A High-Resolution Quantitative Precipitation Estimate over Alaska through Kriging-Based Merging of Rain Gauges and Short-Range Regional Precipitation Forecasts

BRETT T. HOOVER,[a] JASON A. OTKIN,[a] EUGENE M. PETRESCU,[b] AND EMILY NIEBUHR[b]

[a] *Cooperative Institute for Meteorological Satellite Studies, Space Science and Engineering Center, University of Wisconsin–Madison, Madison, Wisconsin*
[b] *NOAA/National Weather Service Anchorage, Anchorage, Alaska*

(Manuscript received 22 October 2021, in final form 3 February 2022)

ABSTRACT: A method is presented to generate quantitative precipitation estimates over Alaska using kriging to merge sparse, unevenly distributed rain gauge observations with quantitative precipitation forecasts from a three-member ensemble of high-resolution numerical weather prediction models. The estimated error variance of the analysis is computed by starting with the estimated error variance from kriging and then refining the variance in *k*-fold cross validation by an empirically derived inflation factor. The method combines dynamical model forecast information with observational data to deliver a best linear unbiased estimate of precipitation, along with an analysis uncertainty estimate, that provides a much-needed precipitation analysis in a region where sparse in situ observations, poor coverage by remote sensing platforms, and complex terrain introduce large uncertainties that need to be quantified. For 6-hourly accumulation estimates produced four times daily from 1 August 2019 to 31 July 2020, three analysis configurations are tested to measure the value added by including model forecast data and how those data are best utilized in the analysis. Several directions for further improvement and validation of the analysis product are provided.

KEYWORDS: Precipitation; Interpolation schemes; Numerical analysis/modeling; Uncertainty

1. Introduction

A high-resolution quantitative precipitation estimate (QPE) is difficult to achieve over Alaska for several reasons. First, the expansive and highly variable geography of Alaska, combined with very low population density outside Anchorage, Fairbanks, and Juneau, makes maintenance of a high-density rain gauge network difficult or impossible. Between rain gauge networks whose data are distributed by MesoWest and the Alaska-Pacific River Forecast Center (APRFC), there are approximately 1400 rain gauge stations, with less than 700 stations typically reporting on any given day, and even fewer stations reporting regularly enough to provide an accumulated QPE over several hours. The majority of these stations are concentrated within high-population-density areas (Fig. 1a). Although the network provides good coverage in these locations, it leaves the rest of Alaska, a landmass that is approximately 21% of the geographic area of the contiguous United States, poorly observed. Gauge observations are further complicated by seasonality, with only a subset of heated gauges capable of providing accurate observations in the winter, while many other gauges do not report during the coldest part of the year or are plagued by inaccurate reports as frozen precipitation is not accurately collected or measured. This can happen when frozen precipitation accumulates around and eventually blocks the gauge orifice (Martinaitis et al. 2015), and underreporting can occur from low catch efficiency in windy conditions, evaporation, or sublimation of precipitation before measurement (Rasmussen et al. 2012), depending on the method of measurement and design of the gauge.

Second, remotely sensed precipitation estimates are also sparse; Alaska has seven National Weather Service (NWS) radars, three of which are located along the Gulf of Alaska (Biorka Island, Middleton Island, and Kenai), one at the base of the Alaska Peninsula (King Salmon), two along the coast of the Bering Sea (Bethel and Nome), and a single inland radar located to the northeast of Fairbanks (Pedro Dome). As shown in Fig. 1c, this leaves the vast majority of Alaska without radar coverage, thereby inhibiting development of radar-derived QPE products. Most of Alaska is located poleward of 60°N latitude, which is too far north for geostationary satellites to observe with high spatial resolution, and polar-orbiting satellite overpasses are neither complete enough nor frequent enough to provide consistent microwave-derived precipitation data. As a result, precipitation products like the Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG; Huffman et al. 2019) only consistently provide coverage for the far southern portions of Alaska.

QPE is highly important for numerical weather prediction (NWP) and hydrological model validation and calibration. APRFC operates 263 river gauges, with some watersheds and long stretches of river covered by only a single river gauge. QPE also has applications in fire weather, with even a few hundredths of an inch (a few quarter millimeters) of rain capable of reducing the risk of fire initiation and growth. Such information is increasingly important as the number of significant fire years continues to grow in frequency, with over 2.5 million acres (1 acre ≈ 0.4 ha) of land burned in Alaska in

Hoover's current affiliation: I. M. Systems Group, NOAA/NWS/NCEP/EMC, College Park, Maryland.

Corresponding author: Brett T. Hoover, brett.hoover@ssec.wisc.edu

DOI: 10.1175/JTECH-D-21-0132.1

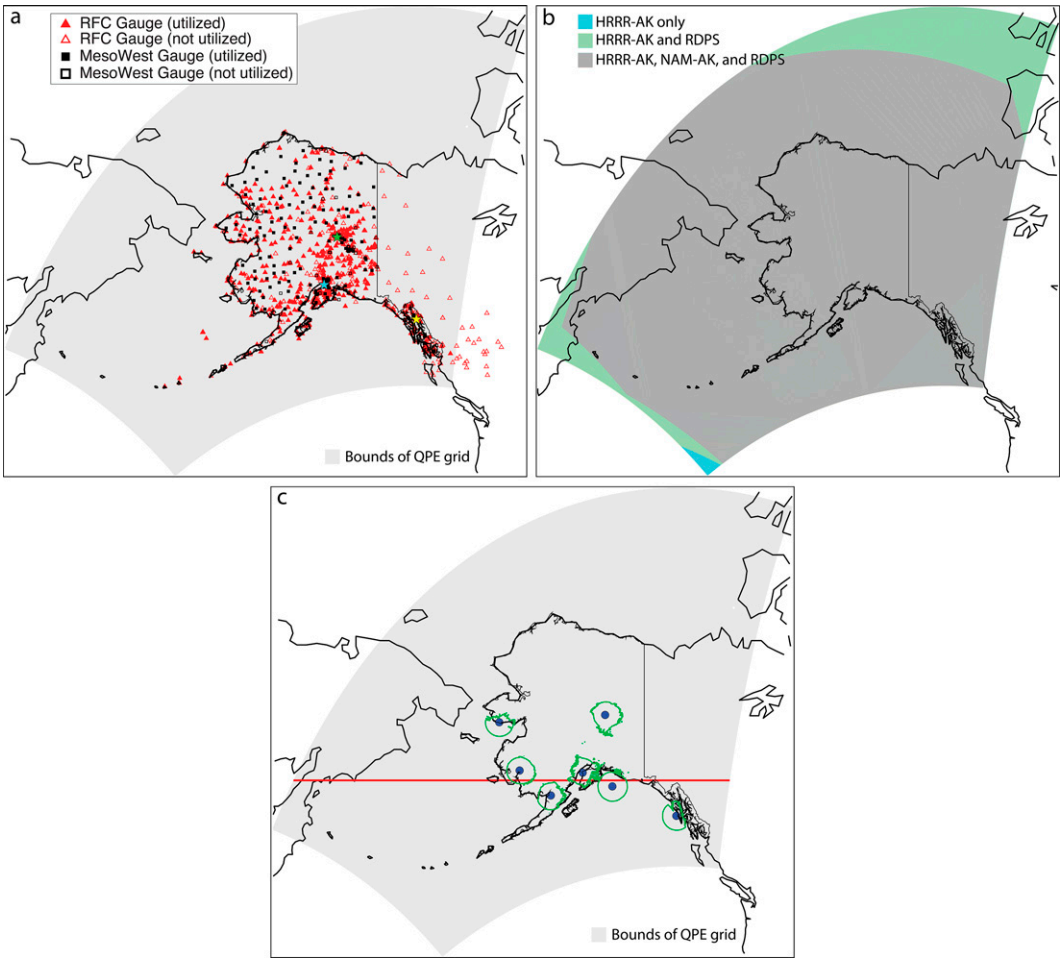

FIG. 1. (a) Map of gauge observation stations. The QPE analysis domain, based on the domain of the HRRR-AK model, is shaded in gray. Stars represent Anchorage (cyan), Fairbanks (green), and Juneau (yellow). (b) Map of QPF domains covered by each member of the QPE first guess. (c) Map of Alaska's NWS NEXRAD stations with representative range provided by nonzero MRMS Quality Indicator for 6-h radar accumulated precipitation at 0000 UTC 12 Jan 2022 (green). The 60°N latitude is highlighted in red.

2019 (https://www.usgs.gov/center-news/about-25-million-acres-alaska-have-burned-states-wildfire-seasons-are-getting-worse). With limited information available, it is important for an Alaska region QPE product to not only provide a best estimate of the accumulated precipitation, but also confidence bounds that represent a range of uncertainty on those estimates.

These challenges leave the Alaska region with few options for QPE. The Parameter–Elevation Regressions on Independent Slopes Model (PRISM; Daly et al. 1994) groups individual gauge stations into topographic facets using a digital elevation model, and then estimates precipitation within a cell through statistical regression to model relationships between precipitation and orography/elevation. The resulting product depicts climatological precipitation formed via a statistical–topographic model designed to map mountainous terrain. The PRISM climatology is used as the basis of reanalysis precipitation products extending over Alaska, like the North American Regional Reanalysis (NARR; Mesinger et al. 2006), and Mountain Mapper, a real-time precipitation product (Schaake et al. 2004). An experimental QPE product from the National Severe Storms Laboratory, the Multi-Radar Multi-Sensor (MRMS; Zhang et al. 2016) QPE, is also produced over the Alaska region and has been operational at NWS since October 2020. MRMS produces a range of QPE products from the Mountain Mapper product to a merged gauge and radar product. As of October 2020, Mountain Mapper and gauge-corrected radar QPE have been replaced or subsumed operationally by the multisensor QPE (MSQPE) product, which combines radar and gauge data with a weighted blend of Mountain Mapper and short-range quantitative precipitation forecast data to fill gaps in radar coverage (Martinaitis et al. 2020). Each of these products attempts to meet the challenges of Alaska's complex orography, large geographic extent, and limited observations by rain gauges, radar, and satellite. Notably, none of the existing QPE products described above provide a well-defined estimate of their uncertainty.

Producing a QPE through merging rain gauge observations and radar-derived accumulated precipitation is accomplished through several different techniques. Radar-derived precipitation is subject to errors from beam blockage and ducting, signal attenuation, variation in the Z–R relationship defining the conversion between reflectivity Z and rain rate R, and differences in the precipitation measured by radar at altitude and what is received at the ground (Wilson and Brandes 1979). These errors can be addressed through calibration of radar-derived precipitation by merging with gauges. The simplest calibration methods include applying a single multiplicative mean bias correction of radar-derived precipitation against gauges computed as a long-term statistical bias (e.g., Wood et al. 2000), on a per-scene basis from one or more available gauges (Wilson and Brandes 1979). More complex methods involve applying a spatially varying calibration by fitting a surface of either multiplicative adjustment factors or by adding an interpolated surface of gauge–radar residuals. These range in complexity from inverse-distance weighting techniques (e.g., Wilson and Brandes 1979) to Barnes analysis producing a field from calibration factors evaluated at each gauge site (Brandes 1975) or a least squares fitting of the relationship between radar and gauge as a function of distance (Michelson et al. 2000).

Merging of radar and rain gauge observations has also been accomplished through kriging—a geostatistical technique that is well designed for interpolation of sparse observational data to a grid. This can involve either kriging an estimate of the radar precipitation error at gauge sites to the grid and subtracting it from radar (Sinclair and Pegram 2005), kriging gauge observations to the grid while using radar-derived precipitation as an external drift variable (Velasco-Forero et al. 2008), more sophisticated techniques involving carefully modeling the precipitation field to retain the mean field of interpolated rain gauges while preserving the mean field deviations from the radar (Ehret et al. 2008), or employing Kalman filtering to minimize uncertainty between gauge-based and radar-based precipitation fields (Wang et al. 2015).

We have developed a kriging-based gridded precipitation product for the Alaska region that utilizes available rain gauge observations and, in the absence of reliable and comprehensive remote sensing observations, merges gauge data with an ensemble of short-range quantitative precipitation forecasts from operational high-resolution regional NWP models. Unlike radar data, NWP model data cover the entire Alaskan region, and NWP data have precipitation errors that are much different from the kind of errors present in radar-derived precipitation. Just as there can be inconsistencies in precipitation estimated between two overlapping radars, inconsistencies can exist between two NWP model predictions, but the source and expression of those inconsistencies can be much different from what is observed with radar. It is therefore not a foregone conclusion that NWP model data will provide value to QPE. QPE for 6-hourly accumulation is generated four times daily from 1 August 2019 to 31 July 2020 using multiple configurations of the kriging-based merging technique. The goals of this study are to 1) test the value added by NWP model data to this QPE framework by benchmarking against a QPE configuration that includes no NWP data, 2) evaluate the best configuration for utilizing NWP model data in a kriging-based merging system, and 3) utilize the kriging (error) variance at each analysis grid point to provide an empirically tuned uncertainty estimate on QPE through cross validation with available observations.

Observational data, NWP forecast data, and the QPE analysis products are discussed in section 2. The kriging method for merging data is discussed in section 3 with example analyses. An evaluation of QPE accuracy and uncertainty is presented in section 4, along with tests to determine the value added by NWP forecast data. Conclusions and potential future directions for research and development are provided in section 5.

2. QPE data, configurations, and products

a. *Rain gauge observations*

NWS rain gauge observations are available from MesoWest (Horel et al. 2002) and distributed by the Public Benefit Corporation (PBC) Synoptic via their mesonet API (https://developers.synopticdata.com/mesonet/). Precipitation observations distributed by MesoWest are produced at varying time resolutions and contain data dropouts. To produce 6-h accumulated precipitation totals from these observations we apply some basic quality assurance. Precipitation is collected from hourly reports or from 12 consecutive 5-min reports, depending on the station. Quality assurance ensures that any hourly observation is collected over a time interval that ends within 10 min of the end of the hour and that consecutive 5-min reports create an unbroken 60-min record. Any hourly observation that does not meet these requirements is flagged as not-a-number (NaN), and then observations from individual stations are summed into the four 6-h intervals for each QPE analysis; any summed 6-h observation that includes an NaN value is removed prior to kriging.

A daily set of 6-h accumulated precipitation gauge observations is also made available by APRFC that is distributed via the NWS raw text product archive maintained by Iowa State University (https://mesonet.agron.iastate.edu/). These gauges are carefully quality controlled by personnel at APRFC and include daily updated blackout station lists for suspicious observations and quality control flags to identify observations involving human intervention. The MesoWest and APRFC datasets do not fully overlap. Comparison of the station lists showed that there are 234 MesoWest stations not found in the APRFC dataset, and 647 APRFC stations not found in the MesoWest dataset. Since observations from APRFC are highly quality controlled, stations that exist in both datasets are always defaulted to APRFC values, and observations in the APRFC blacklist are always excluded. Of the total 1399 potential stations, 764 stations across Alaska are used in this study, with the remaining stations scattered across the Yukon and British Columbia (Fig. 1a). Stations are excluded either because of inconsistent reporting or because they exist outside of the model domain used to define the QPE first guess. Observation density is highest in the Anchorage, Fairbanks,

and Juneau regions, and lowest in the Brooks Range of northern Alaska. Quality assurance and quality control of gauge observations does not appear to affect the number or geographic distribution of gauge observations between 0000 and 0600, 0600 and 1200, 1200 and 1800, and 1800 and 0000 UTC analysis periods.

b. NWP model forecasts

Three publicly available, operational NWP model forecasts are used to define the quantitative precipitation forecast (QPF) grid: the National Centers for Environmental Prediction (NCEP) High Resolution Rapid Refresh (HRRR) over the Alaska region (HRRR-AK), the NCEP North American Model (NAM), and Environment Canada's Regional Deterministic Prediction System (RDPS). The HRRR-AK is a 3-km-resolution model forecast using the Advanced Research version of WRF (WRF-ARW), version 3.8.1 (Skamarock et al. 2008), and is initialized from the operational version 3 Rapid Refresh analysis (Benjamin et al. 2016). The HRRR-AK is initialized eight times daily, at 3-h intervals between 0000 and 2100 UTC, with model output available on National Oceanic and Atmospheric Administration (NOAA) FTP (ftp://ftp.ncep.noaa.gov/pub/data/nccf/com/hrrr/prod/). The NAM is a NOAA Environmental Modeling System (NEMS) framework model using the nonhydrostatic multiscale model on B-grid (NMMB) dynamical core (Janjic and Gall 2012). The Alaska nest (NAM-AK) is a 3-km grid inside of the parent NAM 12-km grid, initialized with the North American Data Assimilation System (NDAS) analysis. The NAM-AK is initialized four times daily at 6-h intervals between 0000 and 1800 UTC and is available on NOAA FTP (ftp://ftp.ncep.noaa.gov/pub/data/nccf/com/nam/prod/). The RDPS is a North American regional version of Environment Canada's global model (Côté et al. 1998) at roughly 10-km resolution initialized with an ensemble–variational data assimilation system four times daily at 6-h intervals between 0000 and 1800 UTC (Caron et al. 2015).

A 6-h QPF grid is collected from each of the three forecasting systems, using the 3- to 9-h forecasts of the HRRR-AK and the 6- to 12-h forecasts of the NAM-AK and RDPS, initialized three and six hours prior to each QPE analysis period, respectively. The forecast period over which QPF is collected is a choice constrained by available initialization times, a desire for a short-range QPF to reduce timing and spatial errors in precipitation, and a desire to avoid problems with model spinup on QPF in the first few hours. The NAM-AK and RDPS QPF are interpolated to the HRRR-AK grid prior to being interpolated to the lower-resolution (12 km) QPE grid within the HRRR-AK model domain. Each of the three model datasets provides complete coverage of Alaska (Fig. 1b). The QPF for a given analysis period is computed as a simple average (or ensemble mean) of the three QPF members. A 12 km grid was chosen for the QPE grid to accommodate the resolution of the RDPS model member, and it was found that reducing resolution of the 3-km HRRR-AK forecast grid to 12 km via Earth System Modeling Framework (ESMF) bilinear interpolation results in no degradation of performance.

The use of an ensemble mean QPF generates precipitation where any of the three QPF members are nonzero. This can result in precipitation features being smeared over larger geographic footprints than exist in any individual QPF member, when members disagree about the existence or precise location of precipitation. This can introduce errors of varying magnitude to QPE depending upon how the ensemble mean QPF is utilized (see section 3e). Strategies may be employed in the future to use remote sensing observations of precipitation or precipitation indicators in order to remove some of these errors (see section 5).

c. QPE configurations and products

Three configurations of the QPE analysis are tested in this study to measure the impact of QPF on the accuracy and uncertainty of the QPE. The three configurations are discussed in detail in section 3e. Two of the three configurations require use of a precipitation climatology in place of the high-resolution QPF to define the QPE first guess. For these configurations, we use a 30-yr climatology derived from the NARR (Mesinger et al. 2006) from the period 1986–2015, available from the Environmental Modeling Center (https://www.emc.ncep.noaa.gov/mmb/rreanl/) and the National Center for Atmospheric Research (https://rda.ucar.edu/datasets/ds608.0/). Precipitation is distributed on a 32-km grid, collected into climatological 6-hourly accumulated precipitation, then interpolated to the QPE grid via ESMF bilinear interpolation.

For this study, QPE is accumulated in 6-h segments four times daily, at 0000–0600, 0600–1200, 1200–1800, and 1800–0000 UTC, and is composed of two types of data: 1) rain gauge observations, and 2) NWP precipitation forecasts. The resultant merged QPE dataset obtained via kriging contains three discrete products: 1) a best-estimate QPE analysis, 2) a lower bound of the QPE analysis representing a 5% confidence limit, and 3) an upper bound of the QPE analysis representing a 95% confidence limit. The intent of these products is to provide both a best estimate of the QPE and a representation of its uncertainty. The QPE analyses are produced on a 12-km-resolution grid centered on Alaska (Fig. 1). The accumulated precipitation at each grid point is represented in millimeters.

An uncertainty estimate for QPE is essential for the Alaska region where in situ precipitation observations are sparse and remote sensing observations are disparate in time and space. While existing QPE products are capable of providing a best guess for the accumulated precipitation, a QPE product that can also quantify the amount of confidence in that best guess can provide forecasters with critical information for assessing the risk of precipitation extremes. The kriging method provides an error variance estimate that is used to derive this uncertainty (see section 3c).

3. Method

a. Kriging

Kriging is a least squares method of obtaining the best linear unbiased estimate (BLUE) of the state at an unobserved point by optimally estimating the state between observed

points by treating observations as realizations of random variables that are spatially correlated (Wackernagel 2010). Two end products are obtained when using kriging: 1) the BLUE representing the best estimate of the state underlying the (often sparse) observations, and 2) an estimate of the variance representing uncertainty in the analysis. The resulting estimation minimizes error variance while maintaining a zero mean of the prediction errors. These are both desirable components of a QPE product, where some measure of uncertainty is necessary, and uncertainty can vary in time and location due to the availability of observations for a particular analysis.

Observations are treated as realizations of random variables with an unknown mean value and a variance that is presumed to be spatially correlated to surrounding grid points. This is achieved by the assumption of second-order stationarity where the variance of the difference between any two points separated by the same distance is presumed to be the same. The second-order stationarity can then be represented by a *semivariogram*, or a one-dimensional model of the semivariance (i.e., one-half of the variance) between observations as a function of the distance between them. The semivariogram model is a function fit to the semivariance data obtained from the observations and is usually selected from a small number of already existing model types (e.g., Gaussian, exponential, and spherical) in part to ensure positive definiteness (Myers 1991). The value of the state at an unobserved point is then modeled as one of these random variables, spatially correlated with observed points based on the semivariogram and the distance to them.

A trend, or gradual variation in space within the data, complicates the semivariogram but can be modeled by allowing the mean of the random variables that are presumed to produce the observations to gradually change over the grid. This requires some information on how the observations covary with some observed quantity that is known universally on the grid. This is referred to as kriging with external drift (Matheron 1969; Hengl et al. 2003). For example, often precipitation correlates with elevation, and elevation (known at each grid point through use of an elevation model) can be used as an external drift variable when kriging gauge observations to a grid (e.g., Goovaerts 2000; Haberlandt 2007). One of the configurations tested in this study uses the ensemble mean QPF as an external drift variable rather than as a QPE first guess (see section 3e).

Kriging has been shown to work well for gridded interpolation of daily meteorological station data for climate applications (Hofstra et al. 2008), interpolation of gauge observations to high-resolution grids for hydrological applications (Tobin et al. 2011), production of long-term gridded precipitation datasets using observations sourced from gauges that are often relocated or provide incomplete records (Libertino et al. 2018), and optimal blending of precipitation observations from gauges and satellite- (Verdin et al. 2016) or radar- (Jewell and Gaussiat 2015) derived precipitation data. When compared with other interpolation techniques, including multiquadratic and inverse distance weighting schemes, kriging is found to provide the best overall error statistics in cross validation, having the most robust and stable scores as observation density is decreased (Jewell and Gaussiat 2015). Kriging also differs from other techniques by producing an interpolation (error) variance product, which contains valuable information on the uncertainty of the interpolation. Many other interpolation techniques, such as inverse distance weighting, provide no measurement of uncertainty (Myers 1994).

b. *QPE kriging product method*

The process for producing a QPE analysis is summarized in Fig. 2. The analysis starts with a first-guess precipitation field (Fig. 2a) and quality-controlled observations (Fig. 2b). Innovations (observation minus first-guess values) are computed at the observation locations (Fig. 2c), and then kriging is used to distribute the innovation information to a gridded increment (Fig. 2d). The increment is then added to the model first guess to produce the final QPE analysis (Fig. 2e), with negative values reassigned to zero. In addition to the gridded increment (i.e., the BLUE of the innovations interpolated to the grid), the kriging variance is also provided (Fig. 2f)—the kriging variance can be interpreted as a measure of the uncertainty in the interpolation of the gridded increment, with higher uncertainty in regions farther from available observations. This uncertainty measurement is used to produce confidence bounds on the QPE analysis (see section 3c). Three configurations of this process are presented in this study, described in section 3e. The three configurations are compared to determine whether the QPF provides any value to the QPE analysis, and if so, whether that value is best utilized by treating the QPF as a first guess or as an external drift variable in kriging.

The method developed during this study to blend NWP model forecast data and gauge observations loosely follows the kriging method used by the National Operational Hydrologic Remote Sensing Center (NOHRSC) National Snowfall Analysis version 2 (NOHRSC_v2; https://www.nohrsc.noaa.gov/snowfall_v2/). The NOHRSC_v2 product blends a first-guess analysis composed of NCEP Stage-IV precipitation (e.g., Lin and Mitchell 2005) data, short-range NWP QPF, and radar-based precipitation estimates with observed snowfall to produce a 24-h accumulated snowfall analysis ending at 0000 and 1200 UTC (Fall et al. 2015). The difference between the snowfall observations and a first-stage analysis product is chosen as the kriged observation, which interpolates corrections to the first guess (or innovations) at observed points to a gridded interpolated correction (or an increment) via kriging (Fall et al. 2020).

c. *QPE confidence bounds*

QPE confidence bounds represent a 5% and 95% confidence interval on precipitation, based on uncertainty from several sources. Interpolation is only one source of uncertainty. In kriging, this uncertainty is represented by the kriging variance, and every interpolated point could be approximated as the center of a normal distribution (presuming BLUE zero-mean prediction error) with a variance equal to the kriging variance (e.g., Lloyd and Atkinson 2001). The 5% and 95% confidence intervals used in this study are computed from this

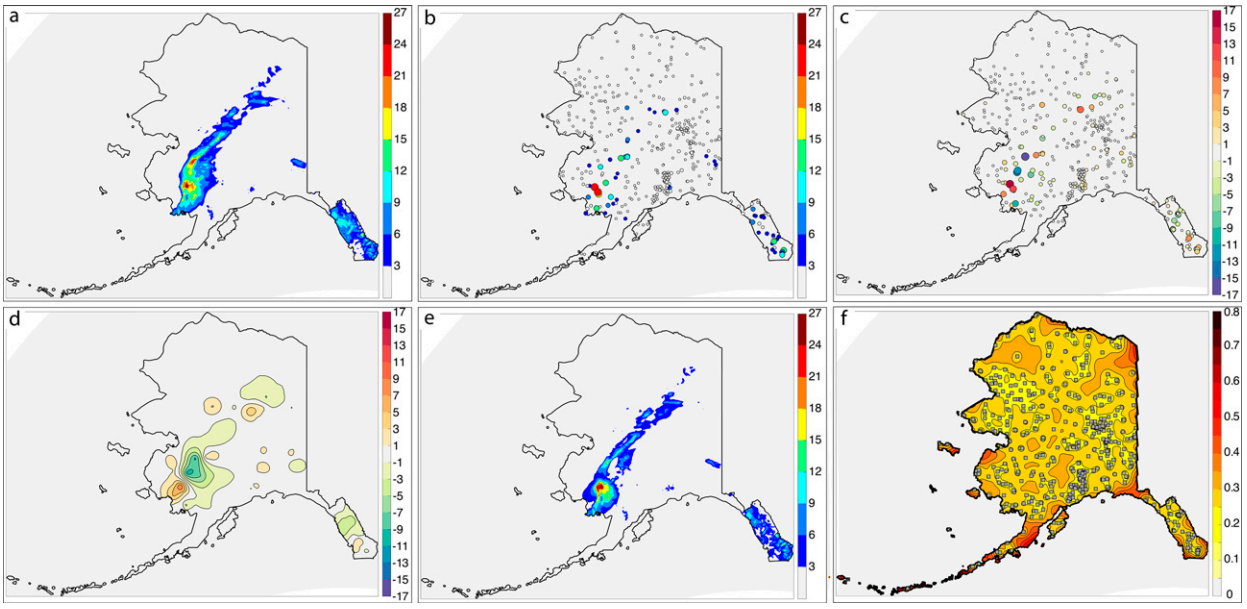

FIG. 2. Example of kriging-based merging of first-guess and gauge data for 0600–1200 UTC 20 Jun 2020. (a) First-guess gridded precipitation. (b) Observed precipitation at gauge sites. (c) Innovations at gauge sites. (d) Gridded increment formed through kriging. (e) Gridded QPE analysis formed from adding increments to the first guess and reassigning negative values to zero. (f) Kriging variance (shaded) with kriged observations plotted as gray squares. The example uses the NWP-BKG configuration (see section 3e), using the high-resolution QPF as a first guess and merging performed via ordinary kriging.

distribution at each analysis time to provide confidence bounds on the gridded increment, that, when added to the first guess, are used to produce 5% and 95% confidence bounds on the QPE analysis. However, there are other sources of uncertainty, such as erroneous, missing, or misplaced precipitation in the first guess that are not accounted for through use of kriging variance alone.

To better represent the uncertainty in the QPE analysis, k-fold cross validation is performed with $k = 40$, whereby 2.5% of the observations are withheld from the analysis on each iteration and then observation–QPE pairs are produced, until cross-validated pairs are computed for each observation. The uncertainty in the QPE analysis is tuned through an inflation-factor α applied to the kriging variance at each cross-validation point such that the probability density function for the QPE increment can be defined by a normal distribution with a mean of the kriged increment and a variance $\alpha\sigma_0^2$, where σ_0^2 is the kriging variance, such that 90% of the cross-validated observations fall between the 5% and 95% bounds of the distribution. The inflation-factor method is similar to methods to inflate ensemble forecast covariance to prevent underdispersion in an ensemble (Anderson and Anderson 1999), or the calibration factor used to correct a radar precipitation field by computing the ratio of radar and gauge observations (Brandes 1975). Note that the inflation factor could be reduced by improving the first guess, increasing the number of observations that are kriged into the analysis, or improving the kriging method itself by including additional external drift variables that correlate strongly with the innovations or using a more effective method to compute the semivariogram. These represent improvements to the kriging method that reduce uncertainty in the analysis.

Treating the kriged, gridded increment as a normal distribution with the presumed (inflated) variance, 5% and 95% confidence bounds are computed along with the best estimate of the increment (Figs. 3b,d,f). These are added to the model first guess (Fig. 3a). The resulting products include the best-guess QPE analysis (Fig. 3e) as well as 5% and 95% confidence bounds of the QPE analysis (Figs. 3c,g).

This method of modeling the estimated QPE error is dependent on an assumption that the errors are Gaussian. Analysis errors in cross validation are found to be distributed with a central peak near zero, but the distribution has much higher kurtosis than a Gaussian and the distribution is more similar to a t distribution with 1 or 2 degrees of freedom (not shown), suggesting that the distribution has fat tails that are more similar to a Cauchy distribution. A Gaussian distribution is approached by a t distribution as the number of degrees of freedom approaches infinity. While the fat-tailed distribution of errors is addressed in the computing of the semivariogram (see section 3e), these deviations from Gaussianity necessarily limit the applicability of the kriging variance to estimate the analysis error with total accuracy.

d. *Software platform*

The software platform for the QPE analysis is open-source and primarily written in Python-3.6.7, other than the NCAR Command Language (NCL) being used to transform NAM-AK and RDPS forecast grids to the HRRR-AK grid using the ESMF regridding function during preprocessing. Kriging is

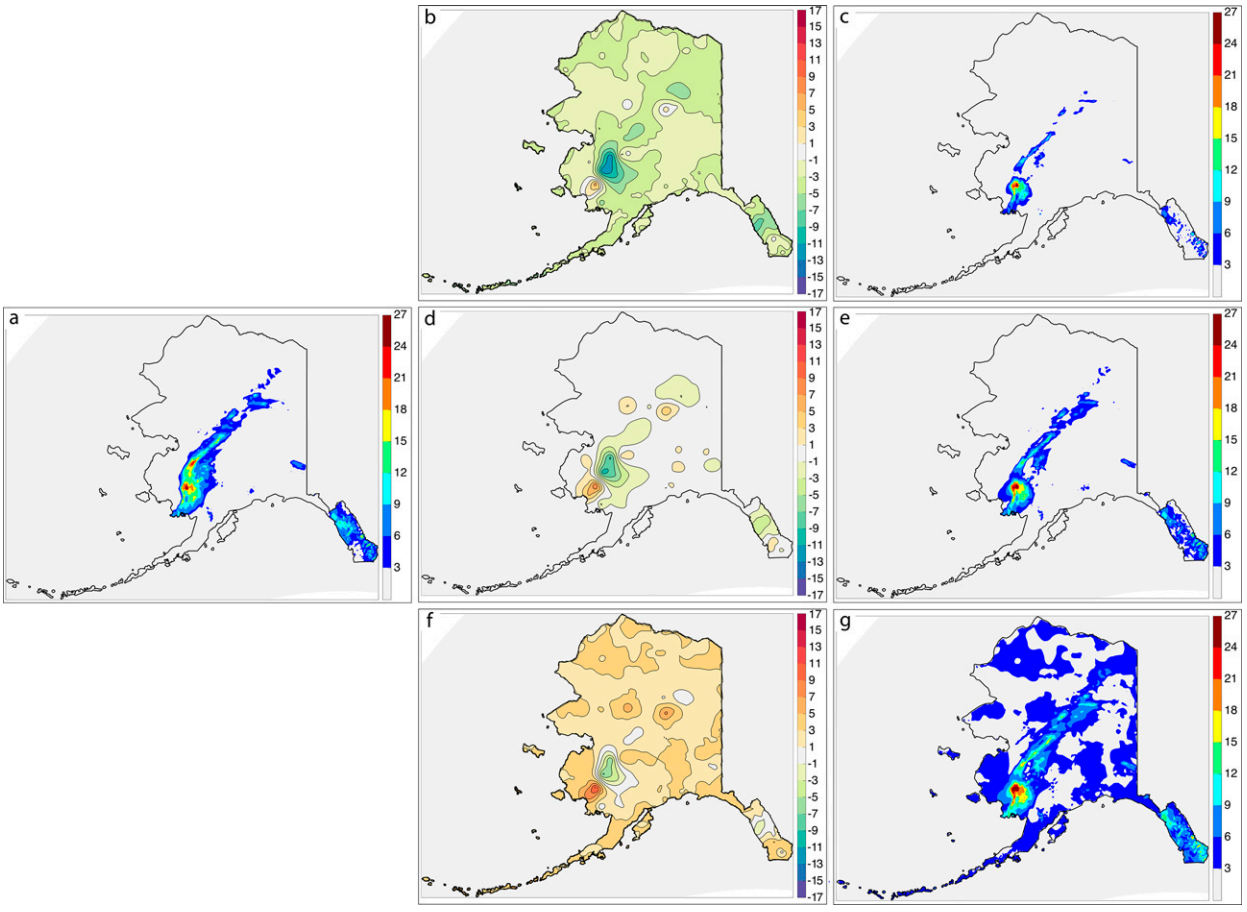

FIG. 3. Example of QPE confidence bounds estimation for the analysis on 0600–1200 UTC 20 Jun 2020. (a) First-guess gridded precipitation, as per Fig. 2a. (b) The 5% confidence bound of the kriged increment. (c) The 5% confidence bound QPE. (d) The best guess of the kriged increment. (e) The best-guess QPE. (f) The 95% confidence bound of the kriged increment. (g) The 95% confidence bound QPE.

accomplished using the GSTAT module in R (Pebesma 2004) and embedded within Python using rpy2. With observation and first-guess data properly preprocessed, the system runs within 2–3 min to produce a single 6-h analysis.

e. *Configurations*

To test the value-added by high-resolution QPF when generating the QPE analyses, three configurations are tested. The most minimal configuration, called NO-NWP, replaces the high-resolution QPF first guess with a climatological precipitation from the 30-yr NARR dataset, with merging of rain gauges accomplished via ordinary kriging. This configuration is the baseline with which other configurations using the high-resolution QPF from the NWP models will be compared. The NWP-BKG configuration uses the high-resolution QPF for the first guess, with merging performed via ordinary kriging. In the third configuration, NWP-KED, the NARR climatology is used as the first guess, but merging is accomplished using kriging with external drift (KED), with the high-resolution QPF serving as the external drift variable. Comparison of NWP-BKG or NWP-KED with NO-NWP will provide a measure of the value of high-resolution QPF to the analysis whereas comparisons between NWP-BKG and NWP-KED will be used to identify which configuration optimally uses information in the high-resolution QPF.

For all configurations, observations are first subjected to a crude quality control in which all observations greater than 75 mm over a 6-h period are automatically rejected. This criterion was selected after analysis of observation and model data revealed that when these very large precipitation amounts are reported, they are often not corroborated with precipitation by nearby gauges or QPF values. It is possible, for example, for a gauge to drastically underreport liquid equivalent of frozen precipitation as the gauge orifice becomes blocked, and then drastically overreport as accumulated precipitation around the orifice is collected all at once (e.g., Martinaitis et al. 2015), making quality control of precipitation of these amounts difficult. However, it is noted that these are not necessarily flawed observations, because climatological precipitation frequency estimates (see NOAA Atlas 14 point precipitation frequency estimates: https://hdsc.nws.noaa.gov/hdsc/pfds/pfds_map_ak.html) meeting this threshold can be found in the Alaskan panhandle at recurrence intervals of 5 years or less. However, the difficulty in validating these observations and the

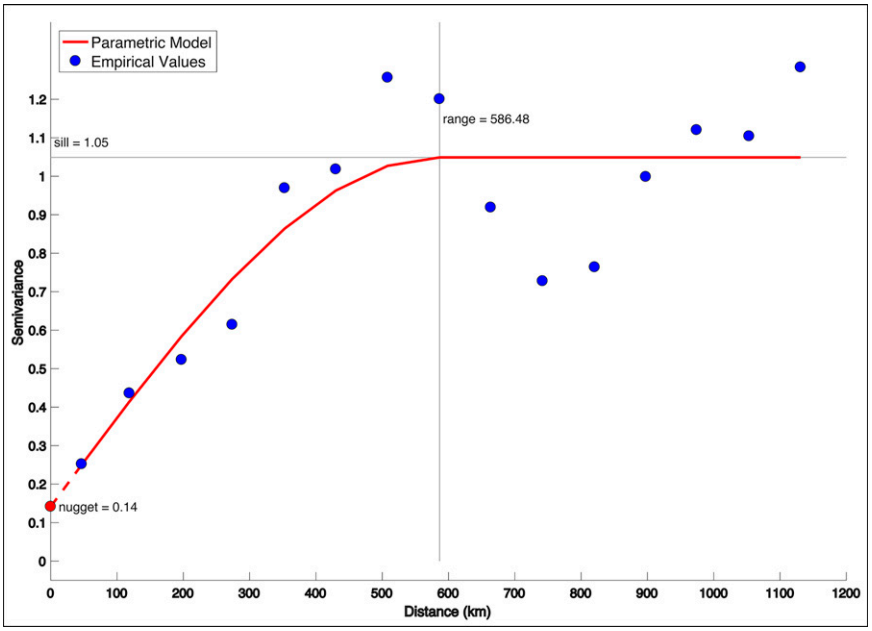

FIG. 4. Example semivariogram for 0600–1200 UTC 20 Jun 2020. Empirical variogram points within discrete distance bins are shown in blue, and the solid red line is the spherical parametric model variogram. The nugget variance is represented by the red point at the *y* intercept. The range and sill are plotted in gray. The semivariogram is computed using R/GSTAT.

extreme innovations often produced by these observations motivated their exclusion in this study.

The remaining observations are then compared with the other observations within a 200-km radius, with any observation that is more than 10 mm different from all observations in its neighborhood rejected. This method of removing significant outliers was adapted from a quality control mechanism by NOHRSC in its kriging-based snowfall maps as an "oddball" test to establish minimal neighborhood corroboration as a part of quality control, though it is noted that this method may remove authentic gauge reports in convective events with significant precipitation gradients and sparse observations. While APRFC gauge observations undergo strict quality control, this second round of quality control is intended mainly to address remaining issues with observations from MesoWest and to remove extreme outliers prior to kriging. Innovation values (observed minus first guess, in observation space) are computed for each of the remaining observations. For any pair of observations within 3 km of each other, one of them is withheld from the analysis to thin observations at distances that are likely well below the smallest distance bin of the variogram.

The remaining observations are blended with the first guess by producing a gridded increment field (observed minus first guess, in grid space), from the innovations through kriging. The first step is to fit a parameterized spherical semivariogram model to the binned semivariance data (the *empirical*, or *sample semivariogram*) to provide a smooth, one-dimensional model of the semivariance between any two grid points as a function of distance (e.g., Fig. 4). The semivariogram is then used to weight the contributions of nearby innovations to infer the value of the increment at an unobserved point on the grid. Observation pairs are collected and sorted into distance bins between 50- and 1130-km separations, and semivariance values are computed for each bin. The number of distance bins and the chosen separation distances are selected by the kriging software (see section 3d) to optimize estimation of the semivariance.

The semivariance is typically zero or near-zero at a separation distance of zero; a positive semivariance at zero-distance is called a *nugget variance* and is related to several sources of error including observation measurement error and uncontrolled fitting of the variogram at distances shorter than the shortest distance between two observations. In this example case, a nugget variance is observed at the *y*-intercept value of the fitted model variogram. The semivariance is typically observed to increase as a function of distance until it levels off at a fixed value, the *sill*, which represents the semivariance of the presumed random process generating the observed data. The semivariance reaches the sill at a distance called the *range*, which represents the limiting distance over which correlation between observed values exists. The sample semivariogram is computed using Cressie's robust variogram estimate (Cressie and Hawkins 1980) to better account for the possibility of an innovation distribution with heavier tails than would be expected for a normal distribution of innovations being kriged to the analysis grid.

The semivariogram's range has significant variability from one analysis to the next. Across the entire period of the experiment, the range has a 90% confidence range of between 40 and 875 km, with the median value of between 260 and 350 km depending on the configuration. Semivariograms with a large range are expressing differences between the first

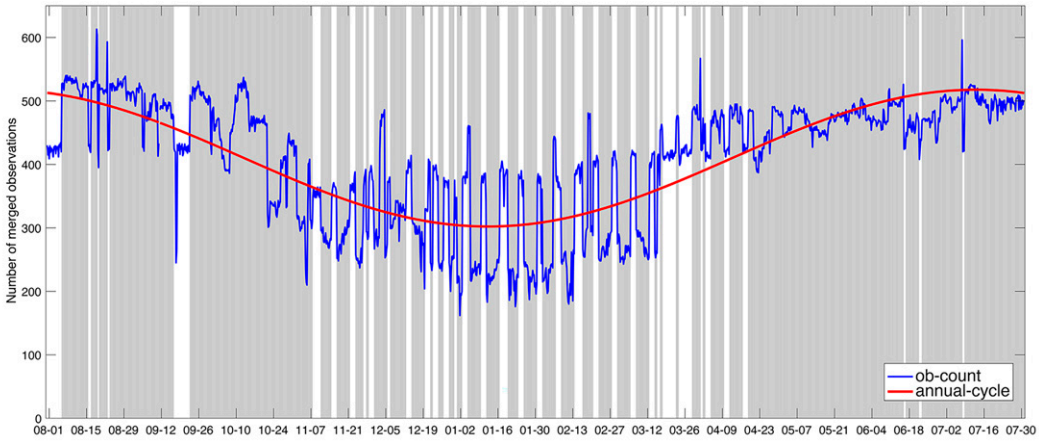

FIG. 5. Observation count for the number of kriged observations per analysis (blue) and annual cycle of observation count (red) in the period 1 Aug 2019–31 Jul 2020. Gray shading represents the analyses for which quality-controlled APRFC gauge observations were available. The annual cycle is computed via harmonic regression.

guess and observations that are highly spatially correlated. This may be caused by significant spatial displacement of large-scale precipitation patterns in the first guess relative to observation. In NWP-BKG, this may be caused by precipitation features in the ensemble mean QPF being significantly displaced from their real location. This source of error could be contributed to significantly by a minority of ensemble members with significant displacement errors. Techniques to improve the first guess may yield reductions in the semivariogram range by removing these kinds of highly spatially correlated errors (see section 5).

4. Results

QPE analyses were produced continuously at 6-h intervals for the period 1 August 2019–31 July 2020. To maximize its uptime, the analysis system is designed to run even if only one of the three high-resolution QPF members is available and regardless of whether the quality-controlled APRFC precipitations observations are available. This approach meant that there was only one missing analysis during the 12-month period. There is a seasonal cycle in the number of kriged observations (Fig. 5), with a minimum in observations appearing in mid-January and a maximum in early July. Inclusion of APRFC observations generally increases (decreases) the number of kriged observations in the warm (cool) season. The number of observations is lower overall during the cool season due to difficulties measuring snowfall.

a. Cross-validation error scores

Errors are computed for cross-validated observation–QPE pairs, to evaluate the accuracy of the analysis where observations are withheld. Chosen scores include the root-mean-square error (RMSE), the mean absolute error (MAE), and the bias error (hereinafter bias)—computed as the mean error. Since the distribution of precipitation observations contains mostly zero or near-zero precipitation amounts (e.g., 82% of observations report at or below 0.1 mm precipitation), QPE accuracy as a function of observed precipitation amount is also examined to focus on QPE performance in high-precipitation cases.

Error scores computed across all available observations reveal that NWP-KED contains the lowest RMSE and MAE, while also having a competitive bias (Fig. 6a). NWP-BKG has an RMSE almost identical to NO-NWP, but with a lower MAE. Because RMSE is computed from squared errors, it is more sensitive to outliers than MAE. It can therefore be inferred from these scores that using the high-resolution QPF as a first guess reduces the overall error relative to using a climatological first guess but that the high-resolution QPF also introduces more outlier errors. Bias is lowest for NO-NWP, implying that the use of high-resolution QPF increases the wet bias of the analysis, though this effect is partially mitigated in NWP-KED by using the high-resolution QPF as an external drift variable instead of a first guess. As an additional benchmark, the equivalent scores for the ensemble mean QPF that defines the first guess in NWP-BKG are an RMSE of 1.71 mm, MAE of 0.61 mm, and bias of 0.30 mm.

From these scores alone, it can be inferred that the high-resolution QPF provides value to the QPE analysis, with the optimal configuration utilizing QPF as an external drift variable to blend observations and a climatological first guess via KED. NWP-KED also expresses the least uncertainty in the analysis, defined by the mean difference between the empirically defined 5% and 95% confidence bounds for QPE in cross validation. The mean 5%–95% confidence interval for NWP-KED across all observations is 1.56 mm, while NO-NWP has an interval of 1.61 mm and NWP-BKG has the largest interval at 1.68 mm. So, in addition to providing the smallest RMSE and MAE, NWP-KED also produces a QPE analysis with the least uncertainty.

Since most observations are of dry or nearly dry environments, the cross-validation error scores are focused mostly on

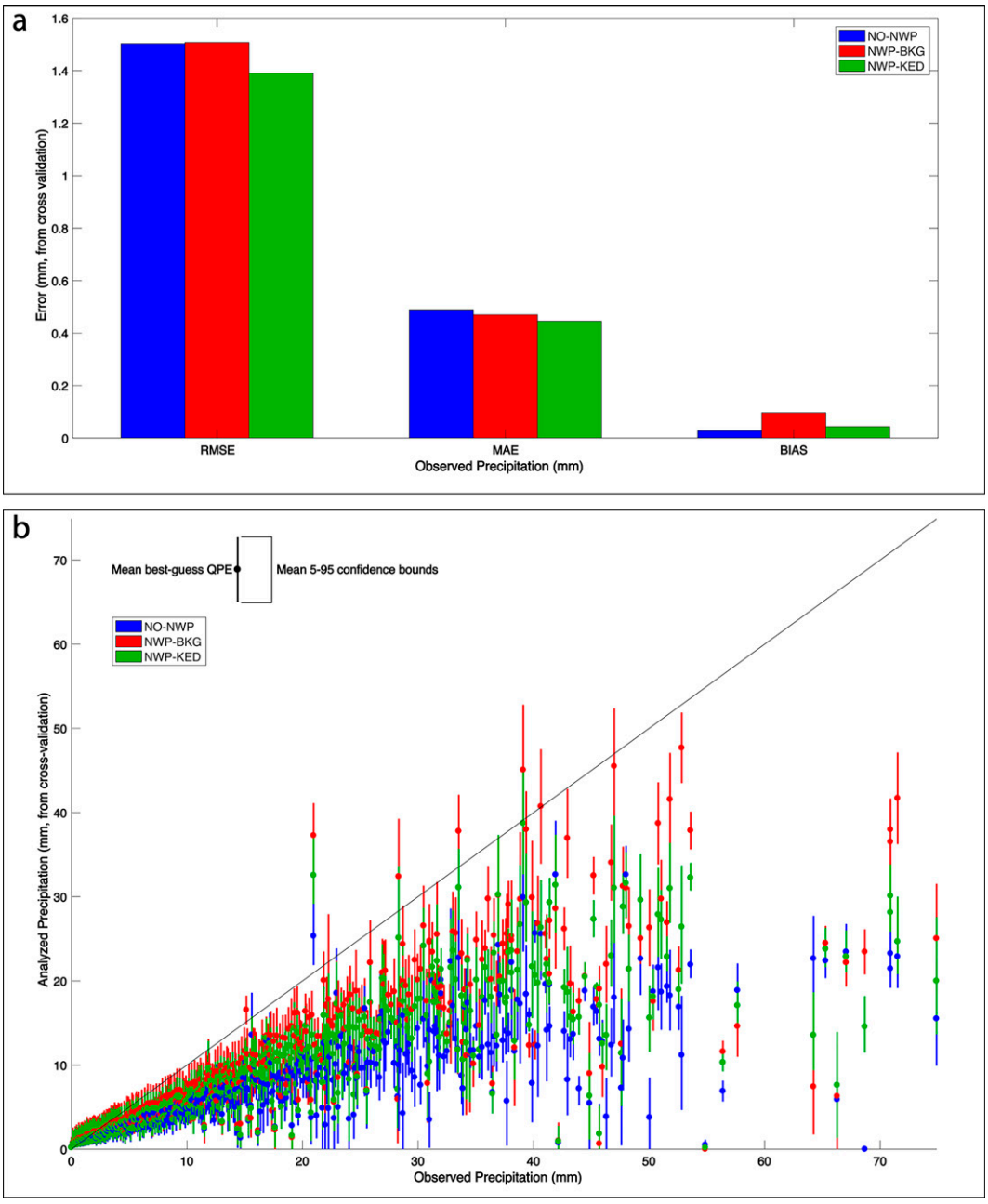

FIG. 6. Cross-validation error statistics for NO-NWP (blue), NWP-BKG (red), and NWP-KED (green) configurations, computed across all available observations. (a) RMSE, MAE, and bias for each configuration. (b) Mean QPE and mean 5%–95% QPE confidence interval for 750 precipitation bins equally spaced between 0 and 75 mm. Perfect QPE for each bin is represented by the black line.

estimating little to no precipitation. To evaluate the accuracy of the QPE analysis in high-precipitation environments, observations are placed in 750 bins at 0.1-mm intervals from 0 to 75 mm. (Fig. 6b). In each bin, a mean best-guess QPE and a mean 5% and 95% confidence interval are computed. Focusing on the minority of observations > 0.1 mm, the QPE analysis from NWP-BKG appears to capture high-precipitation environments better than NWP-KED, yet both configurations are superior to NO-NWP. All three analyses struggle to represent high-precipitation environments in cross validation, almost universally underpredicting precipitation, but NWP-BKG is closest to the observed precipitation across the majority of the spectrum.

b. *Semivariogram differences*

Differences arise in the structure of the semivariogram for each configuration because NWP-BKG uses the high-

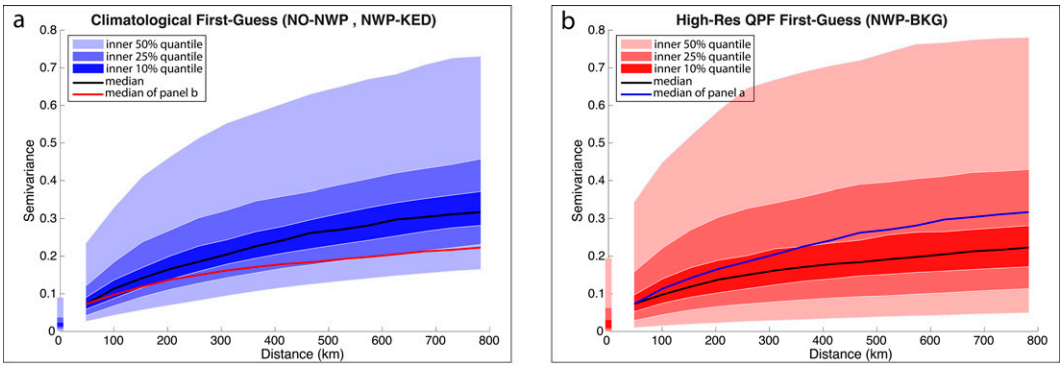

FIG. 7. (a) Median parametric semivariogram (black) and shading representing inner 10%, 25%, and 50% quantiles, for all semivariograms computed from a climatological first guess. (b) As in (a), but for all semivariograms computed from a first guess formed from high-resolution QPF. Quantile shading of nugget variance is provided at $x = 0$ km. The red line in (a) and the blue line in (b) represent the median semivariogram from the other panel.

resolution QPF as a first guess and both NO-NWP and NWP-KED use a climatological first guess. Since innovations are computed as the departure of observations from the first guess, the complexion of the innovations, including their spatial autocorrelation, are different between these two first-guess configurations and by extension there are differences in the semivariogram. Differences in the semivariogram determine how the increment values at discrete, observed points are interpolated to the grid by controlling the similarity between the increment at a grid point and the innovation at some observed point, the amount of variation that exists between the increment and the innovation at zero range, and the smoothness of the increment field. Kriging is a smoother, reducing variance by overestimating small values and underestimating large values (Fig. 6b). The smoothness of the kriged field is dictated by the semivariogram's representation of the spatial autocorrelation estimated from observation pairs via the nugget to sill ratio, with more smoothing as the ratio increases (Oliver and Webster 2014).

Collectively across all computed semivariograms, those from the NWP-BKG configuration display a lower median semivariance at all distances than those from NO-NWP or NWP-KED, but the range of semivariance values is larger at all distances for NWP-BKG as defined by the inner 10%, 25%, and 50% quantile ranges (Fig. 7). The nugget semivariance has twice the inner 50% quantile range for NWP-BKG that it does for NO-NWP or NWP-KED, which skews the distribution of nugget semivariance to higher values in NWP-BKG. For example, 24.7% of the nugget semivariances for NWP-BKG are at values of at least 0.2, whereas for the other configurations it is only 16.1%. The higher nugget semivariances in NWP-BKG allow for larger uncertainties in QPE even in regions of the grid where observations are very close by.

c. 12-month average QPE and uncertainty differences

The 12-month average QPE is defined from the best-guess QPE, and 12-month average uncertainty is defined from the 5% and 95% confidence interval (Fig. 8). Maximum average precipitation (Figs. 8b,d,f) appears along the Gulf of Alaska coast, with average QPE decreasing rapidly with latitude. Regions of high average uncertainty (Figs. 8a,c,e) tend to be collocated with regions of high average precipitation, although there are areas of the North Slope where uncertainty is enhanced by the scarcity of observations.

The 12-month-average QPE and uncertainty are smoothest in the NWP-BKG configuration (Figs. 8a,b) and least smooth in the NWP-BKG configuration (Figs. 8c,d). QPE and uncertainty in NWP-BKG displays local maxima and high variability associated with the Aleutian Range in southwestern Alaska, the Alaska Range through southern Alaska, and the Brooks Range in northern Alaska, with these features smoothed in NO-NWP. NWP-KED is a midpoint between these two configurations (Figs. 8e,f). Average uncertainty in NWP-KED is lower than in NO-NWP or NWP-BKG for the majority of Alaska, with higher average uncertainty appearing along the Aleutian Peninsula and along the Gulf of Alaska coast (Figs. 9a,b). These regions appear to be related to terrain features, with uncertainty in NWP-KED exceeding NO-NWP at high elevations and exceeding NWP-BKG at low elevations (Fig. 9c).

5. Conclusions

A method for producing a QPE analysis from blending first-guess data and rain gauge observations has been presented, utilizing kriging to interpolate innovations to a gridded analysis increment. The method produces both a BLUE of QPE as well as an estimate of the (error) variance at each grid point, which is used to provide 5% and 95% confidence bounds on the analysis. This method is applied to QPE in the Alaska region where gauge observations are sparse and remote sensing options for precipitation estimation are limited by geography, latitude, and resources. A QPE capable of providing an estimate of its uncertainty is critical in this region to properly evaluate risks posed by precipitation on both the high and low ends of the spectrum.

Three configurations of the QPE analysis were tested to evaluate the benefit provided by QPF from high-resolution

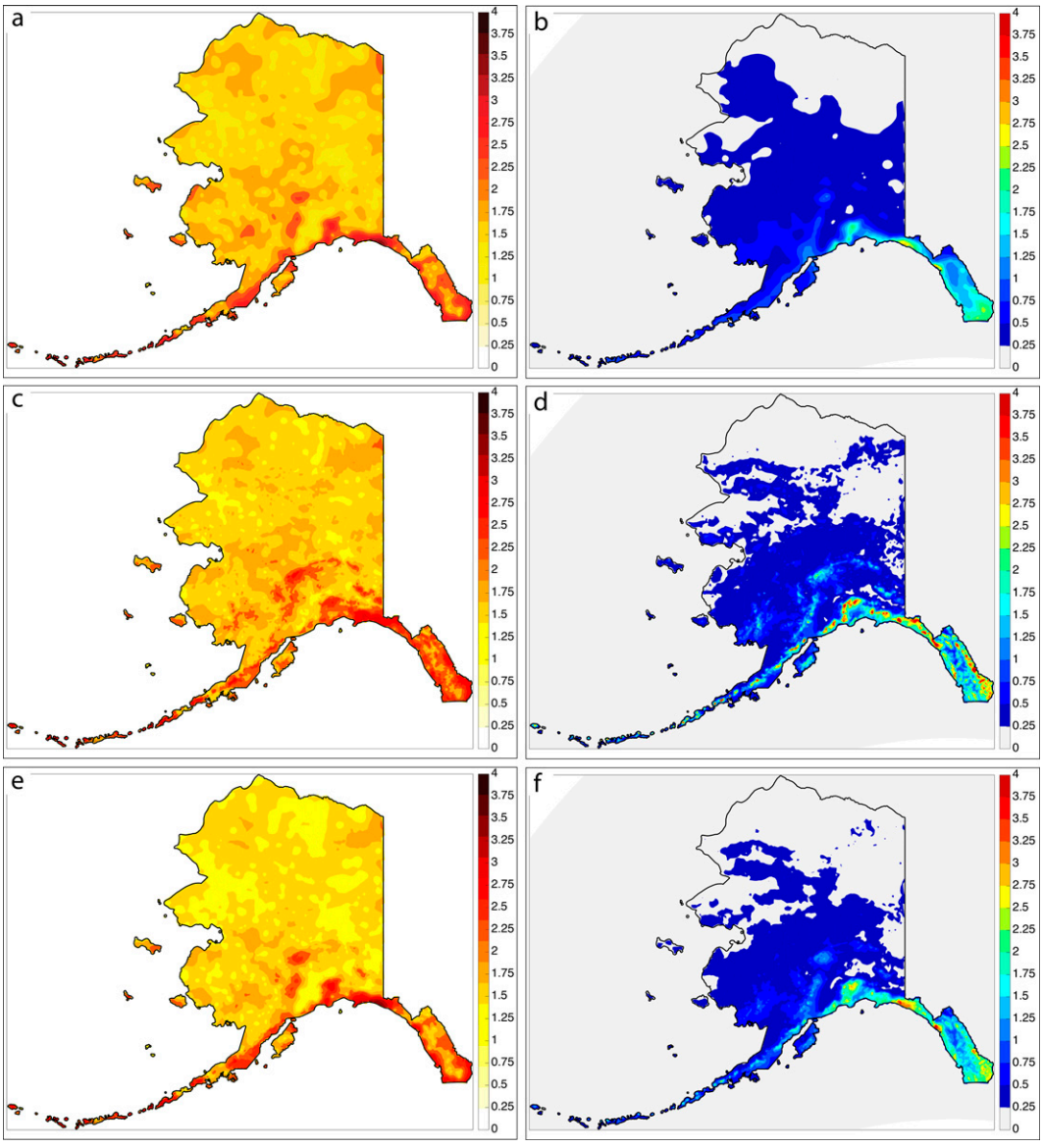

FIG. 8. Twelve-month (left) average QPE uncertainty of the 6-h precipitation analyses and (right) best-guess QPE for the (a),(b) NO-NWP; (c),(d) NWP-BKG; and (e),(f) NWP-KED configurations. Uncertainty is defined by the 5% and 95% confidence interval of the analysis.

NWP and how to best utilize the QPF to reduce analysis errors and uncertainty. In NO-NWP, the QPF is not utilized and instead the analysis is formed by merging gauge observations with a climatological first guess via ordinary kriging. In NWP-BKG, the climatological first guess is replaced with the QPF formed from an ensemble average from three forecast models (HRRR-AK, NAM-AK, and RDPS). In NWP-KED, the first guess is formed from climatological precipitation but QPF is used to merge with gauge observations via kriging with external drift, with the QPF serving as an external drift variable.

Cross-validated error statistics showed that NWP-KED provided the QPE with the lowest RMSE, lowest MAE, and a small bias, although NO-NWP's bias was slightly lower. NWP-KED also produced the analysis with the least uncertainty. However, the cross-validation statistics are collected for a pool of observations that heavily sample dry or nearly dry environments, with 82% of available observations reporting less than 0.1 mm of precipitation. When discretizing cross-validation data by precipitation thresholds, NWP-BKG provides a QPE that more closely aligns with the entire spectrum of sampled precipitation values, although all three configurations tend to underpredict larger precipitation events.

The accuracy of the analysis can be improved through several sources. First, the number of gauge observations could be increased. In addition to APRFC and MesoWest observations

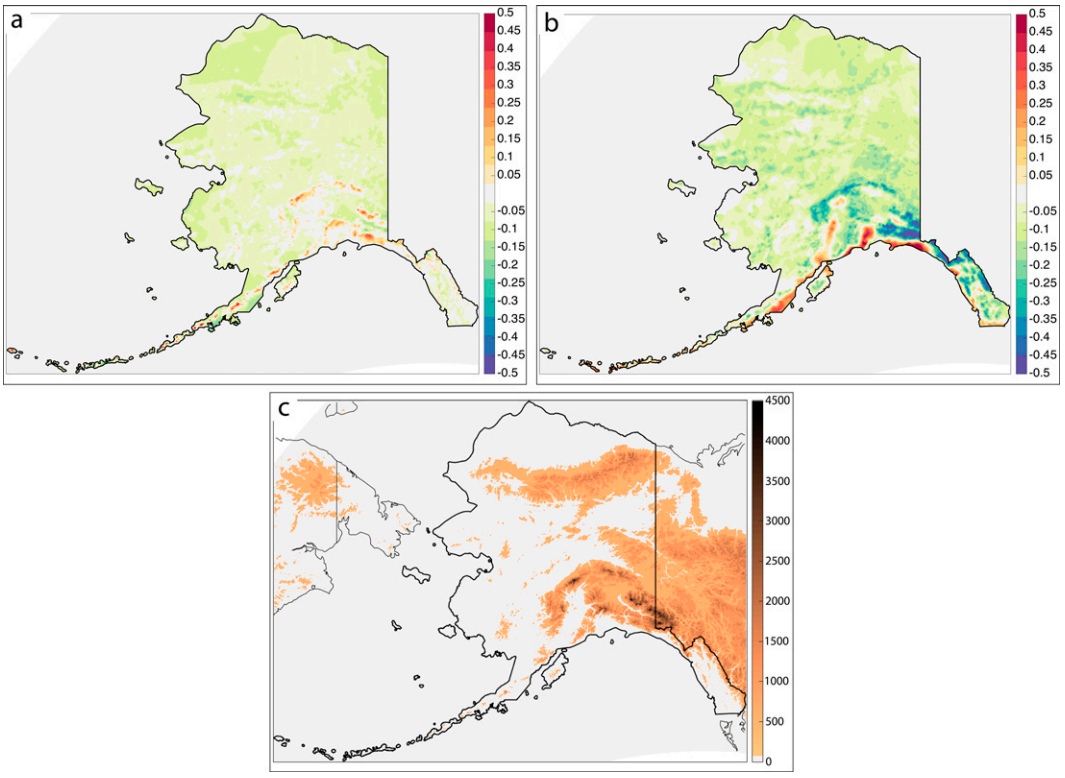

FIG. 9. Difference in 12-month average QPE uncertainty between NWP-KED and (a) NO-NWP or (b) NWP-BKG. (c) Terrain elevation of domain, as defined by the HRRR-AK.

currently being utilized, observations from cooperative observer networks (COOP and CoCoRaHS) are potential candidates to increase observation density. Additional observations would improve the estimation of the semivariogram and reduce kriging variance, both of which can contribute to a reduction in the uncertainty of the analysis.

Second, an improved first guess would reduce the amplitude of innovations and reduce the amount of variance inflation necessary to properly define confidence bounds. Inclusion of additional ensemble members, including time-lagged members from the HRRR-AK, NAM-AK, and RDPS as well as members from global models, could increase the accuracy of the first-guess precipitation field in NWP-BKG or the external drift variable in NWP-KED. Alternatively, the ensemble mean could be replaced with a search for an optimal QPF among available members through examination of contingency table scores [e.g., equitable threat score, receiver operating characteristic (ROC) score]. The selection of a single QPF with the best match to available observations would remove the influence of inaccurate QPF members and retain the spatial distribution of precipitation features in QPF, possibly yielding a first guess with less spatially correlated error affecting the composition of kriging-based merging. Another option is using available remote sensing observations to identify precipitating/nonprecipitating regions and masking the first guess accordingly. False, missing, or mislocated precipitation in regions of the model first guess where there are no gauge observations is a significant source of uncertainty in the analysis, and precipitation clearing based on these data could provide a benefit. This could potentially be achieved through producing binary precipitating/nonprecipitating data from these observations and utilizing indicator kriging (e.g., Solow 1986) to assess the likelihood of precipitation on the grid, masking precipitation that falls below a threshold value. A third option is applying a bias correction procedure to the QPF members. A nonlinear conditional bias correction method such as that developed by Otkin et al. (2018) could be used to handle differences in the bias characteristics between low- and high-precipitation events.

The QPE analysis could be improved through utilizing more covariate information, primarily through defining additional external drift variables that are known universally on the analysis grid and covary with innovations. This could include model variables such as 2-m temperature or moisture, as well as the variance among ensemble members of these fields. Another option may be to use QPE analyses from prior analysis periods as a covariate, which utilizes covarying *temporal* as well as spatial data (e.g., Sideris et al. 2014) and is similar to regression-based model output statistics (MOS; Glahn and Lowry 1972) utilizing prior observations as a predictor (e.g., Lazić et al. 2014). Data that are not universally known on the analysis grid but much more sampled than precipitation observations could be similarly leveraged via cokriging (e.g., Myers 1982), in which case precipitation-rate data swaths from

IMERG or Arctic composite satellite data (Lazzara et al. 2011) could be explored as covariates. Radar data may be useful in any of these roles—as a direct precipitation observation, perhaps after bias correction against nearby rain gauges; as a precipitation indicator for clearing false precipitation from the first guess; or potentially as a cokriging auxiliary variable.

We hope in the future to be able to apply this method to initialization of hydrological forecasting models in the Alaska region, which are often restricted to using the very limited gauge network for initial data and are subject to unknown uncertainties. Application of the QPE analysis and uncertainty estimates within hydrological models would provide another means to indirectly assess their accuracy, relative to river streamflow gauges, and potentially demonstrate value added to the hydrological community through more accurate river forecasts. This could be especially useful for smaller rivers and streams that lack in situ precipitation observations.

Acknowledgments. We thank NWS Anchorage, the NWS Alaska region headquarters, and APRFC for their support and guidance during this project. We also thank Dr. Donald Myers from The University of Arizona for his insight and guidance with application of kriging. This project was funded by NWS Anchorage and the Alaska Region Headquarters via NOAA Cooperative Agreements NA15NES4320001 and NA20NES4320003.

Data availability statement. All datasets used in this study are publicly available. Rain gauge data are available via Synoptic (https://developers.synopticdata.com/mesonet/) and the Iowa Environmental Mesonet page (https://mesonet.agron.iastate.edu/). Model forecast data are available from NOAA FTP (ftp://ftp.ncep.noaa.gov/pub/data/nccf/com/) and the Meteorological Service of Canada (MSC) HTTPS Open Data Server (https://dd.weather.gc.ca/). The developed software is Python based and uses the GSTAT statistical package for R, which is available through the Comprehensive R Archive Network (CRAN; https://cran.r-project.org/).

REFERENCES

Anderson, J. L., and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758, https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2.

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.

Brandes, E. A., 1975: Optimizing rainfall estimates with the aid of radar. *J. Appl. Meteor.*, **14**, 1339–1345, https://doi.org/10.1175/1520-0450(1975)014<1339:OREWTA>2.0.CO;2.

Caron, J., T. Milewski, M. Buehner, L. Fillion, M. Reszka, S. Macpherson, and J. St-James, 2015: Implementation of deterministic weather forecasting systems based on ensemble–variational data assimilation at Environment Canada. Part II: The regional system. *Mon. Wea. Rev.*, **143**, 2560–2580, https://doi.org/10.1175/MWR-D-14-00353.1.

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998: The operational CMC–MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. *Mon. Wea. Rev.*, **126**, 1373–1395, https://doi.org/10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2.

Cressie, N., and D. M. Hawkins, 1980: Robust estimation of the variogram: I. *J. Int. Assoc. Math. Geol.*, **12**, 115–125, https://doi.org/10.1007/BF01035243.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical–topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.*, **33**, 140–158, https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2.

Ehret, U., J. Götzinger, A. Bárdossy, and G. G. S. Pegram, 2008: Radar-based flood forecasting in small catchments, exemplified by the Goldersbach catchment, Germany. *Int. J. River Basin Manage.*, **6**, 323–329, https://doi.org/10.1080/15715124.2008.9635359.

Fall, G., D. Soroka, D. Novak, J. Waldstreicher, T. Alcott, and B. Walawender, 2015: Proposal for national snowfall analysis, version 2. NOAA Doc., 6 pp., https://www.nohrsc.noaa.gov/technology/pdf/Proposal_for_National_Snowfall_Analysis_Phase_II.pdf.

——, M. N. Chanhassen, and K. H. Sparrow, 2020: NOAA's National Snowfall Analysis: Technical description and evaluation. *100th AMS Annual Meeting*, Boston, MA, Amer. Meteor. Soc., 368914, https://ams.confex.com/ams/2020Annual/webprogram/Paper368914.html.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211, https://doi.org/10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2.

Goovaerts, P., 2000: Geostatistical approaches for incorporating elevation in the spatial interpolation of rainfall. *J. Hydrol.*, **228**, 113–129, https://doi.org/10.1016/S0022-1694(00)00144-X.

Haberlandt, U., 2007: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *J. Hydrol.*, **332**, 144–157, https://doi.org/10.1016/j.jhydrol.2006.06.028.

Hengl, T., G. B. M. Geuvelink, and A. Stein, 2003: Comparison of kriging with external drift and regression-kriging. ITC Tech. Note, 17 pp.

Hofstra, N., M. Haylock, M. New, P. Jones, and C. Frei, 2008: Comparison of six methods for the interpolation of daily, European climate data. *J. Geophys. Res.*, **113**, D21110, https://doi.org/10.1029/2008JD010100.

Horel, J., and Coauthors, 2002: MesoWest: Cooperative mesonets in the western United States. *Bull. Amer. Meteor. Soc.*, **83**, 211–225, https://doi.org/10.1175/1520-0477(2002)083<0211:MCMITW>2.3.CO;2.

Huffman, G. J., E. F. Stocker, D. T. Bolvin, E. J. Nelkin, and J. Tan, 2019: GPM IMERG final precipitation L3 half hourly 0.1 degree × 0.1 degree V06. GES DISC, accessed 2 June 2018, https://doi.org/10.5067/GPM/IMERG/3B-HH/06.

Janjic, Z., and R. L. Gall, 2012: Scientific documentation of the NCEP nonhydrostatic multiscale model on the B grid (NMMB). Part 1 dynamics. NCAR Tech. Note NCAR/TN-489+STR, 80 pp., https://doi.org/10.5065/D6WH2MZX.

Jewell, S. A., and N. Gaussiat, 2015: An assessment of kriging-based rain-gauge–radar merging techniques. *Quart. J. Roy. Meteor. Soc.*, **141**, 2300–2313, https://doi.org/10.1002/qj.2522.

Lazić, L., G. Pejanovic, M. Zivkovic, and L. Ilic, 2014: Improved wind forecasts for wind power generation using the Eta model and MOS (model output statistics) method. *Energy*, **73**, 567–574, https://doi.org/10.1016/j.energy.2014.06.056.

Lazzara, M. A., A. Coletti, and B. L. Diedrich, 2011: The possibilities of polar meteorology, environmental remote sensing, communications and space weather applications from artificial Lagrange orbit. *Adv. Space Res.*, **48**, 1880–1889, https://doi.org/10.1016/j.asr.2011.04.026.

Libertino, A., P. Allamano, F. Laio, and P. Claps, 2018: Regional-scale analysis of extreme precipitation from short and fragmented records. *Adv. Water Res.*, **112**, 147–159, https://doi.org/10.1016/j.advwatres.2017.12.015.

Lin, Y., and K. E. Mitchell, 2005: The NCEP Stage II/IV hourly precipitation analyses: Development and applications. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 1.2, https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.

Lloyd, C. D., and P. M. Atkinson, 2001: Assessing uncertainty in estimates with ordinary and indicator kriging. *Comput. Geosci.*, **27**, 929–937, https://doi.org/10.1016/S0098-3004(00)00132-1.

Martinaitis, S. M., S. B. Cocks, Y. Qi, B. T. Kaney, J. Zhang, and K. Howard, 2015: Understanding winter precipitation impacts on automated gauge observations within a real-time system. *J. Hydrometeor.*, **16**, 2345–2363, https://doi.org/10.1175/JHM-D-15-0020.1.

——, and Coauthors, 2020: A physically based multisensory quantitative precipitation estimation approach for gap-filling radar coverage. *J. Hydrometeor.*, **21**, 1485–1511, https://doi.org/10.1175/JHM-D-19-0264.1.

Matheron, G., 1969: *Le krigeage universel.* Vol. 1. *Cahiers du Centre de Morphologie Mathematique*, Ecole des Mines de Paris, 83 pp.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, https://doi.org/10.1175/BAMS-87-3-343.

Michelson, D. B., and Coauthors, 2000: BALTEX Radar Data Centre products and their methodologies. SMHI Meteorology and Climatology Rep. RMK 90, 76 pp.

Myers, D. E., 1982: Matrix formulation of co-kriging. *Math. Geol.*, **14**, 249–257, https://doi.org/10.1007/BF01032887.

——, 1991: Interpolation and estimation with spatially located data. *Chemom. Intell. Lab. Syst.*, **11**, 209–228, https://doi.org/10.1016/0169-7439(91)85001-6.

——, 1994: Spatial interpolation: An overview. *Geoderma*, **62**, 17–28, https://doi.org/10.1016/0016-7061(94)90025-6.

Oliver, M. A., and R. Webster, 2014: A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *Catena*, **113**, 56–69, https://doi.org/10.1016/j.catena.2013.09.006.

Otkin, J. A., R. Potthast, and A. Lawless, 2018: Nonlinear bias correction for satellite data assimilation using Taylor series polynomials. *Mon. Wea. Rev.*, **146**, 263–285, https://doi.org/10.1175/MWR-D-17-0171.1.

Pebesma, E. J., 2004: Multivariable geostatistics in S: The gstat package. *Comput. Geosci.*, **30**, 683–691, https://doi.org/10.1016/j.cageo.2004.03.012.

Rasmussen, R., and Coauthors, 2012: How well are we measuring snow? The NOAA/FAA/NCAR winter precipitation test bed. *Bull. Amer. Meteor. Soc.*, **93**, 811–829, https://doi.org/10.1175/BAMS-D-11-00052.1.

Schaake, J., A. Henkel, and S. Cong, 2004: Application of PRISM climatologies for hydrologic modeling and forecasting in the western U.S. *18th Conf. on Hydrology*, Seattle, WA, Amer. Meteor. Soc., 5.3, https://ams.confex.com/ams/84Annual/techprogram/paper_72159.htm.

Sideris, I. V., M. Gabella, R. Erdin, and U. Germann, 2014: Real-time radar–rain-gauge merging using spatio-temporal co-kriging with external drift in the Alpine terrain of Switzerland. *Quart. J. Roy. Meteor. Soc.*, **140**, 1097–1111, https://doi.org/10.1002/qj.2188.

Sinclair, S., and G. Pegram, 2005: Combining radar and rain gauge rainfall estimates using conditional merging. *Atmos. Sci. Lett.*, **6**, 19–22, https://doi.org/10.1002/asl.85.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Solow, A. R., 1986: Mapping by simple indicator kriging. *Math. Geol.*, **18**, 335–352, https://doi.org/10.1007/BF00898037.

Tobin, C., L. Nicotina, M. B. Parlange, A. Berne, and A. Rinaldo, 2011: Improved interpolation of meteorological forcings for hydrologic applications in a Swiss Alpine region. *J. Hydrol.*, **401**, 77–89, https://doi.org/10.1016/j.jhydrol.2011.02.010.

Velasco-Forero, C. A., A. Seed, D. Sempere-Torres, and G. Pegram, 2008: Optimal estimation of rainfall fields merging radar and rain gauges data in an operational context. *Proc. Fifth European Conf. on Radar in Meteorology and Hydrology*, Helsinki, Finland, Vaisala.

Verdin, A., C. Funk, B. Rajagopalan, and W. Kleiber, 2016: Kriging and local polynomial methods for blending satellite-derived and gauge precipitation estimates to support hydrologic early warning systems. *IEEE Trans. Geosci. Remote Sens.*, **54**, 2552–2562, https://doi.org/10.1109/TGRS.2015.2502956.

Wackernagel, H., 2010: *Multivariate Geostatistics: An Introduction with Applications.* 3rd ed. Springer, 388 pp.

Wang, L., S. Ochoa-Rodriguez, C. Onof, and P. Willems, 2015: Singularity-sensitive gauge-based radar rainfall adjustment methods for urban hydrological applications. *Hydrol. Earth Syst. Sci.*, **19**, 4001–4021, https://doi.org/10.5194/hess-19-4001-2015.

Wilson, J. W., and E. A. Brandes, 1979: Radar measurement of rainfall—A summary. *Bull. Amer. Meteor. Soc.*, **60**, 1048–1058, https://doi.org/10.1175/1520-0477(1979)060<1048:RMORS>2.0.CO;2.

Wood, S. J., D. A. Jones, and R. J. Moore, 2000: Static and dynamic calibration of radar data for hydrological use. *Hydrol. Earth Syst. Sci.*, **4**, 545–554, https://doi.org/10.5194/hess-4-545-2000.

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.