An Efficient Ensemble Technique for Hydrologic Forecasting Driven by Quantitative Precipitation Forecasts

HUMBERTO VERGARA,[a,b] JONATHAN J. GOURLEY,[b] AND MICHAEL ERICKSON[c,d]

[a] *Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma, Norman, Oklahoma*
[b] *NOAA/National Severe Storms Laboratory, Norman, Oklahoma*
[c] *Cooperative Institute for Research in Environmental Science, University of Colorado Boulder, Boulder, Colorado*
[d] *NOAA/Weather Prediction Center, College Park, Maryland*

(Manuscript received 1 July 2022, in final form 23 November 2022)

ABSTRACT: Uncertainty in quantitative precipitation forecasts (QPFs) from numerical weather prediction (NWP) models manifests in errors in the amounts of rainfall, storm structure, storm location, and timing, among other precipitation characteristics. In flash flood forecasting applications, errors in the QPFs can translate into significant uncertainty in forecasts of surface water flows and their impacts. In particular, the QPF errors in location and structure result in errors on flow paths, which can be highly detrimental in identifying locations susceptible to flash flood impacts. To account for this type of uncertainty, the neighboring pixel ensemble technique (NPET) was devised and implemented as a postprocessing algorithm of deterministic or ensemble outputs from a distributed hydrologic model. The aim of the technique is to address displaced hydrologic responses resulting from location biases in QPFs using a probabilistic approach. NPET identifies a sampling region surrounding each forecast pixel and builds an ensemble of surface water flow values considering the pixel's physiographic similarities. The probabilistic information produced with NPET can be calibrated through a set of tunable parameters that are adjusted to account for NWP-specific QPF error characteristics. The utility of NPET is demonstrated for the Ellicott City flash flood event on 27 May 2018, using products and tools routinely used in the U.S. National Weather Service for warning operations. Results from this case demonstrate that NPET effectively conveys uncertainty information about QPF precipitation location in a hydrologic context.

SIGNIFICANCE STATEMENT: This study introduces a new method suitable for operational use called the neighboring pixel ensemble technique (NPET). NPET is an algorithm that generates ensemble-based streamflow forecasts accounting for the location uncertainties in quantitative precipitation forecasts (QPFs) *without the requirement of multiple hydrologic model runs*. NPET is capable of this feat through probabilistic assimilation of a priori QPF displacement information and its uncertainty. The application of NPET with the Flooded Locations and Simulated Hydrographs (FLASH) project shows the technique could be beneficial for flash flood warning operations in the U.S. National Weather Service (NWS). It is envisioned that the application of NPET with Warn-on-Forecast System (WoFS)-forced FLASH outputs will further enhance the quality of flash flood forecasts that support NWS warning operations.

KEYWORDS: Probability forecasts/models/distribution; Ensembles; Hydrologic models; Postprocessing

1. Introduction

Rainfall, as the driving force of natural hazards such as floods, flash floods, debris flow and shallow landslides, is arguably the single most important variable in dynamical models (Baum and Godt 2010; Rasheed et al. 2022; Slater and Villarini 2017) used in early warning systems. Other variables describing the state of the hydrologic system and its physiographic properties are also important in buffering or increasing the effects of precipitation on processes leading to the threat (Bedient et al. 2008; Hall et al. 2014). The capabilities of remote sensing technologies enable the estimation of spatial rainfall fields over continental and even global scales. The radar-based quantitative precipitation estimates (QPEs) are used as forcing to hydrologic forecasting systems that focus on short-fused hydrologic hazards such as rainfall-driven flash floods and debris flows (Gourley et al. 2017). The biggest limitation with QPE-driven forecast systems is the inability to supply rainfall at future time steps, which limits the lead time in forecasting the concomitant hydrologic impacts. Accurately forecasting hydrologic responses at longer lead times would enable forecasters to issue more timely flash flood warnings.

To extend lead time beyond the typical response times of the land surface to observed precipitation, it is necessary to utilize precipitation amounts at future time steps using extrapolation methods and/or quantitative precipitation forecasts (QPFs), the latter of which are nominally provided by numerical weather prediction (NWP) models (Cuo et al. 2011; Ghimire et al. 2021). Forecast rainfall fields have uncertainties due to a number of factors including inaccuracies in depicting the initial states and lateral boundary conditions of the atmosphere, microphysical and boundary layer processes that are often parameterized in the NWP models, and the initiation of convection (Duda and Gallus 2013; Jankov et al. 2007a,b). These problems are particularly pervasive during the warm season when the most intense rainfall rates typically occur over small spatial scales and the operational NWP models exhibit the lowest forecast skill (Fritsch and Carbone 2004).

Corresponding author: Humberto Vergara, humber@ou.edu

DOI: 10.1175/JHM-D-22-0109.1

Wernli et al. (2008) developed a taxonomy for verification metrics attributed to errors in precipitation structure, amplitude, and location (SAL). In general, these errors are considered independently from one another, so efforts are spent to spatially and temporally match QPF objects to analysis-based radar rainfall estimates. Structure refers to the storm organizational mode or storm type as being multicell, squall line, supercell, air mass thunderstorm, etc. An amplitude error indicates that, after the forecast and observed objects are matched, the QPF amounts are biased. Location errors arise from spatial displacements in the forecast rainfall from the observed rainfall. In the SAL framework, contemporary convection allowing models (CAMs) yield QPFs that have minimal structural and amplitude errors but can still have significant locational errors. For example, Duda and Gallus (2013) noted a 100-km spatial displacement on average for 6-h accumulated precipitation at a threshold of 2.54 mm using a Weather Research and Forecasting (WRF) Model run out 24 h. Likewise, Yan and Gallus (2016) noted an average displacement of near 100 km using 3-h accumulated precipitation out to forecast hour 12 using the North American Mesoscale Forecast System (NAM), Global Forecast System (GFS), and a WRF Model also at a threshold of 2.54 mm.

Several approaches have been adopted to address some of the shortcomings of the deterministic model forecasts. One approach is to generate an ensemble of possible outcomes by perturbing boundary conditions, initial conditions, and/or explicit physics schemes or parameterized processes (Stensrud et al. 2000). Ensemble hydrologic forecasts can then be obtained treating each QPF member as an equally likely outcome in addition to perturbing the hydrologic model states, parameters, and/or model structures. For example, the ensemble streamflow prediction (ESP) framework has been used widely in operational streamflow forecasting in the NWS (Day 1985; McEnery et al. 2005; Schaake and Larson 1998; Seo et al. 2006; Smith et al. 1992). These approaches often require multiple hydrologic model runs and are useful in calibrating the ensemble mean and spread for seasonal streamflow prediction.

Direct use of deterministic QPFs in a hydrologic forecasting context can lead to outcomes that can appear precise in location, timing, and magnitude, but can miss the correct flow paths and thus mislead a forecaster. Carlberg et al. (2020) addressed this shortcoming by systematically shifting precipitation objects 55.5 and 111 km in range separated by 45° in azimuth. A larger ensemble size of 81 members was generated and then input to a hydrologic model. Results showed that the method was able to increase the probability of detecting floods. Further improvements were attained by assigning weights to the shifted precipitation members based on climatological shifts in the QPF fields. While providing a basis for considering short-term QPF spatial displacements in streamflow prediction, operational application to flash flood scale remains computationally challenging due to the need for multiple hydrologic model runs in real time.

Hardy et al. (2016) recognized the challenges in applying ensemble QPFs in an operational hydrologic modeling context at flash flood scale without the requirement for multiple hydrologic runs over a large domain. They used a storm scale NWP model to generate ensemble QPFs and then derived the localized probability matched mean (PMM; Ebert 2001) as done by Clark (2017), which addresses shortcomings with the simple arithmetic ensemble mean such as a smoothing of individual features (i.e., increases frequency of lower values while reducing frequency of larger values). The PMM was then used as forcing to a distributed hydrologic model over a single, representative basin. The outcome of this experiment was the distribution of streamflow values plotted as a function of basin area. This information was then mapped to a larger, representative domain to identify the basin scales that were susceptible to flash flooding given the forcing from the localized PMM. While the method required several manual steps and thus limited its operational application, it nevertheless provided preliminary direction on novel uses of ensemble QPFs as forcing to hydrologic models that account for spatial uncertainties in the forecast precipitation fields.

This study introduces a new method suitable for operational use called the neighboring pixel ensemble technique (NPET). NPET is a postprocessing algorithm that generates ensemble-based streamflow forecasts accounting for the location uncertainties in QPF forcing fields *without the requirement of multiple hydrologic model runs* in parallel. NPET is capable of this feat through probabilistic assimilation of a priori QPF displacement information and its uncertainty. Section 2 presents a generalized formulation of the technique, followed by its demonstration in a case study in section 3, including the experiments, materials and methods. Results from the experiments and a discussion describing the capabilities and considerations of NPET are presented in section 4. Finally, section 5 summarizes this work and discusses future studies, and a vision for the application of NPET in operational settings.

2. Formulation of the technique: Building ensembles from deterministic forecasts

NPET postprocesses QPF-forced hydrologic forecasts computed on a rectangular grid of pixels to generate hydrologic ensembles. Ensembles are built for any grid point on the hydrologic forecast grid, sampling values from neighboring, physiographically similar pixels to account for QPF object's locational uncertainty. The basic concept of ensemble generation with NPET is illustrated using the synthetic example for a single location shown in Fig. 1. This particular location (i.e., pixel) for which the ensemble is created will hereafter be referred to as the *NPET node* to simplify the description of the algorithm. A conceptual rainfall object as measured by an observational platform (e.g., radar, satellite, or rain gauge network) is presented in Fig. 1a, highlighting the corresponding hydrologic response on the flow path ending at the NPET node. In Fig. 1b, a rainfall forecast identical in amplitude and structure is created with a northeast displacement bias thus leading to a hydrologic response on the wrong flow paths. In a deterministic framework, the displaced QPF object could potentially highlight areas that will not be truly impacted by the storm (i.e., producing false alarms), while the area of interest

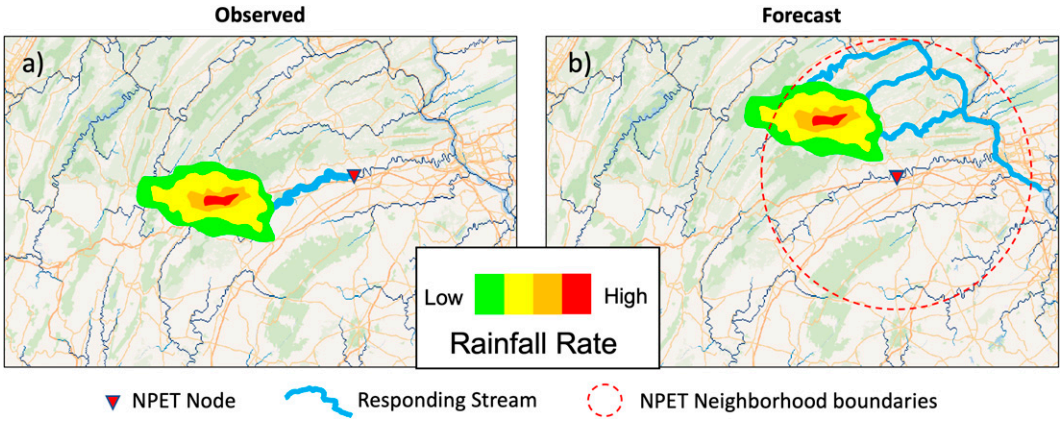

FIG. 1. Synthetic example of displaced rainfall fields for illustration of NPET application: (a) true location of rainfall and hydrologic response and (b) displaced rainfall and hydrologic response.

in this example (i.e., NPET node) would experience impacts but would not be identified (i.e., producing misses). NPET can be used in this scenario to recover the displaced hydrologic response by sampling values that were forecasted at locations surrounding the NPET node to build an ensemble and provide information about the threat likelihood.

In Fig. 1b, the red dashed-line circle delineates the NPET sampling area, which will be referred to hereafter as the *NPET neighborhood*. An important capability of NPET is the ability to characterize the unknown true location of displaced hydrologic responses based on a priori QPF displacement estimates and their uncertainty obtained through characterization of a QPF product's long- or short-term spatial biases. For example, if no information about the direction and magnitude of displacement is available, NPET can be used for sampling large areas around the NPET node, which can result in ensembles skewed toward low or zero values. The latter is common in cases of localized convection such as with single-cell thunderstorms or supercells. However, if information about the displacement is available, NPET can also characterize various degrees of uncertainty by targeting specific regions around the NPET node and producing ensembles with less skewness toward low values.

A weighting function is used to define the NPET neighborhood employing a bivariate distribution to represent the level of uncertainty about the true location of the future precipitation and corresponding hydrologic response with respect to the NPET node. The weighting field is based on the estimation of the joint probability of the distances from the NPET node along the south–north (y or vertical axis) cardinal axis and the distances along the west–east (x or horizontal axis) cardinal axis, for which univariate density distributions are first defined independently. The beta distribution family was selected to model univariate distributions because of its flexibility to model a variety of curve shapes (Jeffreys and Jeffreys 1999; Podlubny 1998) through various combinations of its shape parameters. For NPET, if the distribution is symmetric, it means that the assignment of weights will not be a function of the direction with respect to the NPET node along the corresponding cardinal axis. In those cases, the sampling is said to be isotropic, which is applicable when no information about the direction of the displacement is available to target specific regions within the NPET neighborhood. The opposite occurs when the distribution is skewed, in which case the sampling is said to be anisotropic. The NPET neighborhood is specified with three tunable parameters:

- Its radius r, typically in kilometers (km), which defines the size of the sampling area;
- A Euclidean vector $\mathbf{E} = (E_x, E_y)$, $E_x, E_y \in (-r, r)$ denoting the expected distance and direction of displacement, which controls the skewness of the distribution;
- A unitless scalar parameter $\omega \in [0, \infty)$ that adjusts the spread of the weight's distribution around the expected displacement distance. The range of values of this parameter is not related to any physical quantity, so it must be defined empirically.

Once the univariate distributions have been defined, each coordinate pair of univariate probabilities contained in a circle of radius r are multiplied to arrive to the joint probability distribution following:

$$\mathbf{P}^{xy} = [P(x_i \cap y_j)] = \begin{cases} (p_i^x p_j^y), \forall \left(i - \frac{n}{2}\right)^2 + \left(j - \frac{n}{2}\right)^2 \le \left(\frac{r}{dx}\right)^2 \\ (0)_{i,j}, \forall \left(i - \frac{n}{2}\right)^2 + \left(j - \frac{n}{2}\right)^2 > \left(\frac{r}{dx}\right)^2 \end{cases}, \in \mathbb{R}^{n \times n}, \tag{1}$$

AU2

where $\mathbf{P}^{xy}$ is the 2D joint distribution of n rows and n columns, p_i^x and p_j^y are independently computed probabilities with univariate distribution functions, and dx is the pixel resolution (km). The weighting field delineating the NPET neighborhood $\mathbf{W}_{\text{NPET}}$ is obtained normalizing the elements of the $\mathbf{P}^{xy}$ matrix:

$$\mathbf{W}_{\text{NPET}} = \frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{n} p_i^x p_j^y} \mathbf{P}^{xy} = \left[\frac{p_i^x p_j^y}{\sum_{i=1}^{n}\sum_{j=1}^{n} p_i^x p_j^y} \right] \in \mathbb{R}^{n \times n}, \tag{2}$$

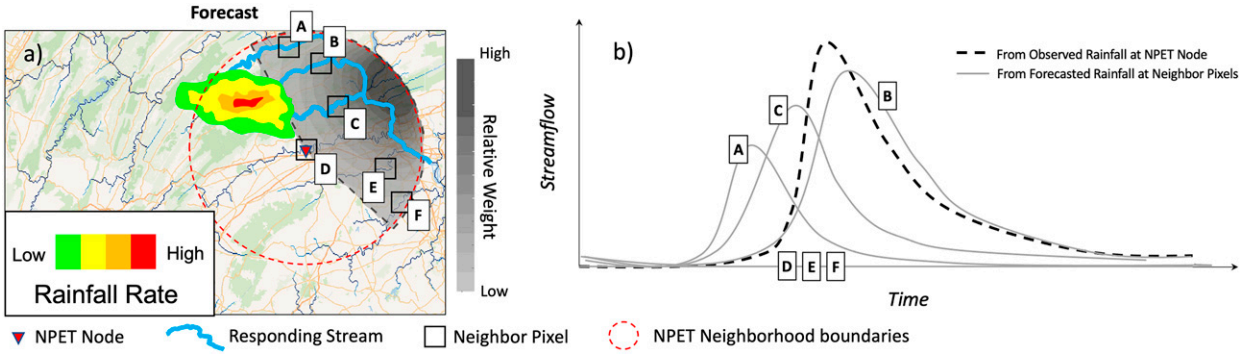

FIG. 2. Implementation of NPET to example case in Fig. 1: (a) NPET neighborhood modeled as weighting field and (b) NPET ensemble time series from sampling locations labeled A–F shown in (a).

where $\sum \mathbf{W}_{\mathrm{NPET}} = 1.0$. Figure 2a shows the NPET weight field computed for the example case in Fig. 1 in grayscale, where the sampling is concentrated over the northeast region of the NPET neighborhood using a radius $r = 150$ km, a Euclidean vector $\mathbf{E} = (120, 90\ \mathrm{km})$, and $\omega = 10$.

NPET builds ensembles borrowing the hydrologic response from nearby pixels (referred to hereafter as *neighbor pixels*) with physiographic attributes and/or antecedent conditions from their contributing drainages that can significantly differ from those of the NPET node's catchment. This can lead to an ensemble variance higher than that intended to represent QPF displacement uncertainty. NPET addresses this through the following: let τ be a *transfer function* of the hydrologic response at a given neighbor pixel to the NPET node such that

$$q_l^{\mathrm{NPET}} = \tau(q_l), \tag{3}$$

where q_l is the deterministic hydrologic response to QPFs at the lth neighbor pixel, and q_l^{NPET} is the hydrologic response transferred to the NPET node for inclusion into the ensemble as the lth member. Notice that Eq. (3) is presented in a generic form such that τ can be any real-valued function that converts q_l into q_l^{NPET} without changing the physical meaning of the hydrologic variable (i.e., q_l and q_l^{NPET} are commensurable quantities). Combining Eqs. (2) and (3), a mathematical expression for the construction of NPET ensembles for any given location can be written as

$$\mathbf{q}^{\mathrm{NPET}}(t) = \begin{pmatrix} W_{\mathrm{NPET}}(1) \times \tau[q_1(t)] \\ W_{\mathrm{NPET}}(2) \times \tau[q_2(t)] \\ \vdots \\ W_{\mathrm{NPET}}(z-1) \times \tau[q_{z-1}(t)] \\ W_{\mathrm{NPET}}(z) \times \tau[q_z(t)] \end{pmatrix}, \tag{4}$$

where $\mathbf{q}^{\mathrm{NPET}}(t)$ is the NPET ensemble at any given time step t, and z is the number of ensemble members that must be less than n^2 (the total number of pixels within the NPET neighborhood). Even though many physiographic attributes affect hydrologic response, the contributing area of a basin is arguably the single most important attribute determining runoff volume and peak flow (Bedient et al. 2008). Consequently, the transfer function used in this study was designed to restrict the selection of neighbor pixels based on how comparable their contributing areas are to that of the NPET node with the following simple piecewise function:

$$\tau(q_l) = \begin{cases} q_l, \text{ if } D_A^{\mathrm{Node}}(1-\mathbb{D}) \le D_A^l \le D_A^{\mathrm{Node}}(1+\mathbb{D}) \\ \varnothing, \text{ if } D_A^l < D_A^{\mathrm{Node}}(1-\mathbb{D}) \text{ or } D_A^l > D_A^{\mathrm{Node}}(1+\mathbb{D}) \end{cases} \tag{5}$$

where D_A^{Node} is the drainage area (km^2) at the NPET node, D_A^l is the drainage area (km^2) at the lth neighbor pixel, $\mathbb{D}$ is a tunable parameter that specifies the limits of variation in contributing area (expressed as a fraction), and $\varnothing$ is the null value symbol. In other words, if the contributing area of the lth neighbor pixel falls within the limits of variation of the contributing area of the NPET node, then the hydrologic response of the lth neighbor pixel is used in the NPET ensemble; otherwise, it is excluded. The six squares with single-letter labels from A to F in Fig. 2a represent neighbor pixels with drainage areas that qualify to contribute values to the ensemble. Notice that the location labeled with the letter D is in fact the NPET node. Also, only six locations have been highlighted in this case for the sake of illustration, while an actual application of NPET more typically results in an ensemble with hundreds of members. Figure 2b presents a time series plot showing the hydrologic response estimated from the observed rainfall and the 6-member NPET ensemble. In this example, the probabilistic information provided by NPET would indicate a nonzero chance of significant hydrologic response at the NPET node, as opposed to the zero-response suggested by the deterministic forecast (member D).

Note that, except for Eq. (5), the formulation of NPET discussed in this section presents a generalized form for the implementation of the technique. Customization is possible via Eqs. (1)–(2) if a different model of bivariate probabilities is available, and/or via Eq. (3) through any real-valued function that generates new values of unit streamflow. Also note that no specific hydrologic model nor QPFs are used in this formulation; this is because NPET is agnostic about modeling physics or QPF bias characteristics.

TABLE 1. Variations of parameters for NPET configurations in the OSE.

Parameter	Value 1	Value 2	Value 3
ω	0.0	17.0	60.0
$\mathbb{D}$	12.5%	25.0%	50.0%
r	150.0 km		
E	(62.9 km west, 77.7 km north)		

3. Study design

a. Experiments

1) SENSITIVITY ANALYSIS

The purpose of this experiment was to test and demonstrate the capabilities of NPET in a controlled setting. The aim was to isolate the effects of spatial displacement from those caused by errors in precipitation amplitude and structure. To accomplish this, an observing system experiment (OSE; Masutani et al. 2010) was designed in which reference QPEs were used to generate displaced rainfall inputs for the simulation of hydrologic response, thus emulating displaced QPFs. A steady-state assumption about precipitation displacement biases was used, in which a unique displacement bias vector is defined for the entire storm. This displacement strategy produces displaced precipitation fields with the same direction and length, thus resulting in a storm track perfectly parallel to the true trajectory with respect to the axis of the weather system's motion.

For the first part, an evaluation of hydrologic forecast errors was conducted to understand the impact of rainfall displacement biases considering different physiographic characteristics. A total of 360 scenarios were generated, consisting of 120 displacement directions in the [0°, 360°) interval and separated by 3° in azimuth, and three displacement lengths: 25, 50, and 100 km. The precipitation values used for this experiment were extracted from the spatial subset labeled as "*hydrometeorological domain*" [see section 3b(1)]. These 360 rainfall scenarios were then used as inputs to generate hydrologic simulations, which were compared against the baseline reference run to compute various error statistics.

In the second part, NPET was implemented for one of the displacement scenarios to assess the impact of various sampling parameter's configurations on the ensemble-based forecast skill to highlight the correct area under threat and encapsulate the reference response. The scenario chosen for this experiment was the one with a displacement of 100 km, 321° azimuth (northwest quadrant). To simplify the analysis, the experiment herein was limited to parameter ω, which controls how concentrated the sampling is over a particular region of the NPET neighborhood, and parameter $\mathbb{D}$, which controls the selection of members and the ensemble's sample size. Parameters r and **E**, which control the size of the NPET neighborhood and expected region of displacement, were specified based on available information about the QPFs projected displacement since this was the main perturbation of the OSE. Table 1 summarizes the parameterizations of NPET used in the experiment.

Figure 3 illustrates the NPET neighborhoods produced with parameters in Table 1. Value 1 of parameter ω results in isotropic sampling, while the other two values result in anisotropic sampling with two different levels of sampling concentration which were assessed herein through the central angle of the NPET neighborhood defining the minor sector that encapsulates at least 95% of the total weights. This 95% threshold value is arbitrary and only intended to describe the level of spatial concentration of the sampling relative to the tuning of parameter ω. Value 2 was tuned to concentrate the sampling within a quadrant sector (i.e., 90° central angle), while value 3 was tuned for an octant sector (i.e., 45° central angle), which results in higher spatial concentration of the sampling. The three values of parameter $\mathbb{D}$ produce three levels of flexibility in terms of the drainage size criterion. The higher the value of $\mathbb{D}$, the more flexibility and the larger the ensemble is. The actual number of members attributed to the NPET node varies in proportion to the spatial distribution of drainage areas.

The ensembles produced with the isotropic implementation of NPET were compared to ensembles generated with a simple, yet computationally more expensive, method based on Carlberg et al. (2020). This method will be referred to hereafter as C2020 for succinctness. First, the selected perturbed

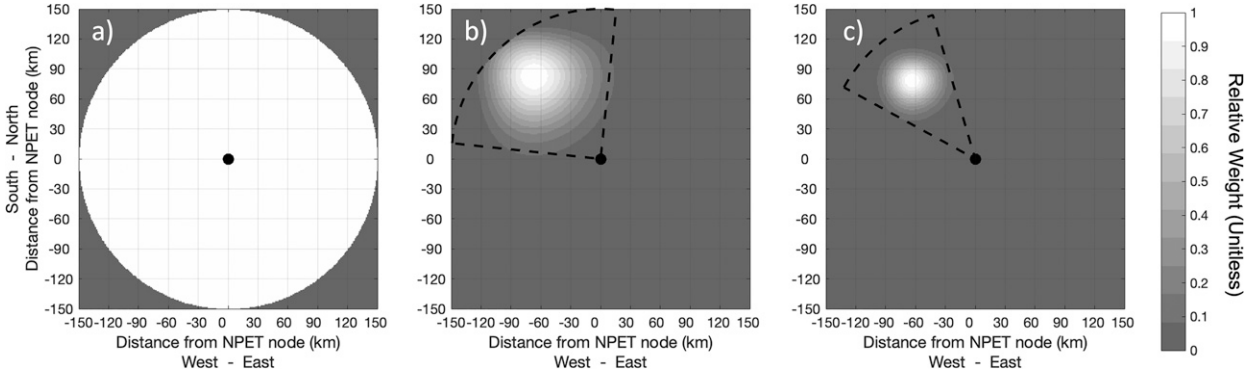

FIG. 3. NPET neighborhoods for the three values of parameter ω in Table 1: (a) value 1—isotropic sampling, (b) value 2—anisotropic with 95% of weights within a quadrant, and (c) value 3—anisotropic with 95% of weights within an octant.

scenario (100 km, 321°) was shifted in the traditional wind direction octants (i.e., N, NE, E, SE, S, SW, W, and NW) with a fixed 100-km displacement length. Then, the hydrologic model was run with each one of the shifted inputs. Note that even though this method was applied to a 9-member ensemble of QPFs in Carlberg et al. (2020) and here it is only applied to a single realization, the mechanics of the method are the same.

2) Real QPFs demo

For this experiment, NPET was tested on hydrologic forecasts driven by an actual QPF product, which is subject not only to displacement biases but also to errors in rainfall intensities and storm structure. As a preliminary step, the forecast precipitation was analyzed with respect to the reference precipitation to determine the overall displacement direction and magnitude. Second, this displacement information was used to configure NPET with an isotropic sampling, which mainly entails defining the size of the NPET neighborhood through the r parameter. Last, a subjective and non-exhaustive calibration was performed on parameters $\mathbb{D}$, $\mathbf{E}$, and ω to produce a tuned anisotropic configuration of NPET for the particular case study. A simple iterative trial and error process based on manual parameter adjustments was followed until the NPET outputs highlighted the true area that observed flash flooding. The aim of this latter exercise was not to generalize NPET parameterizations, but rather to demonstrate the flexibility of the technique and provide insights into possible strategies for optimization.

b. Materials and methods

1) Case study

The experiments in this study were done with observations and simulation data for the flash flood event occurring between 1900 UTC 27 May and 0300 UTC 28 May 2018 over Ellicott City in Maryland. Viterbo et al. (2020) presents a detailed description of the genesis of the storms and subsequent hydrologic response leading to impacts, indicating that flash flood warnings turned into the declaration of a flash flood emergency within a timespan of 30 min. Multiple reports of flash flooding were collected in the NWS Storm Events database (see Fig. 4), while some of the U.S. Geological Survey (USGS) stream gauges measured levels exceeding the minor and moderate flood stages.

For the experiments conducted herein, three domains were delineated and are presented in Fig. 4a. The largest domain (red rectangle) was set for the hydrologic model computational grid, while the domain delineated in blue was used to focus the analysis over a region around the main area where impacts were reported. Four of the most impacted locations identified in Viterbo et al. (2020), including three USGS gauges exceeding flood stages (USGS 01593500—exceeded minor level; USGS 01594000—exceeded moderate level; and USGS 01589352—exceeded minor level) and an ungauged location on the Tiber River where postmortem analysis by USGS estimated peak values exceeding historic flooding levels, were selected to examine time series of unit streamflow.

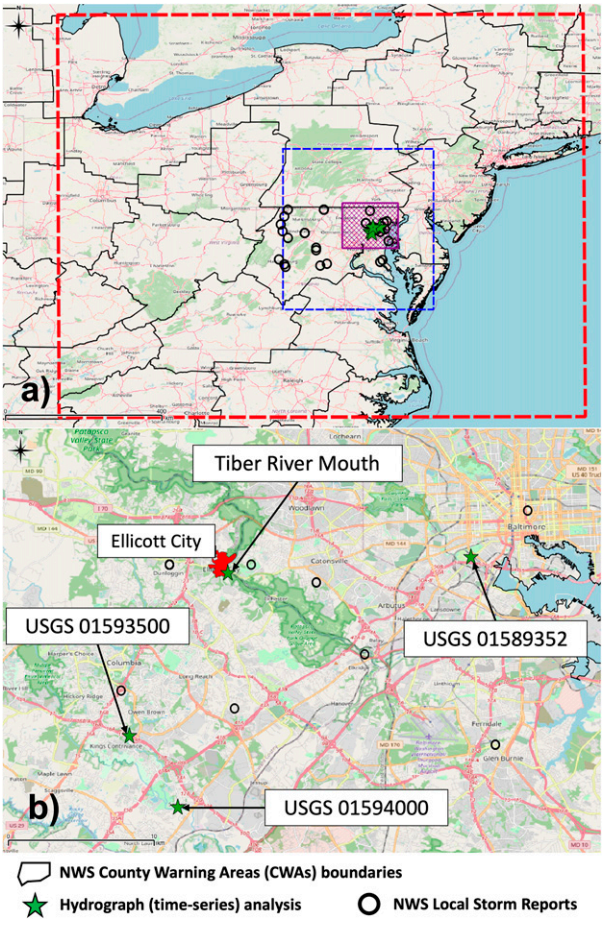

FIG. 4. Spatial domains for the Ellicott City case in May 2018: (a) large-scale view with hydrological model computational domain (red dashed-line rectangle), analysis domain (blue dashed-line rectangle), and hydrometeorological domain (purple textured rectangle) and (b) main area impacted around Ellicott City district boundaries.

These four locations fall in the small basin category on the flash flood scale (Clark et al. 2014) with values of unit streamflow exceeding 1.0 $m^3 s^{-1} km^{-2}$, which has been associated with the occurrence of impacts on lives and property (Martinaitis et al. 2017).

2) Precipitation data

QPEs used to produce reference simulations were based on the radar-only product from the Multi-Radar Multi-Sensor (MRMS; Zhang et al. 2016) system, available every 2 min at a 1-km pixel resolution over the conterminous United States (CONUS). For the sensitivity analysis, a 20-min series of QPEs were used instead of the complete 2-min series to keep computational costs low. Because 20-min QPEs were used in the OSE, the same temporal frequency was used for both the reference runs and displaced scenarios. For the Real QPFs Demo experiment, however, the full 2-min series was employed to create the reference data.

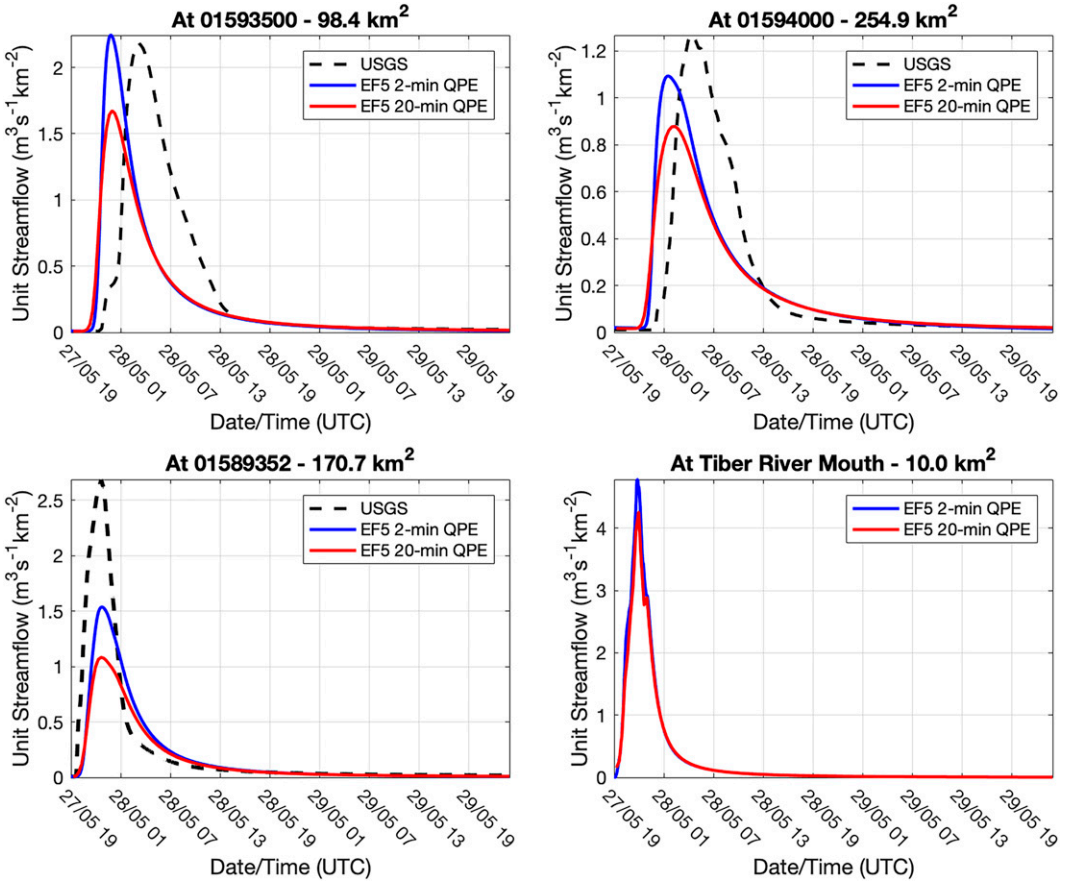

FIG. 5. Time series of simulated unit streamflow with EF5 using 2- and 20-min MRMS QPEs inputs for the Ellicott City event. Measurements are included at the USGS gauged locations.

QPFs from the High Resolution Rapid Refresh (HRRR) model, a rapidly updating CAM in the U.S. National Weather Service (NWS) since 2014 (Benjamin et al. 2016; Dowell et al. 2022; James et al. 2022), were used for the Real QPFs Demo. The version of the HRRR employed herein was version 3 with hourly temporal resolution, an 18-h forecast horizon, and a 3-km pixel resolution. The forecast run valid at 1300 UTC 27 May 2018 was used to demonstrate how NPET can be used on hydrologic forecasts forced by this QPF product.

3) HYDROLOGIC MODELING

The Ensemble Framework For Flash Flood Forecasting (EF5), a distributed hydrologic modeling software in the NWS as part of the Flooded Locations And Simulated Hydrographs (FLASH) project (Flamig et al. 2020; Gourley et al. 2017) was used to conduct the hydrologic simulations. FLASH was transitioned to the NWS in 2016 and yields contemporary products used for the issuance of flash flood warnings and watches. At present, the QPE-forced EF5 yields streamflow forecasts out to 12 h in the future on a 1-km grid every 10 min. EF5 must complete its forecasts in less than 10 min at over 10 million grid points across CONUS and outer U.S. territories. Given the limited computational resources available at the National Centers for Environmental Prediction (NCEP) where FLASH runs operationally, it is not possible to have multiple instances of EF5 forecasts given forcing from ensemble storm scale QPFs.

EF5 models the macroscale processes of the hydrological cycle on a rectangular computational grid with multiple physics options for water balance calculations and flow routing along hillslopes and streams. The physics options used herein are based on the Coupled Routing and Excess Storage (CREST; Wang et al. 2011) coupled with the kinematic wave (KW) approximation to the Saint-Venant Equations of one-dimensional unsteady open channel flow, which has produced the most skillful estimates of unit streamflow among all physics options in FLASH (Gourley et al. 2017). The setup for EF5 used herein was based on its CONUS-wide implementation in FLASH and does not require hydrologic model parameter tuning against any reference streamflow (e.g., from USGS stream gauges). CREST-KW parameters in FLASH were estimated from readily available geospatial information of digital terrain, soil, and land cover/use datasets (Flamig et al. 2020; Vergara et al. 2016), which produces a reasonably skillful model for a priori parameter settings as evidenced in the comparison of simulated unit streamflow against USGS measurements for the Ellicott City event at the three gauged selected locations in Fig. 5. EF5 was

configured to produce simulations with both a 20-min and a 2-min time step consistent with the MRMS QPE settings previously discussed. Additionally, EF5 was configured to produce hourly estimates of unit streamflow from the HRRR QPFs.

4) HYDROLOGIC FORECAST PERFORMANCE ASSESSMENT

The performance and error metrics were chosen to describe attributes of hydrologic forecasts relevant for flash flood warning operations with FLASH: 1) the magnitude of streamflow values; 2) the exceedance of streamflow levels associated with flash floods; and 3) the spatial characteristics of streamflow objects that delineate areas of flash flood threat. For attribute 1, maps of unit streamflow differences were calculated as follows:

$$\mathbf{q}_{\text{diff}} = (q_{i,j}^{\text{sim}} - q_{i,j}^{\text{ref}})_{n\times m}, \tag{6}$$

where $\mathbf{q}_{\text{diff}}$ is a matrix of unit streamflow differences ($\text{m}^3\ \text{s}^{-1}\ \text{km}^{-2}$) with n rows and m columns defined on the domain of interest on the 1-km grid, $q_{i,j}^{\text{sim}}$ and $q_{i,j}^{\text{ref}}$ are, respectively, the simulated and reference unit streamflow at the ith row and jth column. Likewise, the ability of ensembles to encapsulate the reference response was assessed through visual inspection of maps of unit streamflow ensembles and time series at the four selected locations. Percentiles were extracted from the ensembles to describe the distribution of unit streamflow values modeled with NPET. For attribute 2, the number of pixels exceeding threshold values of 1.0, 2.0, and 5.0 $\text{m}^3\ \text{s}^{-1}\ \text{km}^{-2}$ were reported. These threshold values were based on levels of severity determined from prior research (e.g., Martinaitis et al. 2017).

Last, for attribute 3, an object-oriented scheme based on the contiguous rain areas (CRA; Ebert and McBride 2000) method was employed to facilitate the description of spatial hydrologic information. The scheme was applied to highlight areas of flash flood threat from maps of maximum unit streamflow. A flash flood entity was defined following guidance from NWS forecasters about spatial consistency of high-resolution products (Gourley and Vergara 2021). The scheme herein was found to encapsulate the main area of threat and impacts from the Ellicott City flash flood event using a threshold of 0.35 $\text{m}^3\ \text{s}^{-1}\ \text{km}^{-2}$ with a minimum contiguous area of 40 km^2 applied to the reference simulation with MRMS QPEs. These settings were fixed and used for the scenarios with displaced rainfall and resulting hydrologic response.

4. Results

a. Sensitivity analysis

1) IMPACT OF PRECIPITATION DISPLACEMENT ON HYDROLOGIC FORECASTS

Figure 6 presents a sample of eight maps of unit streamflow differences [Eq. (6)] corresponding to the traditional wind direction octants from the 120 scenarios of displaced rainfall inputs corresponding to a 100-km displacement [section 3a(1)]. Pixels in blue colors indicate underestimation of unit streamflow values, while pixels in red colors show overestimation. Because this experiment only used precipitation from the isolated area of the hydrometeorological domain (see Fig. 4a) for both the baseline and displaced runs, underestimation only occurs over the correct area of hydrologic response for all scenarios. Therefore, the analysis herein focuses on the features observed over the areas of overestimation, which correspond to the displaced hydrologic response. The spatial patterns of overestimation differ among the different scenarios, which is further confirmed by the different shapes of the flash flood entities. Recall that all scenarios share the exact same storm structure and rainfall intensities, and that the displacement is produced by a steady-state assumption. Therefore, these differences can only be explained by differences in physiographic characteristics. Clear evidence of this can be seen on the panels for the south, southeast, and east displacements of Fig. 6 where the storm is mainly placed over the Chesapeake Bay estuary. This resulted in either significantly smaller objects or not enough hydrologic response to produce a flash flood entity.

This analysis was extended to include all 120 directions and for the three displacement distances, for which a summary is presented in Fig. 7. The top-left and center plots present information about the magnitude of unit streamflow values in the displaced hydrologic response, while the top-right plot describes the size of flash flood entities. The three bottom plots show the variability of unit streamflow values determining the severity of flash flood impacts. All metrics illustrate the anisotropic nature of the impact of displaced rainfall fields on hydrologic response. In some cases, the differences in hydrologic response among neighboring areas can be large, as shown by the maximum difference metric with values that ranged between 10.0 and 20.0 $\text{m}^3\ \text{s}^{-1}\ \text{km}^{-2}$ (top-left plot), although the mean differences in unit streamflow were only between 0.5 and 1.0 $\text{m}^3\ \text{s}^{-1}\ \text{km}^{-2}$ (top-center plot). Moreover, a direct correlation between any of the metrics and displacement length cannot be generalized. All metrics are consistently higher toward the southwest-south direction for the 25- and 50-km displacement runs because large portions of the Washington, D.C., metropolitan area are within these distances, where high values of unit streamflow result from direct runoff over impervious surfaces. Similar observations can be made for the same distances to the East (slightly northeast) in the general Baltimore area. Areas to the north-northwest are generally more rural, while areas to the southeast are dominated by estuaries. At distances close to 100 km, the terrain is either rural or over the estuaries, which explains why, in general, the 100-km run presents the lowest values in all metrics. This analysis demonstrates that the disparities in physiographic attributes that buffer or exacerbate hydrologic response to rainfall can be significant over relatively small distances and highlights the importance of taking this heterogeneity into account when using QPF-based hydrologic forecasts subject to displacement biases.

2) NPET TESTBED

Figure 8 presents the results of applying the C2020 method to the selected displacement scenario. Only the 90th, 95th, and 99th percentiles are included because these ensembles

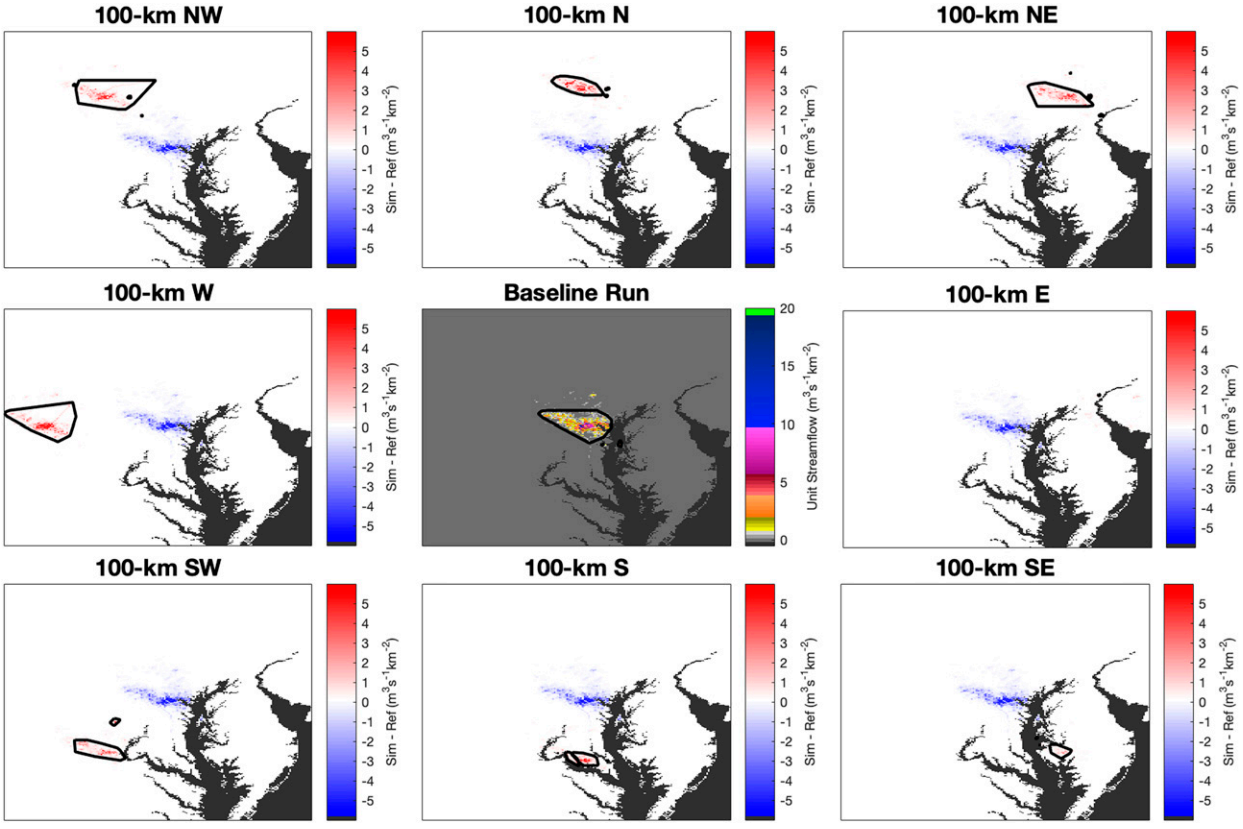

FIG. 6. Sample of unit streamflow difference maps for the eight conventional wind directions and 100-km length from the OSE scenarios. The reference maximum unit streamflow map using MRMS QPE is presented in the center panel. Black polygons are the objects representing a flash flood entity based on grids of maximum unit streamflow.

are highly skewed toward low values. The maps noticeably show the footprint of each individual ensemble member, preserving the spatial features of the correct hydrologic response. However, and even though one of the members was very close to the actual area of the event, the discrete nature of this method results in an incomplete picture of location uncertainty as evidenced by the gaps in between the shifted directions. The impact of these gaps can be further evidenced in the time series from the USGS 01589352 and the Tiber Mouth River locations (bottom A and C panels) on the northern part of the true impacted area, where no C2020 members had coverage. Consequently, the ensembles failed to encapsulate the reference run due to underdispersion. The other two locations (bottom B and D panels) had a better performance because the closest C2020 member was centered around the southern part of the true impacted area where the USGS stations are situated (see Fig. 4b). Note that because the displacement length was known a priori, the shifting of precipitation fields was done with a fixed length. Clearly, shifting in more directions and distances would result in a more complete representation of location uncertainty. However, this can only be accomplished at the cost of high computational demands.

Figures 9 and 10 present the results of the application of NPET for the same displacement scenario. Maps of NPET ensemble percentiles are presented in Fig. 9 to assess areal flash flooding forecasting skill. As discussed in section 2, these NPET ensembles are highly skewed toward low values. This is particularly evident for the isotropic implementation (top panels), for which only the 99th percentile shows meaningful values and spatial patterns of unit streamflow largely because of the inherent high degree of smoothing that comes with this sampling strategy. Nonetheless, this basic configuration of NPET accomplishes the goal of producing forecasts that can communicate spatial uncertainty due to QPF location errors. In contrast with the C2020 simulation, the description of uncertainty is spatially continuous, and so most of the impacted area is encapsulated by the NPET fields. However, the intensity and small-scale features of the hydrologic response are dampened. This is an expected outcome that is also due to the inherent smoothing of the probabilistic sampling in NPET. As the sampling is targeted over a smaller region, the intensity of unit streamflow increases as evidenced at the different percentiles, which is due to ensembles that are less skewed toward low values. Moreover, the sampling that used the octant sector (i.e., the highest degree of sampling precision) shows a very strong signal over the correct flash flood threat area.

The effect of varying parameter $\mathbb{D}$ on ensemble generation with NPET is illustrated in Fig. 10. The time series of ensembles at the four selected locations are obtained by an isotropic implementation of NPET. The matrix of ensemble plots is

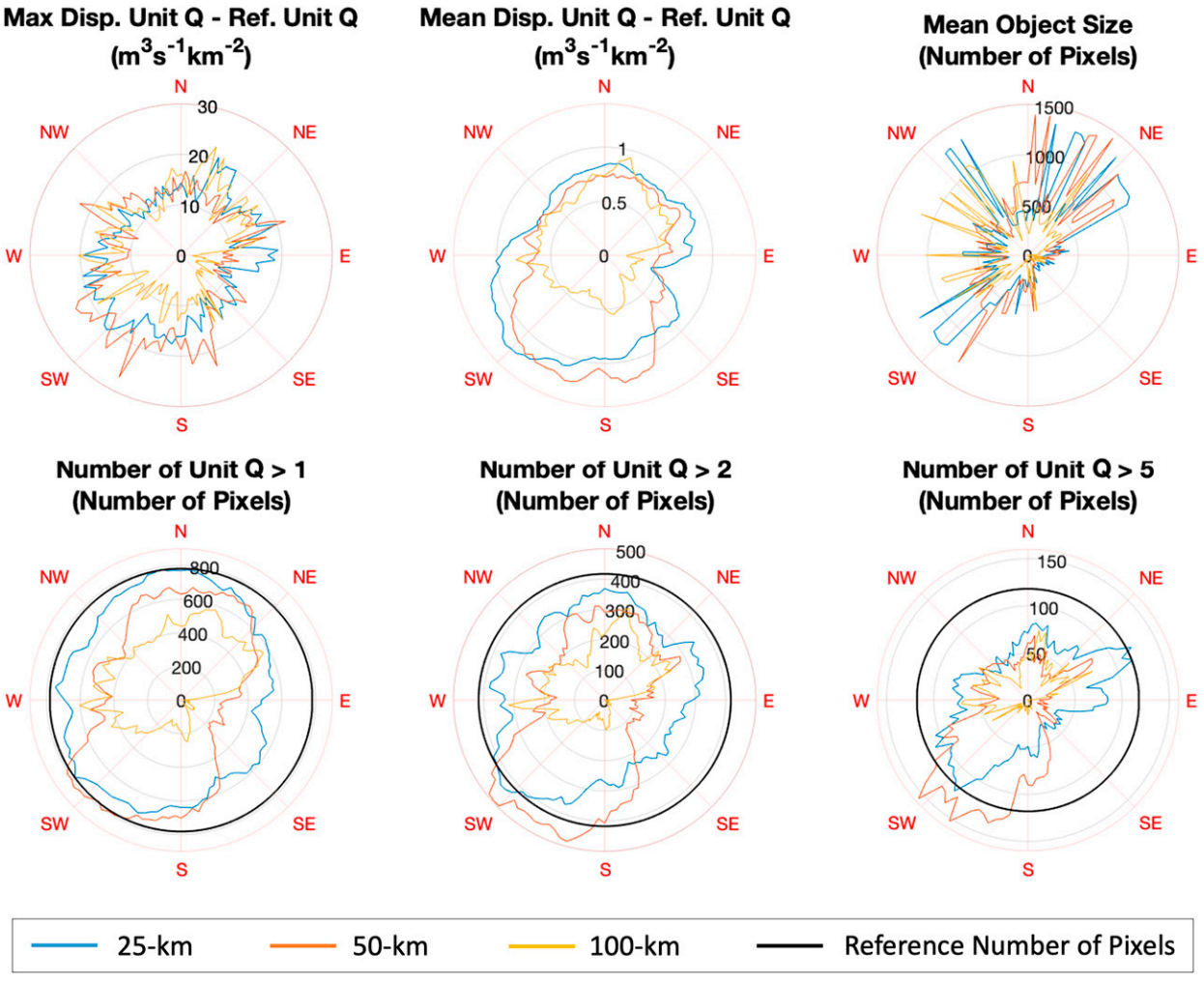

FIG. 7. Polar plots of various error metrics over the areas of overestimation for the 360 OSE scenarios of displacement with the steady-state assumption.

organized by placing each location on a column, and each particular value of $\mathbb{D}$ on a row sorted in descending order. A noteworthy general observation is that these configurations of NPET result in ensembles with very large sample sizes (on the order of thousands of members). The effect of regulating the hydrologic similarity restriction with $\mathbb{D}$ on the resulting size of the ensembles is clear. As the value of $\mathbb{D}$ increases, the restriction relaxes and the number of members dramatically increases. This in turn leads to ensembles with more diverse values of unit streamflow, which is demonstrated by the changes in the distributions. However, the size of the ensembles produced by NPET with Eq. (5) depend on the catchment-scale: as the drainage area increases, the ensemble size decreases, which limits the capability to modify the ensemble attributes through $\mathbb{D}$. For the largest basin (USGS 01594000), only when the restriction is relaxed the most (i.e., $\mathbb{D} = 50\%$) was the ensemble able to encapsulate the reference simulation. Although these large ensembles include members with significant hydrologic response consistent with the reference simulation, their distributions are heavily skewed toward low values as evidenced by only the 99.9th percentile encapsulating the reference simulation. As discussed before, this is an expected outcome particularly for an isotropic sampling that implies a high degree of uncertainty about the displacement of rainfall. Nevertheless, when comparing all ensembles to the deterministic forecast, it is evident that the application of NPET offers better information about the potential magnitude of the hydrologic response. Furthermore, NPET produces ensembles that can encapsulate the reference simulation at all locations in contrast to the C2020 method (Fig. 8).

b. *Real QPF demo*

Figure 11 presents the preliminary analysis of HRRR forecast showing two distinct displacement characteristics between the northern and southern halves of the analysis domain. For the northern area, the storm systems that impacted Ellicott City moved from the northwest to the

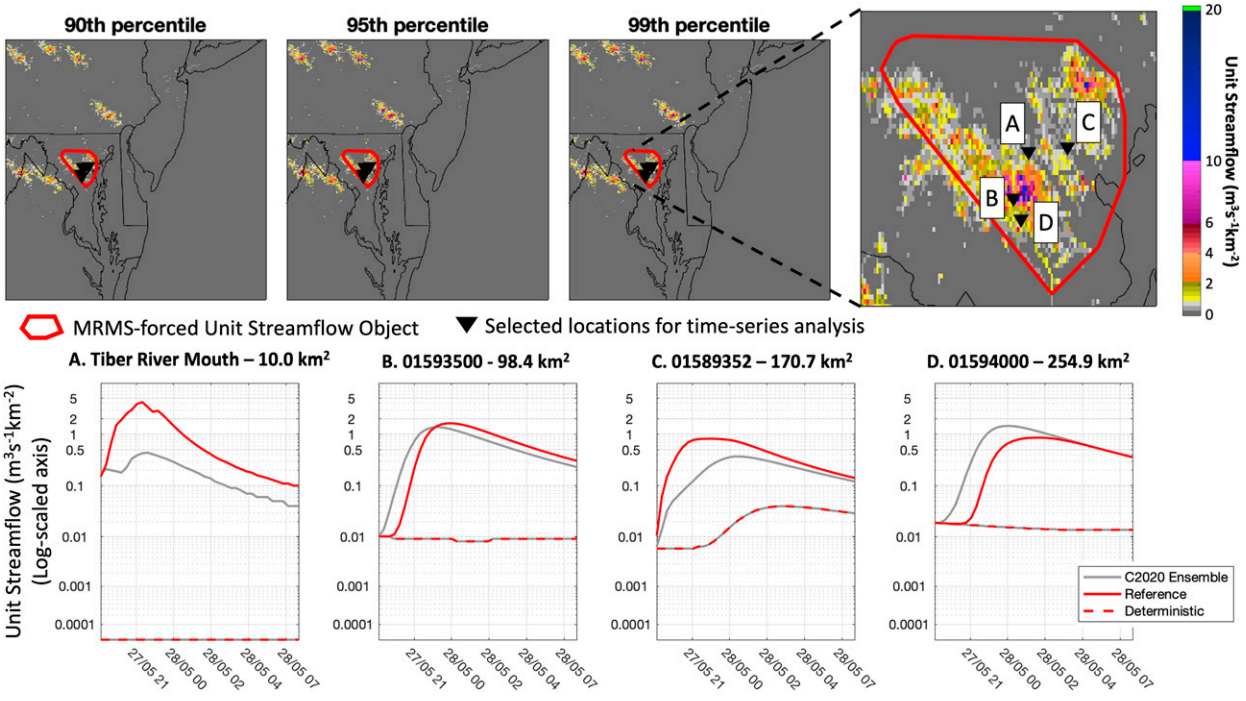

FIG. 8. C2020 8-member ensemble simulations: (top) maps of selected ensemble percentiles and (bottom) time series for the four selected locations. The A–D labels in inset zoomed in map correspond to the titles of four time series plots.

southeast, but the HRRR moved the storms too slowly, which resulted in rainfall forecasts generally displaced and dispersed toward the northwest. The southern area, however, did not show a well-defined displacement tendency.

The above QPF displacement analysis for the HRRR was used to implement NPET with isotropic sampling and a neighborhood size of 150-km radius (top panels of Fig. 12). Then, the subjective and nonexhaustive calibration produced an implementation with two subdomains separated by the white dashed line shown in the bottom panels of Fig. 12. An anisotropic sampling was defined for the northern subdomain with parameters $r = 150$ km, $\mathbb{D} = 50\%$, $\mathbf{E} = (-40.0, 70.0$ km), and $\omega = 100$. The southern subdomain was set up for an isotropic sampling with parameters $r = 50$ km, $\mathbb{D} = 50\%$, $\mathbf{E} = (0, 0$ km), and $\omega = 0$. Similar to the OSE scenario, the isotropic sampling applied over the entire domain yields a large, smoothed field of unit streamflow values, with only the 99th percentile presenting nonzero values around the impacted area. Once again, it is worthwhile underlining that even with this basic configuration of NPET, the ensembles can convey the uncertainty in the hydrologic forecast that is due to location biases in the QPFs. Likewise, the calibrated configuration of NPET can more precisely characterize the location uncertainty with values along the northwest path of displaced rainfall evidenced on the 90th–99th percentiles fields, while correctly highlighting the impacted area. These results, first, are indicative of a better characterization of the distribution of unit streamflow values by the ensembles, and second, demonstrate the value of the capability of NPET to take advantage of available information about the displacement in QPFs. Likewise, the use of subdomains illustrates the application of NPET with configurations that can vary in space, which is applicable to location biases that are region-specific.

The hydrographs in Fig. 13 show variable skill in characterizing the distribution of unit streamflows and encapsulating the reference simulation. First, the deterministic forecast greatly underestimates the hydrologic response for three out of the four selected locations, so it is immediately clear that all the ensemble-based forecasts offered better information in those cases. The deterministic forecast at the USGS 01589352 stream gauge location is the exception, underestimating the peak unit streamflow by approximately 25%, which can be considered a reasonably skillful forecast. Second, and like the results from the OSE experiments, the ensembles generated with the isotropic configuration are heavily skewed toward low values. Nonetheless, these ensemble-based forecasts still present a more complete picture of possibly higher values of unit streamflow than the deterministic ones. Last, the anisotropic configuration produced ensembles that are significantly less skewed toward low values, which shows that the level of sampling precision can improve the attributes of the distributions. However, the spread of the ensembles is consistently low for all locations and so these NPET forecasts are not able to encapsulate the reference simulations. Considering these results and those presented in Fig. 12, placing too much confidence over a particular sampling region can result in better areal description of the probability of flash flood threat but at the expense of biased descriptions at specific locations.

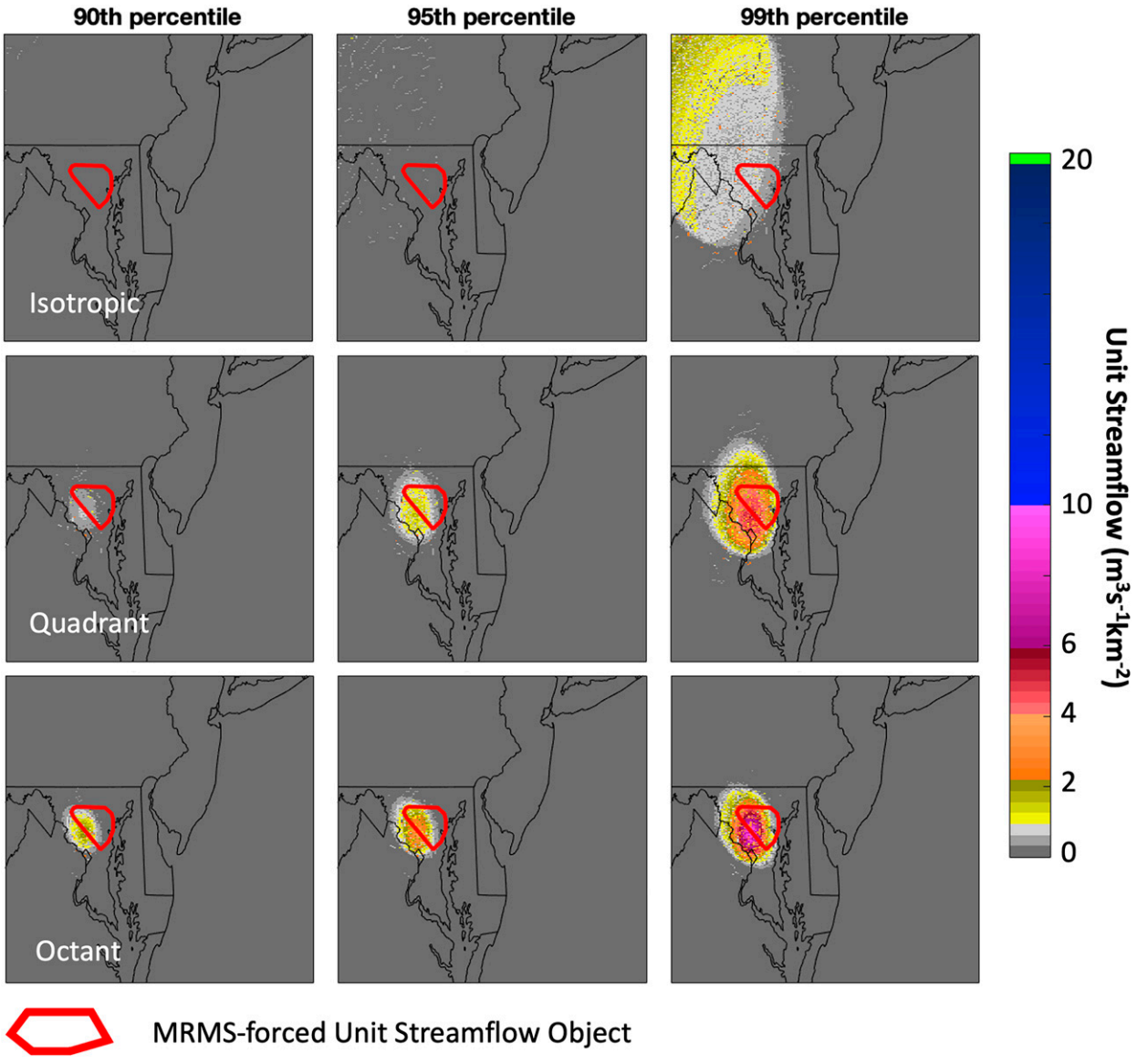

FIG. 9. Selected ensemble percentiles for ensembles with NPET. Each row corresponds to each of the NPET neighborhoods in Fig. 4 with parameter $\mathbb{D}$ fixed to 25%.

5. Summary and conclusions

This work introduces the neighboring pixel ensemble technique (NPET), a novel method that can generate ensembles of hydrologic forecasts from a single, deterministic hydrologic model run to account for locational uncertainties in QPFs. The main motivation behind the development of the technique is the challenge in increasing lead times for flash flood warnings using QPFs as inputs to the FLASH system used in U.S. NWS operations. Doing this is challenging because, first, QPFs are subject to various sources of uncertainty that lead to different types of errors, including location errors, which propagate to the resultant hydrologic forecasts. Second, ensemble-based methods commonly used to characterize uncertainties in QPFs require multiple hydrologic model runs, one for each member of the ensemble of QPFs, to produce the associated ensemble of hydrologic outputs. The computational costs of this approach are directly proportional to the size of the ensembles of QPFs, which must be large to properly approximate the correct distribution of precipitation forecast's attributes. The computational resources available at NOAA's NCEP, where FLASH runs for flash flood warning operations covering the CONUS and outer U.S. territories, limit the implementation of methods requiring large number of hydrologic model iterations, and warrants efficient techniques like NPET for probabilistic methods to be feasible in operations.

Because NPET is agnostic of the QPFs or the hydrologic model, this study presented the technique in a generalized form so it can be applied to a variety of problems involving

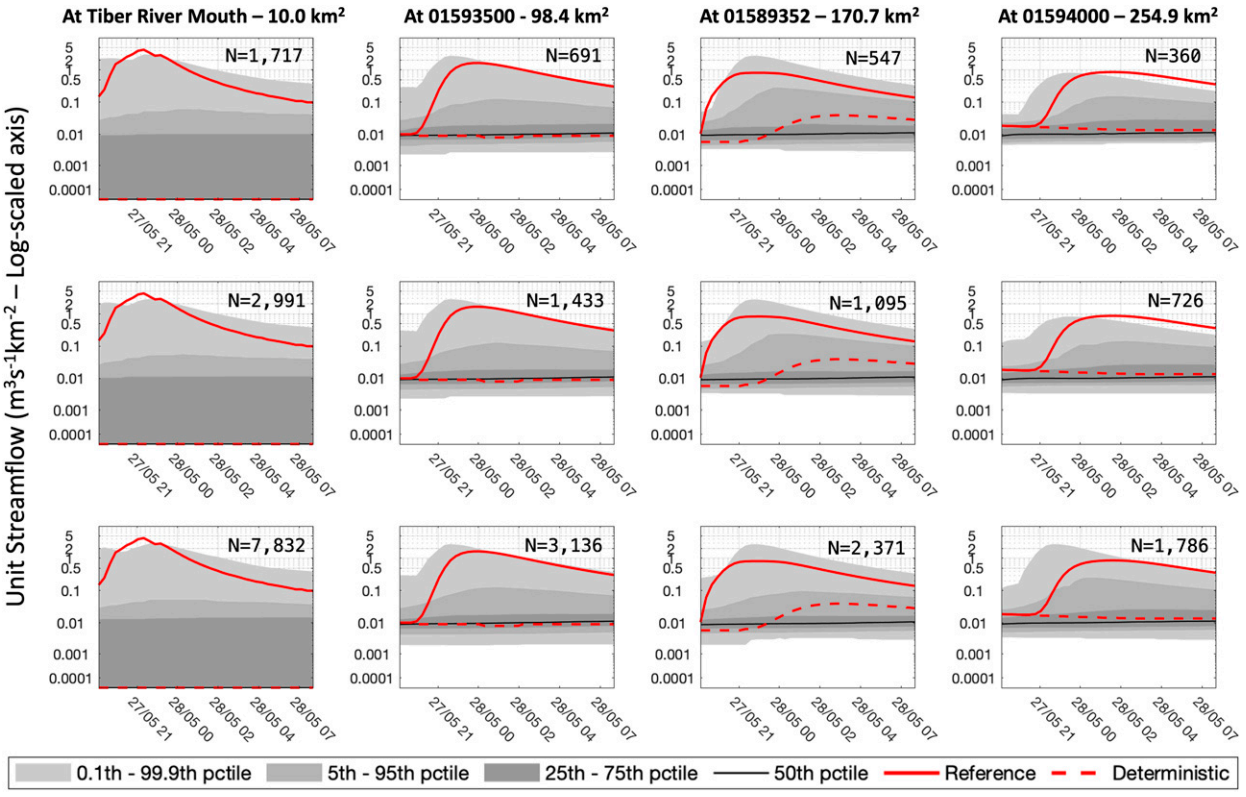

FIG. 10. Isotropic NPET ensemble percentile time series for (top) $\mathbb{D} = 12.5\%$, (middle) $\mathbb{D} = 25\%$, and (bottom) $\mathbb{D} = 50\%$, for the four selected locations. The value of N within each graph box is the number of ensemble members.

rainfall forecasts as forcing in a hydrologic forecasting system. The technique was demonstrated for the flash flood event in Ellicott City on 27 May 2018 using operational tools employed for warning operations in the U.S. NWS: MRMS QPEs used for real-time rainfall estimation, HRRR QPFs for rainfall forecasting, and a FLASH hydrologic model used for flash flood forecasting. The demonstration included an OSE to explore basic aspects of hydrologic forecast errors stemming from QPF's locational uncertainty and the way to address them with NPET. Last, an NPET application exercise was conducted emulating an operational run with QPFs from the HRRR.

Results from the OSE highlighted the importance of the terrain heterogeneity when assessing the hydrologic impact of displaced rainfall. Differences between the reference hydrologic simulation (nature run or truth) and those based on displaced rainfall fields (perturbations) were significant in magnitude and spatial structure. These differences can dramatically change the interpretation of the hydrologic response as it relates to the issuance of a flash flood warning. Therefore, methods attempting to transpose displaced hydrologic responses need to pay attention to this consideration. NPET is designed to be able to take this into account with an equation that can take any real-valued function formulated to transfer displaced hydrologic responses to forecast locations. A limitation of this function is its inability to recover displaced hydrologic response from areas where the model does not produce outputs (e.g., over lakes and the ocean). In this demonstration work, a simple equation based on the contributing area was employed. The role of this function in NPET was only evident when examining ensembles at specific locations. Its main effect was on the size of the ensemble by controlling the eligibility of sampled pixels. Possibly, a more sophisticated transfer function can meaningfully change the distribution attributes, which invites exploration in future endeavors.

One of the main important takeaways from this demonstration study is about the capability of NPET to produce very large ensembles of hydrologic outputs from a single deterministic output, which was on the order of thousands of members for the configurations used herein. Furthermore, the comparison with the C2020 method clearly showed NPET's superior ability to describe location uncertainty through the spatial continuity of the probabilistic fields and encapsulation of the reference simulation, even though the intensity and small-scale features of unit streamflows were smoothed out. The uncalibrated nature of the ensembles and the inherent smoothing effect of the spatial sampling in NPET can explain their tendency to be heavily skewed to low values, which poses challenges for the interpretation of the conveyed probabilistic information. However, results from the experiments that adjusted the sampling based on displacement information showed that there are opportunities to improve the ensemble's attributes. An advisable consideration when applying higher degrees of precision in the spatial sampling

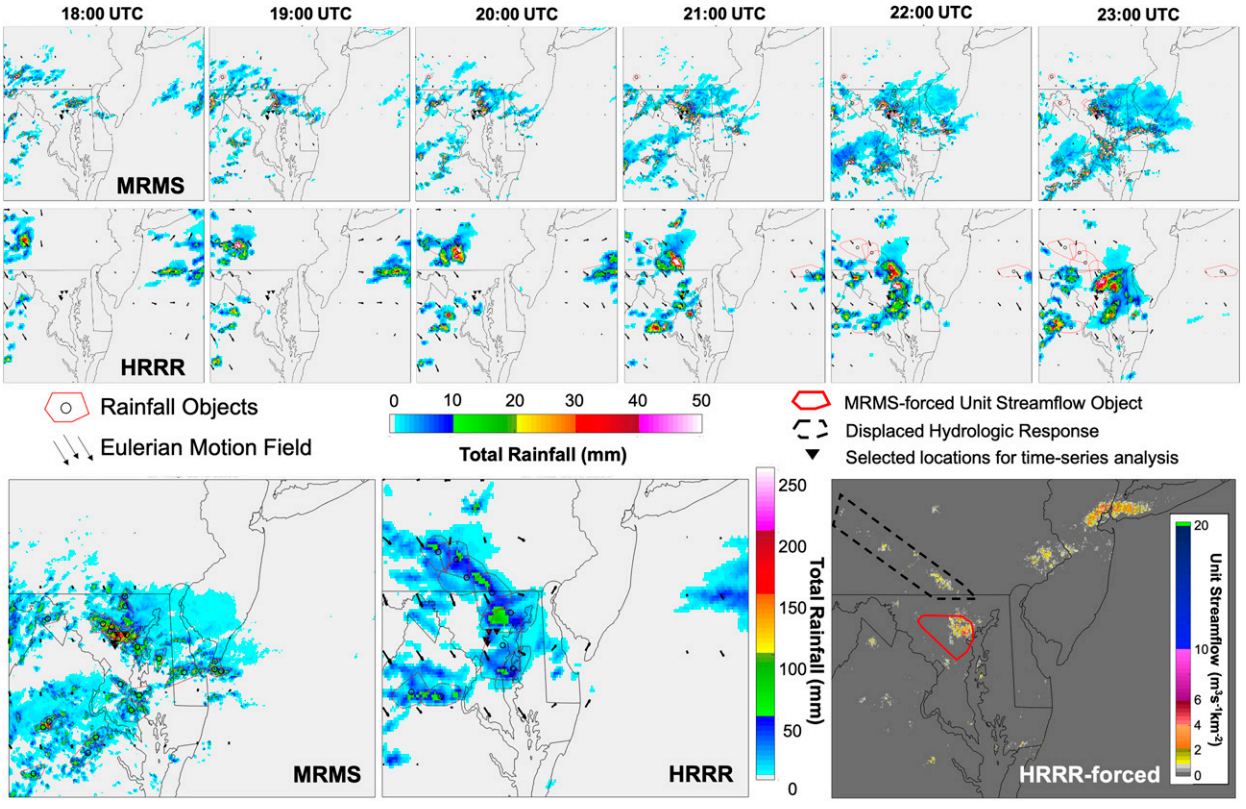

FIG. 11. Displacement analysis of HRRR forecast valid at 1300 UTC. Hourly accumulations of rainfall for (top) MRMS and (middle) the HRRR for the period 1800–2300 UTC 27 May 2018. (bottom) Summary plots of the same period are also included.

with NPET is the trade-off between improving the characterization of location uncertainty for spatial analysis and the improvements achieved at the pixel scale for specific location analysis. Also, and as discussed above, it is possible to attain improvements using more sophisticated transfer functions. The results presented herein demonstrate that NPET ensembles effectively accomplish the goal of conveying uncertainty about precipitation location in QPFs.

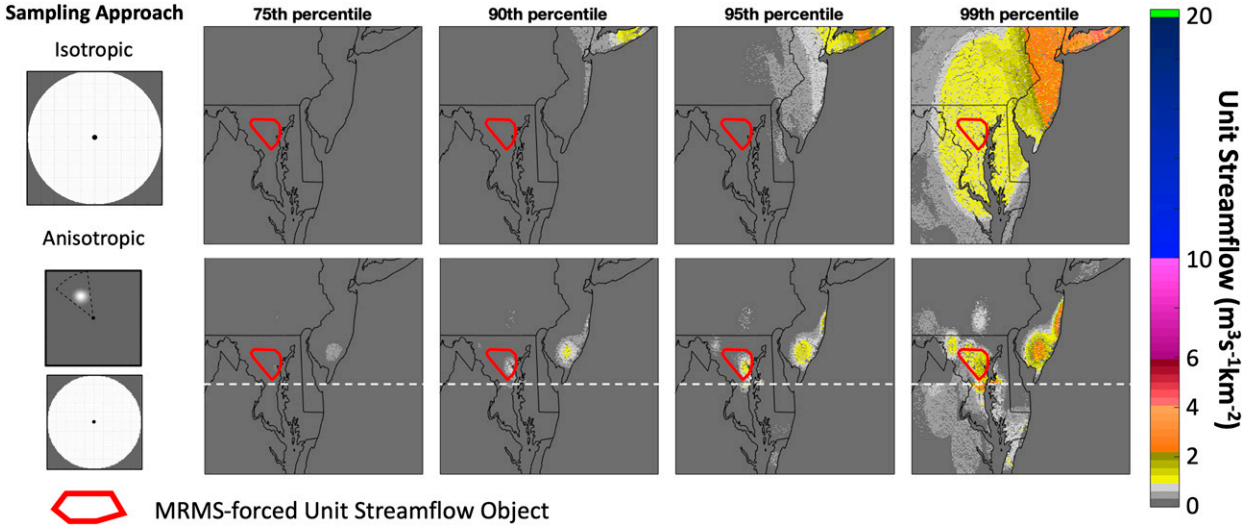

FIG. 12. NPET ensemble percentiles for the HRRR-driven hydrologic forecast: (top) isotropic sampling and (bottom) hybrid anisotropic (above the horizontal white dashed line) and isotropic sampling (below the horizontal white dashed line).

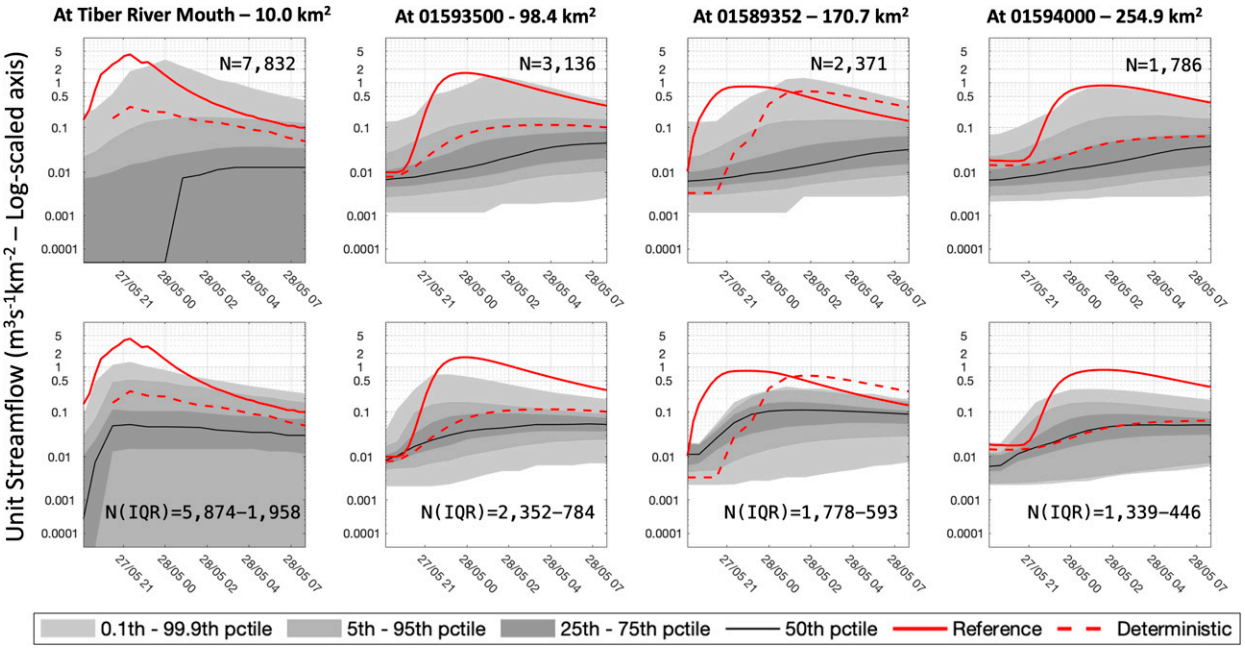

FIG. 13. NPET ensemble percentile time series for the two configurations applied to HRRR-driven hydrologic forecasts: (top) isotropic sampling and (bottom) anisotropic sampling. Sample sizes are provided for the isotropic configuration, while interquartile ranges (IQR) of sample sizes are provided for the anisotropic configuration.

Overall, the application of NPET to FLASH products suggests that the technique could yield beneficial applications in flash flood warning operations in the U.S. NWS. Subsequent research will be needed, however, to propose ways in which the QPF location uncertainty information content in NPET forecasts can be communicated and translated into flash flood warnings, similar to other efforts aimed at learning how forecasters interpret and use probabilistic outputs (e.g., Martinaitis et al. 2023; Yussouf et al. 2020). The demonstrated value of NPET in the presence of a priori precipitation displacement knowledge will also warrant further studies focused on strategies to characterize these errors. Developments at NOAA's Weather Prediction Center (WPC) with the Tracking of Heavy Precipitation Objects (THePrO; Erickson and Nelson 2020) has resulted in a database of QPF object-based biases which can inform NPET for its use in operations to specify parameters **E** and *r*. This object tracking framework could also be extended to generate short-term information of the most recent HRRR forecasts to continuously update these NPET parameters. Although the main reason to develop NPET was to operate without the need of multiple hydrologic model runs, it is feasible to extend the implementation of the technique to systems that utilize limited sized ensembles of QPFs, which is typical of operational ensemble-based NWPs (e.g., NOAA's GEFS v12 with 31 members).

Acknowledgments. This research was funded by the NOAA/OAR–Joint Technology Transfer Initiative (JTTI) program under Federal Award NA20OAR4590354, U.S. Department of Commerce, National Oceanic and Atmospheric Administration. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NOAA, The University of Oklahoma, or the U.S. Department of Commerce.

Data availability statement. The numerical model simulations and datasets upon which this study is based are too large to archive or to transfer in their entirety. We provide all the information, datasets, and computer programs needed to replicate the simulations using a combination of code written in MATLAB and Python. This material is available at https://github.com/HyDROSLab/NPET-Study [https://doi.org/10.5281/zenodo.6787491].

REFERENCES

Baum, R. L., and J. W. Godt, 2010: Early warning of rainfall-induced shallow landslides and debris flows in the USA. *Landslides*, **7**, 259–272, https://doi.org/10.1007/s10346-009-0177-0.

Bedient, P. B., W. C. Huber, and B. E. Vieux, 2008: *Hydrology and Floodplain Analysis*. Prentice Hall, 795 pp.

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Carlberg, B., K. Franz, and W. Gallus Jr., 2020: A method to account for QPF spatial displacement errors in short-term ensemble streamflow forecasting. *Water*, **12**, 3505, https://doi.org/10.3390/w12123505.

Clark, A. J., 2017: Generation of ensemble mean precipitation forecasts from convection-allowing ensembles. *Wea. Forecasting*, **32**, 1569–1583, https://doi.org/10.1175/WAF-D-16-0199.1.

Clark, R. A., J. J. Gourley, Z. L. Flamig, Y. Hong, and E. Clark, 2014: CONUS-wide evaluation of national weather service flash flood guidance products. *Wea. Forecasting*, **29**, 377–392, https://doi.org/10.1175/WAF-D-12-00124.1.

Cuo, L., T. C. Pagano, and Q. J. Wang, 2011: A review of quantitative precipitation forecasts and their use in short- to medium-range streamflow forecasting. *J. Hydrometeor.*, **12**, 713–728, https://doi.org/10.1175/2011JHM1347.1.

Day, G. N., 1985: Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plann. Manage.*, **111**, 157–170, https://doi.org/10.1061/(ASCE)0733-9496(1985)111:2(157).

Dowell, D. C., and Coauthors, 2022: The High-Resolution Rapid Refresh (HRRR): An hourly updating convection-allowing forecast model. Part I: Motivation and system description. *Wea. Forecasting*, **37**, 1371–1395, https://doi.org/10.1175/WAF-D-21-0151.1.

Duda, J. D., and W. A. Gallus Jr., 2013: The impact of large-scale forcing on skill of simulated convective initiation and upscale evolution with convection-allowing grid spacings in the WRF. *Wea. Forecasting*, **28**, 994–1018, https://doi.org/10.1175/WAF-D-13-00005.1.

Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461–2480, https://doi.org/10.1175/1520-0493(2001)129<2461:AOAPMS>2.0.CO;2.

——, and J. L. McBride, 2000: Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrol.*, **239**, 179–202, https://doi.org/10.1016/S0022-1694(00)00343-7.

Erickson, M. J., and J. A. Nelson, 2020: Development and usage of the Heavy Precipitation Object Tracker (HPOT). *2020 Flash Flood and Intense Rainfall Experiment (FFaIR)*, College Park, MD, NOAA, https://origin.wpc.ncep.noaa.gov/verification/mtd/about.php.

Flamig, Z. L., H. Vergara, and J. J. Gourley, 2020: The ensemble framework for flash flood forecasting (EF5) v1.2: Description and case study. *Geosci. Model Dev.*, **13**, 4943–4958, https://doi.org/10.5194/gmd-13-4943-2020.

Fritsch, J. M., and R. E. Carbone, 2004: Improving quantitative precipitation forecasts in the warm season: A USWRP research and development strategy. *Bull. Amer. Meteor. Soc.*, **85**, 955–966, https://doi.org/10.1175/BAMS-85-7-955.

Ghimire, G. R., W. F. Krajewski, and F. Quintero, 2021: Scale-dependent value of QPF for real-time streamflow forecasting. *J. Hydrometeor.*, **22**, 1931–1947, https://doi.org/10.1175/JHM-D-20-0297.1.

Gourley, J. J., and H. Vergara, 2021: Comments on "flash flood verification: Pondering precipitation proxies." *J. Hydrometeor.*, **22**, 739–747, https://doi.org/10.1175/JHM-D-20-0215.1.

——, and Coauthors, 2017: The FLASH project: Improving the tools for flash flood monitoring and prediction across the United States. *Bull. Amer. Meteor. Soc.*, **98**, 361–372, https://doi.org/10.1175/BAMS-D-15-00247.1.

Hall, J., and Coauthors, 2014: Understanding flood regime changes in Europe: A state-of-the-art assessment. *Hydrol. Earth Syst. Sci.*, **18**, 2735–2772, https://doi.org/10.5194/hess-18-2735-2014.

Hardy, J., J. J. Gourley, P.-E. Kirstetter, Y. Hong, F. Kong, and Z. L. Flamig, 2016: A method for probabilistic flash flood forecasting. *J. Hydrol.*, **541**, 480–494, https://doi.org/10.1016/j.jhydrol.2016.04.007.

James, E. P., and Coauthors, 2022: The High-Resolution Rapid Refresh (HRRR): An hourly updating convection-allowing forecast model. Part II: Forecast performance. *Wea. Forecasting*, **37**, 1397–1417, https://doi.org/10.1175/WAF-D-21-0130.1.

Jankov, I., P. J. Schultz, C. J. Anderson, and S. E. Koch, 2007a: The impact of different physical parameterizations and their interactions on cold season QPF in the American River basin. *J. Hydrometeor.*, **8**, 1141–1151, https://doi.org/10.1175/JHM630.1.

——, W. A. Gallus Jr., M. Segal, and S. E. Koch, 2007b: Influence of initial conditions on the WRF–ARW Model QPF response to physical parameterization changes. *Wea. Forecasting*, **22**, 501–519, https://doi.org/10.1175/WAF998.1.

Jeffreys, H., and B. Jeffreys, 1999: *Methods of Mathematical Physics*. 3rd ed. Cambridge University Press, 730 pp.

Martinaitis, S. M., and Coauthors, 2017: The HMT multi-radar multi-sensor hydro experiment. *Bull. Amer. Meteor. Soc.*, **98**, 347–359, https://doi.org/10.1175/BAMS-D-15-00283.1.

——, and Coauthors, 2023: A path toward short-term probabilistic flash flood prediction. *Bull. Amer. Meteor. Soc.*, https://doi.org/10.1175/BAMS-D-22-0026.1, in press.

Masutani, M., and Coauthors, 2010: Observing system simulation experiments. *Data Assimilation*, W. Lahoz, B. Khattatov, and R. Menard, Eds., Springer, 647–679, https://doi.org/10.1007/978-3-540-74703-1_24.

McEnery, J., J. Ingram, Q. Duan, T. Adams, and L. Anderson, 2005: NOAA's advanced hydrologic prediction service: Building pathways for better science in water forecasting. *Bull. Amer. Meteor. Soc.*, **86**, 375–386, https://doi.org/10.1175/BAMS-86-3-375.

Podlubny, I., 1998: *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of their Solution and Some of their Applications*. 1st ed. Academic Press, 340 pp.

Rasheed, Z., A. Aravamudan, A. Gorji Sefidmazgi, G. C. Anagnostopoulos, and E. I. Nikolopoulos, 2022: Advancing flood warning procedures in ungauged basins with machine learning. *J. Hydrol.*, **609**, 127736, https://doi.org/10.1016/j.jhydrol.2022.127736.

Schaake, J., and L. Larson, 1998: Ensemble streamflow prediction (ESP): Progress and research needs. *Special Symp. on Hydrology*, Boston, MA, Amer. Meteor. Soc., J19–J24.

Seo, D. J., H. D. Herr, and J. C. Schaake, 2006: A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble streamflow prediction. *Hydrol. Earth Syst. Sci. Discuss.*, **2006**, 1987–2035, https://doi.org/10.5194/hessd-3-1987-2006.

Slater, L. J., and G. Villarini, 2017: Evaluating the drivers of seasonal streamflow in the U.S. Midwest. *Water*, **9**, 695, https://doi.org/10.3390/w9090695.

Smith, J. A., G. N. Day, and M. D. Kane, 1992: Nonparametric framework for long-range streamflow forecasting. *J. Water Resour. Plann. Manage.*, **118**, 82–92, https://doi.org/10.1061/(ASCE)0733-9496(1992)118:1(82).

Stensrud, D. J., J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107, https://doi.org/10.1175/1520-0493(2000)128<2077:UICAMP>2.0.CO;2.

Vergara, H., P.-E. Kirstetter, J. J. Gourley, Z. L. Flamig, Y. Hong, A. Arthur, and R. Kolar, 2016: Estimating a-priori kinematic wave model parameters based on regionalization for flash flood forecasting in the conterminous United States. *J. Hydrol.*, **541**, 421–433, https://doi.org/10.1016/j.jhydrol.2016.06.011.

Viterbo, F., and Coauthors, 2020: A multiscale, hydrometeorological forecast evaluation of national water model forecasts of

the May 2018 Ellicott City, Maryland, flood. *J. Hydrometeor.*, **21**, 475–499, https://doi.org/10.1175/JHM-D-19-0125.1.

Wang, J., and Coauthors, 2011: The Coupled Routing and Excess Storage (CREST) distributed hydrological model. *Hydrol. Sci. J.*, **56**, 84–98, https://doi.org/10.1080/02626667.2010.543087.

Wernli, H., M. Paulat, M. Hagen, and C. Frei, 2008: SAL—A novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Wea. Rev.*, **136**, 4470–4487, https://doi.org/10.1175/2008MWR2415.1.

Yan, H., and W. A. Gallus Jr., 2016: An evaluation of QPF from the WRF, NAM, and GFS models using multiple verification methods over a small domain. *Wea. Forecasting*, **31**, 1363–1379, https://doi.org/10.1175/WAF-D-16-0020.1.

Yussouf, N., K. A. Wilson, S. M. Martinaitis, H. Vergara, P. L. Heinselman, and J. J. Gourley, 2020: The coupling of NSSL Warn-on-Forecast and FLASH systems for probabilistic flash flood prediction. *J. Hydrometeor.*, **21**, 123–141, https://doi.org/10.1175/JHM-D-19-0131.1.

Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638, https://doi.org/10.1175/BAMS-D-14-00174.1.