Software*, 93, 381–397.

Palmer, T.N. (2001) A nonlinear dynamical perspective on model error: a proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quarterly Journal of the Royal Meteorological Society*, 127, 279–304.

Pilling, C. (2016) New developments at the Flood Forecasting Centre: operations and flood risk guidance. *WIT Trans. Built Environ*, 165, 237–248.

Prein, A.F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., Lipzig, N.P.M. and Leung, R. (2015) A review on regional convection-permitting climate modeling: demonstrations, prospects, and challenges. *Reviews of Geophysics*, 53, 323–361.

Price, D., Hudson, K., Boyce, G., Schellekens, J., Moore, R., Clark, P., Harrison, T., Connolly, E. and Pilling, C. (2012) Operational use of a grid-based model for flood forecasting. In: *Proceedings of the Institution of Civil Engineers—Water Management*, vol. 165, p. 65.

Rezacova, D., Sokol, Z. and Pesice, P. (2007) A radar-based verification of precipitation forecast for local convective storms. *Atmospheric Research*, 83, 211–224.

Roberts, N.M. and Lean, H.W. (2008) Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review*, 136, 78–97.

Schwartz, C.S., Kain, J.S., Weiss, S.J., Xue, M., Bright, D.R., Kong, F., Thomas, K.W., Levit, J.J., Coniglio, M.C. and

Wandishin, M.S. (2010) Toward improved convection-allowing ensembles: model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Weather and Forecasting*, 25, 263–280.

Schwartz, C.S. and Sobash, R.A. (2017) Generating probabilistic forecasts from convection-allowing ensembles using neighborhood approaches: a review and recommendations. *Monthly Weather Review*, 145, 3397–3418.

Surcel, M., Zawadzki, I. and Yau, M. (2014) On the filtering properties of ensemble averaging for storm-scale precipitation forecasts. *Monthly Weather Review*, 142, 1093–1105.

Tennant, W. (2015) Improving initial condition perturbations for MOGREPS-UK. *Quarterly Journal of the Royal Meteorological Society*, 141, 2324–2336.

Theis, S., Hense, A. and Damrath, U. (2005) Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach. *Meteorological Applications*, 12, 257–268.

Vincendon, B., Ducrocq, V., Nuissier, O. and Vié, B. (2011) Perturbation of convection-permitting NWP forecasts for flash-flood ensemble forecasting. *Natural Hazards and Earth System Sciences*, 11, 1529–1544.

Wood, N., Staniforth, A., White, A., Allen, T., Diamantakis, M., Gross, M., Melvin, T., Smith, C., Vosper, S., Zerroukat, M. and Thuburn, J. (2014) An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quarterly Journal of the Royal Meteorological Society*, 140, 1505–1520.

Yu, J., Li, X.-F., Lewis, E., Blenkinsop, S. and Fowler, H.J. (2020) UKGrsHP: a UK high-resolution gauge–radar–satellite merged hourly precipitation analysis dataset. *Climate Dynamics*, 54, 2919–2940.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Böing SJ, Birch CE, Rabb BL, Shelton KL. A percentile-based approach to rainfall scenario construction for surface-water flood forecasts. *Meteorol Appl*. 2020;27:e1963. https://doi.org/10.1002/met.1963