High-Resolution GFS-Based MOS Quantitative Precipitation Forecasts on a 4-km Grid

JEROME P. CHARBA AND FREDERICK G. SAMPLATSKY

NOAA/National Weather Service, Meteorological Development Laboratory, Silver Spring, Maryland

(Manuscript received 14 September 2009, in final form 13 August 2010)

ABSTRACT

The Meteorological Development Laboratory (MDL) of the National Weather Service (NWS) has developed high-resolution Global Forecast System (GFS)-based model output statistics (MOS) 6- and 12-h quantitative precipitation forecast (QPF) guidance on a 4-km grid for the contiguous United States. Geographically regionalized multiple linear regression equations are used to produce probabilistic QPFs (PQPFs) for multiple precipitation exceedance thresholds. Also, several supplementary QPF elements are derived from the PQPFs. The QPF elements are produced (presently experimentally) twice per day for forecast projections up to 156 h (6.5 days); probability of (measurable) precipitation (POP) forecasts extend to 192 h (8 days). Because the spatial and intensity resolutions of the QPF elements are higher than that for the currently operational gridded MOS QPF elements, this new application is referred to as high-resolution MOS (HRMOS) QPF.

High spatial resolution and enhanced skill are built into the HRMOS PQPFs by incorporating finescale topography and climatology into the predictor database. This is accomplished through the use of specially formulated "topoclimatic" interactive predictors, which are formed as a simple product of a climatology- or terrain-related quantity and a GFS forecast variable. Such a predictor contains interactive effects, whereby finescale detail in the topographic or climatic variable is built into the GFS forecast variable, and dynamics in the large-scale GFS forecast variable are incorporated into the static topoclimatic variable. In essence, such interactive predictors account for the finescale bias error in the GFS forecasts, and thus they enhance the skill of the PQPFs.

Underlying the enhanced performance of the HRMOS QPF elements is extensive use of archived fine-grid radar-based quantitative precipitation estimates (QPEs). The fine spatial scale of the QPE data supported development of a detailed precipitation climatology, which is used as a climatic predictive input. Also, the very large number of QPE sample points supported specification of rare-event (i.e., ≥1.50 and ≥2.00 in.) 6-h precipitation exceedance thresholds as predictands. Geographical regionalization of the PQPF regression equations and the derived QPF elements also contributes to enhanced forecast performance.

Limited comparative verification of several 6-h model QPFs in categorical form showed the HRMOS QPF with significantly better threat scores and biases than corresponding GFS and operational gridded MOS QPFs.

Limited testing of logistic regression versus linear regression to produce the 6-h PQPFs showed the feasibility of applying the logistic method with the very large HRMOS samples. However, objective screening of many candidate predictors with linear regression resulted in slightly better PQPF skill.

1. Introduction

Forecast skill of quantitative precipitations forecasts (QPFs) in the National Weather Service (NWS) is low, especially during the warm season (Fritsch and Carbone 2004). While gradual improvements in QPF skill have been made over the years, the rate of improvement has lagged corresponding improvements in large-scale

circulation forecasts by operational numerical weather prediction (NWP) models. The skill disparity is widely believed to be due to the relatively coarse spatial resolution in operational NWP models: intense precipitation is associated with mesoscale systems, whereas these NWP models primarily resolve synoptic-scale circulations. This implies that a key to improving QPF resides in downscaling predictive information from current NWP models or using finer-resolution models.

The performance of precipitation forecasts from operational NWP models reflects the large-scale nature of the models. Light precipitation is overpredicted, while intense precipitation is underpredicted (Jensenius 1990),

Corresponding author address: Jerome P. Charba, National Weather Service, 1325 East-West Highway, Silver Spring, MD 20910-3283.
E-mail: jerome.charba@noaa.gov

which the authors have found in archived National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS; Kanamitsu 1989; Kanamitsu et al. 1991; Iredell and Caplan 1997) precipitation forecasts from 2001 to the present. We have also seen a similar bias trend in the NCEP Nonhydrostatic Mesoscale Model (Rodgers et al. 2005), though to a somewhat lesser degree. (Documentation of these findings is planned for a forthcoming article.) Such model bias suggests inadequacy of the parameterization of convective precipitation in these coarse-mesh operational models, as recent experimentation with fine-grid nonhydrostatic ("convection allowing") numerical models to improve precipitation forecasts have yielded encouraging results (e.g., Clark et al. 2009).

Model output statistics (MOS) postprocessing of NWP output has been used by the NWS to produce unbiased QPFs and provide uncertainty (probability) estimates for over three decades (Antolik 2000; Ruth et al. 2008). However, conventional MOS QPF applications are hampered by short samples from stable NWP models (Carter et al. 1989; Antolik and Baker 2009), the necessity of using broad precipitation categories as predictands, poor sharpness of forecast probabilities (Wilks 2006) for rare heavy-precipitation categories, and the inability to predict very heavy precipitation events. Another limitation of the operational MOS QPFs is that the MOS station distribution is quite irregular over the forecast domain (Gilbert et al. 2009).

The recent implementation of the gridded National Digital Forecast Database (NDFD; Glahn and Ruth 2003) by the NWS has created a strong demand for corresponding National Digital Guidance Database (NDGD) grids. In response to the demand, NDGD grids for a number of weather elements, including 6-h and 12-h QPFs, are being operationally produced from MOS station forecasts (Glahn et al. 2009; Gilbert et al. 2009) through application of tailored objective analysis techniques.

The objective of this study is to enhance the spatial and intensity resolution of the current gridded MOS-based (GMOS) QPFs through application of new data inputs and new techniques. A fundamental departure here from the GMOS QPF approach is that the QPFs are produced by applying the predictive equations directly on a fine-mesh (4 km) grid instead of (initially) at irregularly spaced MOS stations. While such a direct-grid approach is not new—it has been applied previously for the prediction of thunderstorms (Hughes 2004; Charba and Samplatsky 2009), severe local storms (Hughes 1999), and QPF (Charba 1998)—the present use of a fine-mesh grid together with a geographically regionalized approach (Lowry and Glahn 1976) presented new challenges and new opportunities.

In a companion article, Charba and Samplatsky (2011, hereinafter CS) describe how geographical regionalization

together with use of a fine-mesh grid necessitated the development of procedures to avoid spatial discontinuities in the QPF products across regional boundaries. CS also briefly describe how the use of finescale terrain and climatological data resulted in enhanced spatial resolution and skill in the gridded QPF products. Thus, this new MOS QPF application is referred to as high-resolution MOS (HRMOS) to distinguish it from the current GMOS QPF method.

This article provides an overview of the HRMOS QPF model. The precipitation predictand is described in section 2 and the various types of potential predictors are discussed in section 3. Geographical regionalization of the model is briefly discussed in section 4, highly ranked predictors in the probabilistic quantitative precipitation forecast (PQPF) linear regression equations are examined in section 5 along with limited tests with the nonlinear logistic regression model, and the development of several derived precipitation products is described in section 6. A brief examination of the properties and performance of the QPF elements is presented in section 7, and a summary and comments are contained in section 8.

2. Precipitation predictand specification

At the NWS, QPF products are commonly issued for 6-, 12-, and 24-h periods. The 12-h QPF products form an important source of input into public weather forecasts, and 24-h products are used primarily in water management operations. An important application of the 6-h QPFs is their ingestion into streamflow prediction models run at NWS River Forecast Centers (RFCs). These models predict flow levels for small streams and rivers across the United States. Since flooding along these water bodies poses a major threat to human life and property, the production of high-quality 6-h QPF products was the focus in the HRMOS model development; the production of similar 12-h QPF products was a secondary objective.

a. Precipitation database

The database used for specifying the 6- and 12-h precipitation predictands consisted of "stage III" 6-h quantitative precipitation estimates (QPEs), which are produced in real time at RFCs. At each of the 12 conterminous U.S. (CONUS) RFCs, the QPEs are specified on a polar stereographic grid with a 4.762-km mesh at 60°N for a local subset of the national Hydrologic Rainfall Analysis Project (HRAP) grid. This grid is commonly referred to as the 4-km HRAP grid, as the mesh is just over 4 km at midlatitudes. For the nine CONUS RFC service areas east of the Continental Divide (see map online at http://www.nws.noaa.gov/ahps/rfc/rfc.php), the stage III QPE analyses

are based on the combined inputs from Weather Surveillance Radar-1988 Doppler (WSR-88D) precipitation estimates and quality-controlled (QCed) gauge observations (Fulton et al. 1998). The QPEs for the three western RFCs are produced with a climatology-enhanced objective analysis of QCed gauge observations (Henkel and Peterson 1996); radar-based precipitation estimates are not used in this rugged mountainous region of the CONUS because they are not considered reliable there.

The QPEs for the HRAP grid subsets are composited at the NCEP Central Computer System (CCS) to form a "stage IV" national mosaic (information online at http://www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/stage4/), an example of which is shown in Fig. 1a. For use in this study, an archive of the stage IV data spanning January 2001–March 2010[1] was obtained from the National Precipitation Verification Unit of NCEP.

Although human quality control of the stage III analyses is performed at each RFC (Fulton et al. 1998), we found random and systematic errors in the national grids. Random errors, due largely to erroneous gauge observations, were common in the early years of the archive, but such error in recent grids is rare. Systematic error in the QPE grids is manifested mainly as precipitation underreporting, which results from poor radar coverage, beam blockage, and beam overshooting of precipitation. To a far lesser degree, systematic error in the QPEs also arises from the nonrandom incidence of radar beam anomalous propagation (AP), which was more common in the early part of the archive. Improvements in the multisensor precipitation processing algorithms over time (Fulton et al. 1998) have resulted in AP reduction, but precipitation underreporting due to incomplete radar coverage is unchanged.

Thus, we developed QC procedures to detect and screen random and systematic errors in the QPE data in two serial steps. First, random error was detected and screened through a "human in the loop" dynamic process, whereby individual grids were subjected to automated spatial and temporal consistency checks. From a preliminary automated dynamic QC run, the lead author inspected computer-generated error flags and subgrid masks for individual grids. When the inspector disagreed with an error flag or subgrid mask, he issued either an override of the error flag or an adjustment to the subgrid mask for inclusion in a final dynamic QC run.

The systematic data error detection and masking process, which comprised the second QC step, is "grid static" in the sense that QPE values were set to missing for a seasonally fixed set of grid points in every warm season (April–September) or cool season (October–March) grid. The data-reject grid points were predetermined from careful examination of 6-h precipitation seasonal mean relative frequency (SRF) maps for multiple precipitation exceedance thresholds (see section 3b). Unrealistic SRFs at individual 4-km grid points were identified through cross-checking against detailed geophysical and documented precipitation climatology maps (Daley et al. 1994; Charba et al. 1998). The QPE data were set to missing at the invalid SRF grid points in an ensuing grid-static data masking computer run.

Results from the QC processing revealed that the QPE discard rate with the dynamic and static masks combined was about 6% during the warm season and about 12% during the cool season, the vast portion of which resulted from the grid-static masking. While the discarded fraction from the dynamic masking was extremely small (averaging about 0.02% over the entire archive), it is considered significant because the rejected data often fell near the upper bound of valid precipitation amounts. The higher discard rate for the cool season resulted from the grid-static masking of radar-underdetected shallow stratiform precipitation. In fact, the broad coverage of the grid-static data masking, especially over the high plains of the United States, prompted subsequent insertion into those areas of QCed 6-h precipitation gauge observations from aviation routine weather reports (METARs). Corresponding to the raw QPE field in Fig. 1a, the QCed QPE field, supplemented with these gauge data, is shown in Fig. 1b.

It is worth noting that a routine aspect of the QPE QC processing is a statistical comparison of 24-h precipitation based on the QPE data with QCed precipitation gauge observations at roughly 7000 stations over the CONUS. For this comparison the 6-h QPE data are accumulated over 24 h; the QCed gauge observations were supplied by the NCEP Hydrometeorological Prediction Center (HPC). Averaged over the years 2001–09 and the full CONUS domain, the root-mean-squared difference, mean absolute difference, and mean difference for the cool (warm) season were 0.13, 0.04, and 0.0 in. (0.17, 0.05, and 0.0 in.), respectively. The smallness of these 24-h precipitation difference measures indicates that the QCed QPE and gauge data could be used interchangeably for most meteorological applications.[2]

---

[1] The NPVU archive of the QPE data begins on 1 October 2000, but data for the period October–December 2000 were not used because of their very poor quality.

[2] It is noted the gauge versus QPE differences may be larger east of the U.S. Continental Divide than to the west because the QPE analyses in the latter domain are based only on gauge data.
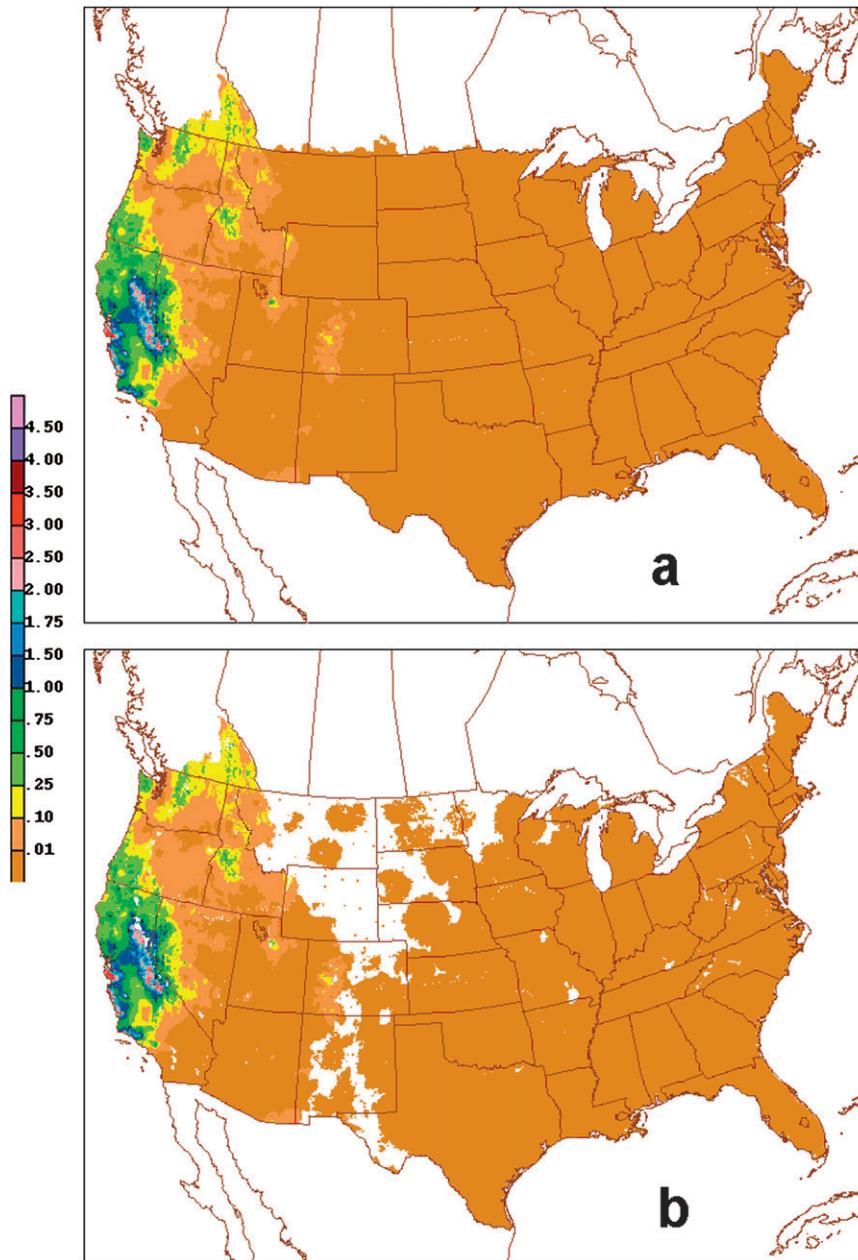
FIG. 1. Stage IV quantitative precipitation estimates (in.) for the 6-h period ending 0000 UTC 5 Jan 2008 (a) before and (b) after QC procedures were applied. Missing data (white areas) in (b) reflect application of QC procedures and brown specks in those locations show where precipitation gauge data were subsequently inserted.

b. Predictand specification

The HRMOS predictand was specified from the QCed QPE data for individual 4-km HRAP grid boxes as a binary variable with a value of 1 when a precipitation exceedance threshold was met and 0 otherwise. The 6-h exceedance thresholds are ≥ 0.01, ≥ 0.10, ≥ 0.25, ≥ 0.50, ≥ 0.75, ≥ 1.00, ≥ 1.50, and ≥ 2.00 in. It is noted that the peak 6-h precipitation threshold used in the currently operational MOS QPF product is ≥ 1.00 in. (Antolik 2000). The addition of the very rare event (i.e., ≥ 1.50 and ≥ 2.00 in.) 6-h precipitation exceedance thresholds was possible because the number of grid points with QCed QPE data is far greater than the number of CONUS MOS stations with trustworthy gauge data ($\sim$420 000 versus $\sim$300). With such a large number of sampling points and
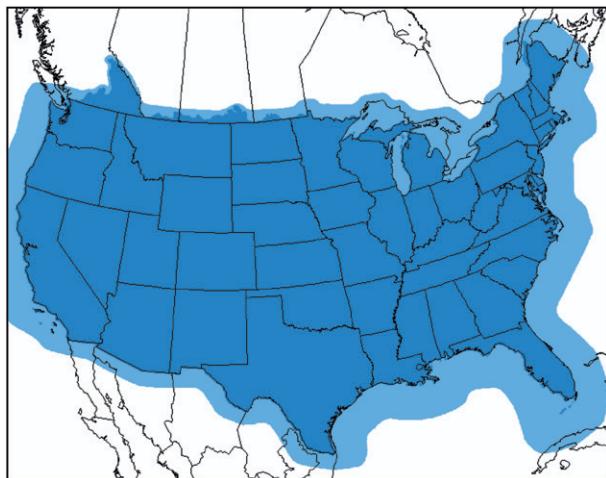
FIG. 2. HRMOS developmental domain (medium blue shading) and the required coverage for NWS operational QPF products (light blue shading). The map border coincides with the bounds of the HRAP grid.

TABLE 1. (left) Basic potential predictor variables and (right) the applicable height level(s). All references to precipitation pertain to standard 6-h periods. Abbreviations are DOY, day of the year; RF, relative frequency; precip, precipitation; CTG, cloud to ground; u, westerly wind component; v, southerly wind component; MSL, mean sea level; RH, relative humidity; and CAPE, convective available potential energy.

Variable	Level (s)
(a) Constant data	
Terrain elev	—
Lat	—
Lon	—
Sine, cosine DOY	—
Seasonal RF for multiple precip categories	—
Monthly RF for multiple precip categories	—
PRISM monthly-mean precip	—
Monthly RF of CTG lightning	—
(b) GFS forecasts	
Tot (grid scale + convective) precip	—
Convective precip	—
Vertical velocity	700 mb
u, v wind components	Multiple levels
MSL pressure	
Pressure height	Multiple levels
Mean RH	Multiple layers, precipitable water 1000–500 mb
Terrain-forced vertical velocity*	10 m, 850 mb
Divergence of the moisture flux	850 mb
Equivalent potential temperature	850 mb
Best lifted index	—
CAPE	Surface based
K index	—
Total totals index	—

* Based on GFS u, v wind components at 10 m and 850 mb.

a historical archive of QPE data spanning 2001–10, statistically robust samples were formed even for the ≥1.50- and ≥2.00-in. 6-h precipitation thresholds. Finally, the exceedance thresholds for the 12-h periods are identical to those for 6 h, except for the addition of a peak 12-h threshold of ≥3.00 in. (The peak 12-h threshold for MOS is ≥2.00 in.)

The predictand for development of the HRMOS model was specified for 4-km grid boxes over the stage IV data coverage area (denoted HRMOS developmental domain in Fig. 2). The corresponding candidate predictor variables, discussed in the next section, are defined for the same locations. However, to meet the NWS CONUS service requirements, operational QPFs must be produced for the expanded coverage area shown in Fig. 2. Thus, predictor fields used for operational application of the HRMOS model must also be defined for this expanded area, which posed a challenge for predictor variables derived from data confined to the developmental area.

3. Candidate predictor specification

Forecast output from the NCEP GFS provided the dynamic predictor base for the HRMOS application. Although convective parameterization in the GFS can result in mesoscale detail in the forecast precipitation field, other model variables are of the synoptic scale. Thus, to increase mesoscale information in the predictor data, finescale topography and climatological data were incorporated as "static" predictor inputs. The approach used for incorporating the "topoclimatic" data is similar

to that used in a previous short-range statistically based QPF model (Charba 1998), but here these data inputs are more extensive and they contain enhanced spatial resolution. The basic forms of these topoclimatic inputs are listed in Table 1a under the heading "constant data," and the basic GFS forecast variables are listed in Table 1b. These variables are denoted as "basic" as additional candidate predictors (described later) are derived from them. In the following subsections, we first address how each of the topoclimatic variables was defined.

a. Topography

Terrain elevations provided by the U.S. Geological Survey on a 30-arc-s latitude–longitude grid were interpolated onto the 4-km HRAP grid. This initial terrain field was then smoothed through conventional grid
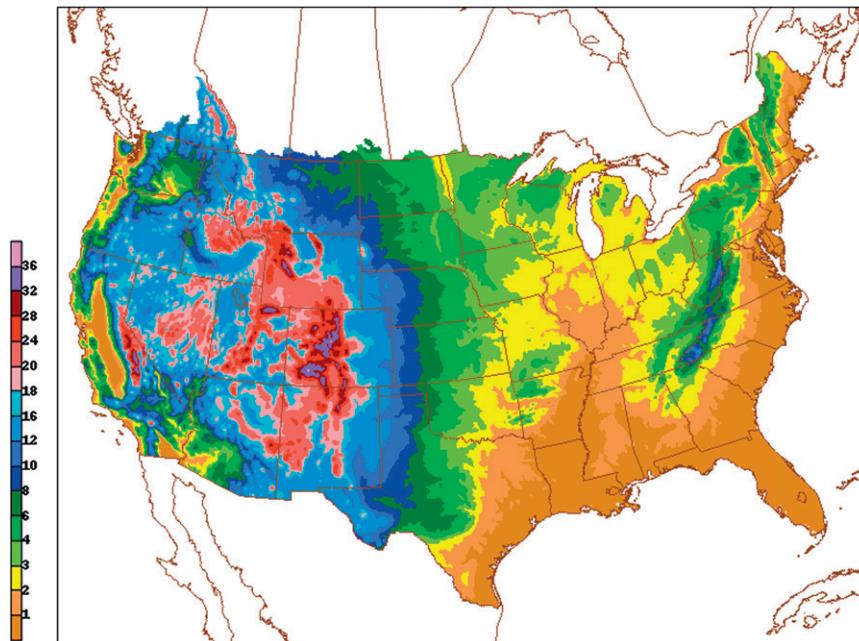
FIG. 3. Terrain elevation (hundreds of meters) for the developmental area.

smoothing to optimize the predictive effectiveness of a stand-alone terrain elevation predictor and terrain-induced vertical velocity (the primary predictor usage of the terrain data). The utility of these terrain-based predictors was tested on the basis of sensitivity experiments, whereby PQPF skill from predictive regression equations was objectively measured for various levels of terrain detail (not shown). The level of terrain detail that resulted in optimal PQPF skill is shown in Fig. 3; nevertheless, substantial smoothing of the derived terrain-induced vertical velocity was necessary.

b. Precipitation climatology

Three types of gridded precipitation climatologies were included as predictor inputs (Table 1a), each with unique properties and limitations. One type consisted of 6-h precipitation SRFs, which were computed for each 4-km grid box and each of the eight 6-h precipitation exceedance thresholds from an 8-yr (2001–08) archive of the HRMOS predictand (section 2b). To obtain adequate statistical samples at individual grid boxes, the SRFs were averaged over the 6-month warm or cool seasons and over the four 6-h periods of the day.

An example of the "raw" cool season SRFs is shown in Fig. 4a, which is for the ≥0.10-in. 6-h precipitation exceedance threshold. This raw SRF field contains several impediments to direct predictor usage, which include (a) extensive data coverage gaps, especially from

West Texas to the Canadian border (results from the QC screening of the QPE data); (b) excessive spatial incoherency throughout the SRF field, which is due largely to shortness of the underlying sample; c) fictitious flower-shaped patterns (due to range dependency of radar estimated precipitation), which is especially problematic in the U.S. central and northern plains; and (d) insufficient geographical coverage for operational application of the HRMOS model.

The treatment of these deficiencies involved reconstruction and expansion of the raw SRF field. A key aspect of the reconstruction involved localized grid filling and smoothing, which was performed manually with the aid of interactive grid-editing software (Wier et al. 1998) together with guidance from precipitation climatology maps from Charba et al. (1998) and Daley et al. (1994). The geographical coverage expansion also involved manual grid filling, whereby miscellaneous precipitation climatology information as well as cloud-to-ground lightning climatology maps (section 3c) was used as guidance. These localized grid-filling and smoothing processes were followed by conventional automated grid smoothing to remove residual finescale noise over the full domain.

The reconstructed SRF field corresponding to Fig. 4a is shown in Fig. 4b. Note that it contains high spatial coherency as well as mesoscale detail, which includes strong SRF maxima (minima) on the windward (lee) slopes of significant mountain chains across the United States (Fig. 3) and also weak maxima along the lee
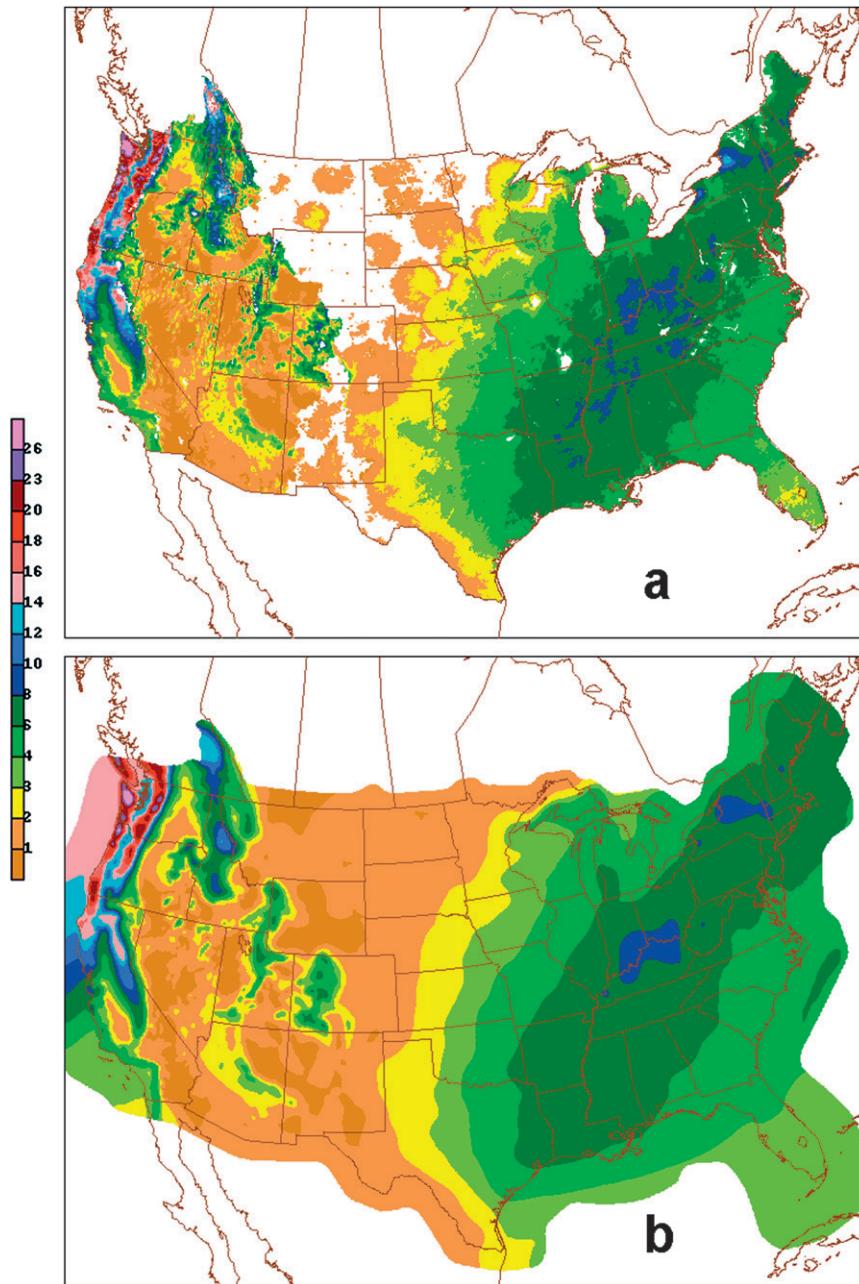
FIG. 4. (a) Raw cool SRF (%) of ≥0.10 in. and (b) the corresponding reconstructed SRF field for the expanded coverage area (Fig. 2).

shores of the Great Lakes.[3] Similarly, reconstructed cool season SRF fields were also produced for the seven other 6-h precipitation exceedance thresholds and all thresholds for the warm season.

A second type of precipitation climatology (Table 1a) consisted of monthly relative frequencies (MRFs) for each of five 6-h precipitation exceedance thresholds (≥0.10, ≥0.25, ≥0.50, ≥1.00, and ≥2.00 in.), which were developed on a 20-km grid from 33 yr of hourly precipitation gauge measurements from the U.S. cooperative observer network (Charba et al. 1998). For application

---

[3] An unfortunate consequence of the automated grid smoothing used to suppress finescale SRF noise throughout the domain is the amplitude reduction of the real small-scale SRF features such as are found along the lee shores of the Great Lakes (Fig. 4). Subsequent manual editing would be required to restore this detail, the cost of which was considered excessive relative to the likely benefits.
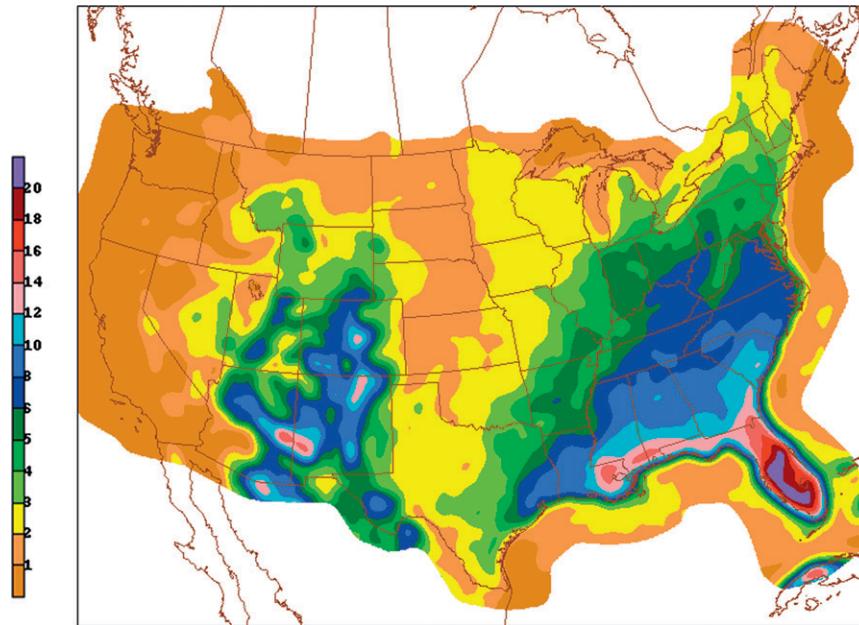
FIG. 5. July relative frequency (smoothed; %) of one or more CTG lightning strikes in 4-km HRAP grid boxes during 1800–0000 UTC for the expanded coverage area.

here, the MRF fields were interpolated onto the HRAP grid. It is noted that the MRF and SRF precipitation climatologies (discussed above) complement one another, as the MRFs vary by month and time of day whereas the SRFs contain higher spatial detail and additional precipitation thresholds (≥0.01, ≥0.75, and ≥1.50 in.).

The third precipitation climatology type (Table 1a), commonly known as the Parameter-elevation Regressions on Independent Slopes Model (PRISM; Daley et al. 1994), consists of monthly-mean precipitation amounts on a 30-arc-s latitude–longitude grid that spans the CONUS. This precipitation climatology type was incorporated here to complement the SRF and MRF climatologies, as it includes both statistical analyses and finescale orographic precipitation modeling in its formulation.

As for the SRFs, it was necessary to extend the areal coverage of the MRF and PRISM climatologies to the expanded CONUS domain (Fig. 2). These expansions were performed objectively with a simple iterative diffusion procedure, whereby the reconstructed and expanded SRF field for ≥0.10 in. (Fig. 4b) was used as a key control variable.

c. Lightning climatology

The fourth climatological predictor type (Table 1a) consists of monthly relative frequencies (LRFs) of one or more cloud-to-ground (CTG) lightning strikes within the 4-km HRAP grid cells and each of the four 6-h periods of the day. The LRFs were derived from a 15-yr (1994–2008) archive of CTG strikes from the National

Lightning Detection Network (NLDN; Cummins et al. 1998). Spatial and month-to-month temporal smoothings were applied to raw LRF grids to enhance spatial coherency and intermonth consistency at the smallest resolvable scales. An example map valid for 1800–0000 UTC during July is shown in Fig. 5. The LRF climatology complements the three precipitation climatologies as it contains increased spatial detail in the eastern United States, good temporal resolution, and coverage beyond the CONUS boundaries.

d. GFS model output variables

Candidate predictor variables from GFS output were prepared by first interpolating the model fields to the HRAP grid from a Meteorological Development Laboratory (MDL) archive on a polar stereographic grid with a 95.25-km mesh length (called an 80-km grid since the grid spacing is about 80 km at midlatitudes).[4] Of the full list of GFS variables in Table 1b, the first six are basic model variables, whereas the remaining ones are mostly derived quantities. Also, while this list of variables may appear short, the actual number of candidate GFS predictors was far larger (approaching 100) because of several factors. First, continuous variables were converted to multiple grid binary variables (Jensenius 1992) to help account for nonlinear associations with the precipitation

---

[4] The GFS grids in the MDL archive are interpolated from the native model grid.

TABLE 2. Derived interactive (potential) predictor variables formed as a product of variable 1 and variable 2 (and variable 3 for three-component variables). All references to precipitation pertain to std 6-h periods. Nonstandard abbreviations, in addition to those in Table 1, are vert. vel., vertical velocity; TVV, terrain vertical velocity; GB, grid binary; comp., component; convec., convective; equiv. pot., equivalent potential; and neg., negative.

Variable 1	Variable 2	Variable 3
(a) Topography–climatology		
1) TVV (upward; GFS 850-mb u, v comp.)	GFS total precip ≥0.04 in. (GB)	—
2) TVV (upward; GFS 10-m u, v comp.)	GFS 1000–500-mb mean RH	—
3) Monthly RF CTG lightning	GFS K index ≥30.0 K (GB)	—
4) Monthly RF precip ≥0.10 in.	Seasonal RF precip ≥0.10 in.	GFS tot precip ≥0.10 in. (GB)
5) Monthly RF precip ≥0.50 in.	Seasonal RF precip ≥0.50 in.	GFS convec. precip ≥0.16 in. (GB)
6) Monthly RF precip ≥0.10 in.	PRISM monthly precip	GFS tot precip ≥0.10 in. (GB)
7) Monthly RF precip ≥0.10 in.	GFS tot precip ≥0.10 in. (GB)	—
8) Monthly RF precip ≥0.25 in.	GFS tot precip ≥0.25 in. (GB)	—
9) Monthly RF precip ≥0.50 in.	GFS convec. precip ≥0.16 in. (GB)	—
10) Seasonal RF precip ≥0.10 in.	GFS tot precip ≥0.10 in. (GB)	—
11) Seasonal RF precip ≥0.25 in.	GFS tot precip ≥0.25 in. (GB)	—
12) Seasonal RF precip ≥0.50 in.	GFS tot precip ≥0.40 in. (GB)	—
13) Seasonal RF precip ≥0.75 in.	GFS tot precip ≥0.65 in. (GB)	—
(b) GFS only		
1) GFS 700-mb vert. vel. (upward)	GFS K index ≥30.0 K (GB)	—
2) GFS 850-mb equiv. pot. temperature	GFS 850-mb u-wind comp.	—
3) GFS 850-mb equiv. pot. temperature	GFS 850-mb v-wind comp. (southerly)	—
4) GFS 700-mb vert. vel. (upward)	GFS precipitable water	—
5) GFS 700-mb vert. vel. (upward)	GFS 850-mb moisture divergence (neg.)	—

predictands and to instill consistency between the binary predictands and the predictors. Where continuous predictor variables are inherently bounded (such as for relative humidity), the continuous form was also used. Second, many of the more useful variables were specified at multiple forecast projections to account for possible phase speed bias in the GFS. Finally, some of the more useful variables were specified not only from the latest GFS cycle, but also from as many as three previous 6-hourly cycles (up to 18 h earlier). The aim of this "ensemble" approach was to enhance forecast skill and intercycle consistency in the HRMOS QPF products.

e. Interactive variables

As noted before, GFS model output variables contain predominantly synoptic-scale predictive information, whereas the "topoclimatic" variables contain finescale detail. A technique by which detail in the latter is infused into the former is through specification of "interactive" variables. Such a variable was defined as a simple product of a selected topography or climatology variable (from Table 1a) with a selected GFS variable (from Table 1b), similar to that used previously by Reap and Foster (1979) and Charba (1998). These product variables approximate "true" interactive variables used by Bocchieri (1979) and Charba (1987), where such a variable was defined as the (statistically derived) relative frequency of occurrence of the predictand for any combination of

values of the component predictor variables. The simple product surrogate was adopted here because the development and implementation costs are far less. This factor together with the extensive topoclimatic inputs available in this study provided the opportunity to specify many interactive variables of this type. The full list of these topoclimatic interactive (TCIA) variables is shown in Table 2a.

A key attribute of a TCIA variable is that it combines, in an interactive manner, predictive aspects of each component variable into a single predictor. Specifically, fine spatial detail that characterizes a topography or climatology variable is infused with large-scale information in a GFS forecast variable. Also, the dynamic aspects of the GFS forecast instill temporal variability into the time-static topography or climatology variable. An example of these interactive effects is shown in Fig. 6b for the product of the cool season 6-h precipitation SRF of ≥0.10 in. (Fig. 4b) and a grid binary representation (with a 0.10-in. exceedance threshold) of the GFS 6-h total (grid scale plus convective) precipitation at the 12-h forecast projection (Fig. 6a). Note that the product results in a conditioning of SRF for ≥0.10 in. with the GFS prediction of 6-h precipitation ≥0.10 in. Close inspection of Figs. 4b, 6a, and 6b reveals that the SRF variable is responsible for the finescale detail in the product field, whereas the GFS forecast of ≥0.10 in. (a binary variable) controls the geographical coverage of nonzero values. From one perspective, the GFS forecast
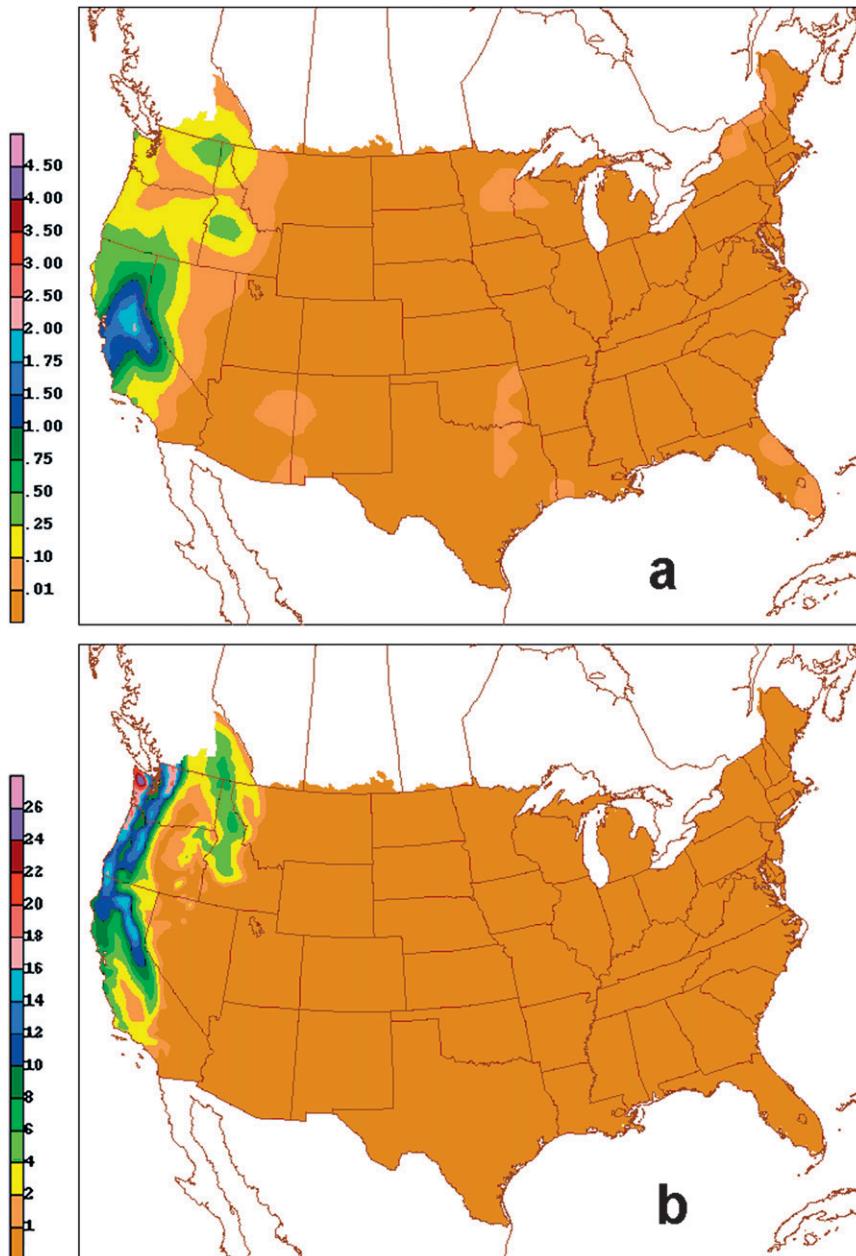
FIG. 6. (a) GFS total precipitation (in.) with a 12-h forecast projection for the 6-h period ending 0000 UTC 5 Jan 2008, and (b) the product of cool season relative frequency (%) of 6-h precipitation ≥ 0.10 in. (Fig. 4b) and the GB representation of (a) with a ≥ 0.10-in. exceedance threshold.

acts as a dynamic control on the static climatology variable. From another perspective, the climatology serves to account for the finescale bias error in the GFS forecast.

This concept underlies other TCIA variables in Table 2a. All variables numbered 4–13 involve one or more types of precipitation climatology together with a GFS forecasted 6-h precipitation threshold. In the case of the

three-component interactive variables (variable numbers 4–6), the aim was to combine the complementary aspects of different precipitation climatology types (section 3b). For variables 1 and 2, the finescale terrain-induced vertical velocity (computed via low-level GFS wind components) is conditioned by GFS measureable precipitation (≥ 0.01 in.) and lower-tropospheric mean relative humidity forecasts, respectively. Finally, for variable number

3, lightning climatology is dynamically conditioned by the GFS-predicted K (convective instability) index.

Another attribute of a TCIA variable is that it allows for a nonlinear relationship between the predictand and the individual component variables. This attribute may be significant because it provides a mechanism by which nonlinear effects can be accounted for in the linear regression model used in the HRMOS QPF application.

It is noted that the level of smoothing applied to a topoclimatic interactive variable was an important aspect of its predictive effectiveness. Since taking the product of two variables essentially multiples the spatial detail in the component fields, the product field tends to exhibit excessive spatial variability ("noise"). The noise was controlled by applying conventional grid smoothing, where the level of smoothing was "tuned" by trial and error to optimize PQPF skill. The specific tuning criterion used was to maximize spatial detail to the degree that PQPF skill is not sacrificed. The tuning was initially tedious, but, with experience, good smoothing estimates could be made from inspection of test PQPF fields.

A smaller set of additional product variables was specified (Table 2b), where both component variables were selected from the GFS group (Table 1b). The aim was to account for possible nonlinear relationships between these GFS product variables and the precipitation predictands. The component pairs were chosen to reflect precipitation mechanisms. For instance, 700-mb vertical velocity is paired with a convective instability index, precipitable water, and divergence of the low-level moisture flux. It is important to note that rare outlier values in some of these product variables were (initially) problematic in the PQPF linear regression equations, as an outlier value could result in a very poor PQPF. The problem was treated by truncating outlier values, where truncation bounds were set at about the 1 (or 99) percentile values of a product distribution. Finally, the avoidance of extreme outlier values was a factor in the decision to use (inherently bounded) grid binary variables as the GFS components of the TCIA variables (Table 2a).

4. Geographical regionalization

Details of the geographical regionalization of the PQPF linear regression equations are contained in CS. It is useful to briefly review the regionalization process here because slight changes have subsequently been made to the regions. Also, regionalization was used in the formulation of the derived QPF elements, which was not addressed in the earlier article.

In an early developmental stage of the HRMOS model, the geographical regions were the nonoverlapping areas shown in Fig. 7, and CS showed these regions resulted in
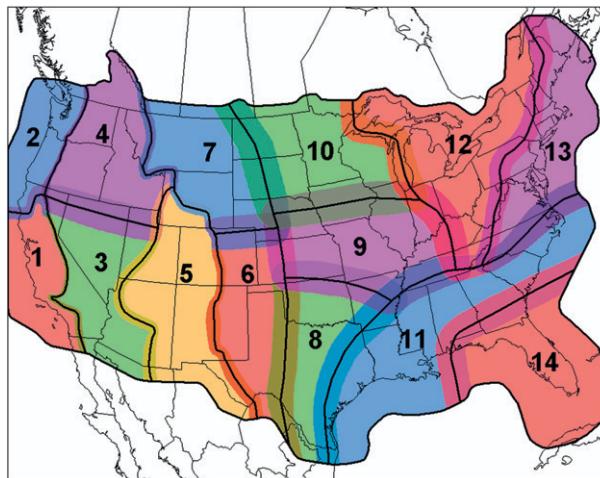

FIG. 7. Overlapping regions (depicted by color shades that change in overlap areas) numbered 1–14, which are used for development and application of the PQPF regression equations and nonoverlapping regions (boldface black lines) used for formulation of the derived QPF products.

artificial discontinuities in PQPF patterns along regional boundaries. The problem was treated by outwardly expanding the boundaries of each nonoverlapping region, and then regression equations were developed for the resulting overlapping regions. Since application of these equations to the overlapping regions results in multiple PQPFs in the overlap zones, a weighting technique was used to blend the multiple PQPFs in those areas. CS showed that the region overlap and weighting techniques mitigated the artificial PQPF discontinuities without degrading forecast skill. Daily inspection of experimental real-time PQPF maps by the authors over a period of almost 2 yr reveals that the current regions do not result in significant regional discontinuities anywhere over the CONUS domain.

The recent slight modification of the regions (noted above) consisted of two changes. One was that a new region (region 14 in Fig. 7) was formed by partitioning off from the original region 11 (Fig. 4 in CS) an area slightly larger than the state of Florida. This was done to treat a unique problem of strong warm season overforecasting of heavy precipitation amounts at very long forecast projections (120 h and longer) over the Florida peninsula. Another change was a westward extension of the northwest portion of region 12 to include the southern shore of Lake Superior. This change consolidates the area of cool season precipitation enhancement of the Great Lakes (Charba et al. 1998 and references therein) entirely within region 12.

The weighting technique constitutes a key component of the regionalization process, as it is used for blending

multiple PQPFs (and other variables involving the derived QPF elements) within the regional overlap zones. Here, we note a few aspects of the weight fields and how they are applied. For example, Fig. 8 shows the weight field for region 7, which features an inner core coinciding with the nonoverlapping portion of the region (Fig. 7) where the weights have a constant value of 1.0. Within bands of variable width along the regional perimeter (which constitute the regional overlap zones), the weights decrease smoothly to 0.0 according to a continuous mathematical function (CS). A key feature of the weights is that for a grid point in an overlap zone the set of weights from all regions sums to 1.0. Thus, the application of the weights to overlapping fields results in seamless transitions across the regions.

5. PQPF regression equations

a. Equation development and PQPF scoring

Separate PQPF linear regression equations were developed for the cool and warm seasons, each of the 14 overlapping geographical regions, 6- and 12-h valid periods, and each forecast projection in 6-h increments[5] in the 12–156-h range for the 1200 and 0000 UTC cycles. In addition, the development was extended to 192 h (8 days) for the ≥0.01-in. threshold (probability of precipitation, POP); PQPFs for higher precipitation thresholds for this very long forecast range are not produced because of their very low skill. The samples for the equation development were formed by aggregating predictor–predictand data within each overlapping region for all cool (or warm) season days for the period 1 January 2001–30 September 2009.

The Brier skill score (BSS),

$$BSS = 1 - \frac{BS}{CL}, \qquad (1)$$

which measures the fractional improvement in Brier score (BS; Brier 1950) of the PQPFs over the BS for climatology (CL), was used to assess the skill of the PQPFs. Here, BS is given by

$$BS = \frac{1}{n}\sum_{k=1}^{n}(f_k - o_k), \qquad (2)$$

where n is the number of forecasts in the verification sample, f_k is the PQPF whose range is 0.0–1.0, and o_k is the observation with a value of 1 when the precipitation

amount is greater than or equal to the threshold and 0 otherwise. Finally, CL is the climatology score given by (2), where the predictand SRF defined in section 3b was used as the climatological forecast.[6]

To assess whether the BSS corresponding to one set of PQPFs is significantly different (in a statistical sense) from the BSS for another set, we used the two-tailed paired t test (Wilks 2006), where the null hypothesis is that there is no difference in skill between the two sets. In these tests the BSS was computed separately for individual days based on all CONUS grid points, all days of an independent season (about 180 days) were used as the verification sample, and we assume that the sample average of $(BSS_1 - BSS_2)$ values (where the subscripts distinguish the two PQPF sets) approximate a Gaussian distribution.[7] This formulation was used in all significance tests conducted in this study.

b. Unconditional versus conditional predictands

Two approaches (UC and CND) for obtaining unconditional PQPFs were tested. With the UC method the predictand (section 2b) is unconditional, and thus the associated regression equations (directly) yield unconditional PQPFs. Also, for a given cycle, season, and forecast projection, the same set of predictors is used for all precipitation thresholds (predictands). This set was obtained through forward-selection screening regression (Glahn and Lowry 1972; Wilks 2006) applied simultaneously to the eight 6-h precipitation thresholds, whereby a given predictor is selected as it yields the maximum incremental reduction of variance (among all unused predictors) for one of the thresholds. Predictor screening over all predictands simultaneously greatly reduces the computational cost, and the common predictor set may enhance PQPF consistency among the multiple precipitation thresholds.

With the CND method, only the measureable precipitation predictand is unconditional; the predictands

---

[5] Consecutive valid times of the 12-h periods overlap since they are in 6-h increments.

[6] Because the MRF data (described in section 3b) are stratified by month and time of the day, they should be more representative of the true precipitation climatology; however, these data were not used because the ≥0.01-in. precipitation threshold is not available.

[7] It is cautioned that for a full-season sample, dependencies in the QPE observations may exist among consecutive days. Since the paired t test assumes a random sampling of independent observations, consecutive-day samples may underestimate the true variability of the BSS differences and thus overestimate the statistical significance of the BSS differences. A random withholding of days from the developmental dataset would provide a more robust verification sample, and it will be considered in future work. Also, cross validation (Wilks 2006) is another preferred method of forecast model testing, but this computationally expensive approach was not feasible in this study because of its very large cost.

FIG. 8. Map of weight constants for region 7.

FIG. 9. BSS (%) of 6-h PQPFs for multiple precipitation thresholds (in.) over the CONUS domain with the UC and CND predictand approaches (see text). The scores are averaged over four day 1 forecast projections (12, 18, 24, and 30 h) from the 1200 UTC cycle for the 2007/08 cool season.
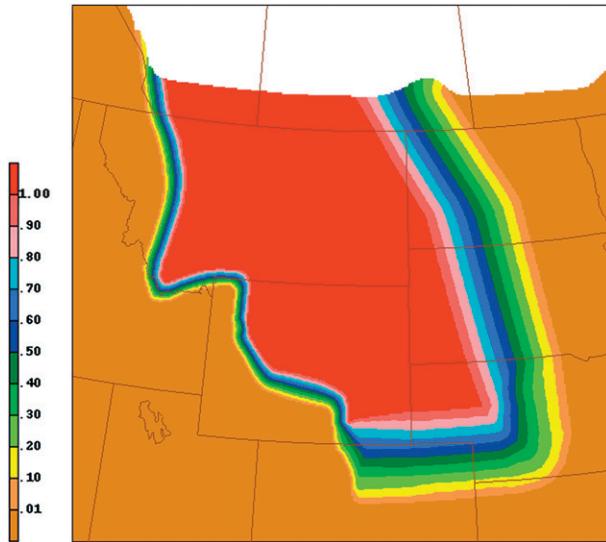
for the seven remaining precipitation thresholds are conditioned on the occurrence of measureable precipitation. Also, one predictor set is screened for the unconditional POP predictand, and a different set is screened for the seven conditional predictands. Since regression equations for the latter predictands yield conditional PQPFs (CPQPFs), unconditional probabilities are obtained by taking the product of POP and CPQPF. The CND method has long been used at MDL especially for applications where the distribution of predictand occurrences among the multiple thresholds is highly nonuniform and event occurrences for some thresholds are very rare (the case here).

BSS comparisons of unconditional 6-h PQPFs with the UC and CND methods showed CND yielded slightly higher skill for both seasons and cycles and all forecast projections. For example, Fig. 9 shows comparative BSSs averaged over four day 1 forecast projections (12, 18, 24, and 30 h) for the 2007/08 cool season. Note that the BSS with the CND method (BSS_{CND}) is slightly higher than for BSS_{UC} for all but the (inherently unconditional) ≥ 0.01-in. threshold. The t test indicated that the BSS_{CND} improvement on BSS_{UC} for thresholds in the ≥ 0.10 to ≥ 1.00-in. range is significant at the 99% confidence level, whereas the small improvement for ≥ 2.00 in. is significant at the 95% confidence level.

Increased skill with the CND method is likely due to two factors. One is that the predictors for the conditional predictands are substantially different from those for the unconditional POP predictand, as shown in the following subsection; that is, the two predictor sets are tailored to the unique predictands. A second factor is that the CND method can account for nonlinear relationships
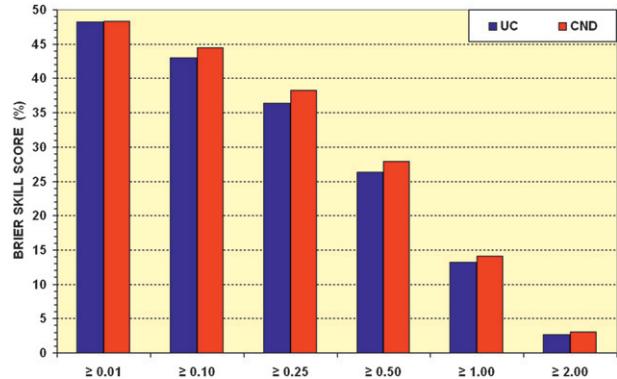
between the unconditional predictands and the predictors even with the use of linear regression equations. Since taking the product of POP and CPQPF to obtain an unconditional PQPF is equivalent to taking the product of the underlying linear regression equations, individual predictor terms now involve products of the original predictors. Accordingly, the unconditional predictands are functions of the (inherently nonlinear) product variables. Thus, all subsequent references to unconditional PQPFs in this article are based on the CND method.

c. Predictor properties

With forward-selection screening regression, the number of predictors selected for inclusion in a POP or CPQPF regression equation was optimized on the basis of several criteria. The criteria consist of an arbitrary upper bound (19) on the number of predictors, a minimum incremental reduction of variance condition, and a predictor colinearity restriction. Parameter values associated with the latter two controls were chosen by optimizing statistical forecast performance properties of the unconditional PQPFs [such as BSS, reliability, and sharpness (Wilks 2006)] for independent samples.

The number of predictors among the regression equations varied widely as a function of the predictand type, geographical region, and forecast projection. The overall range for the cool season was 4–18 among the POP equations, with an average of about 10; the corresponding range for the CPQPF equations was 12–19, with an average of about 18. There was little difference in the number of predictors for the warm season. Since the upper bounds of these ranges slightly exceed the number of predictors commonly used in MOS applications at MDL (Glahn 1985; Antolik 2000), a test of BSS versus the number of predictors was conducted for
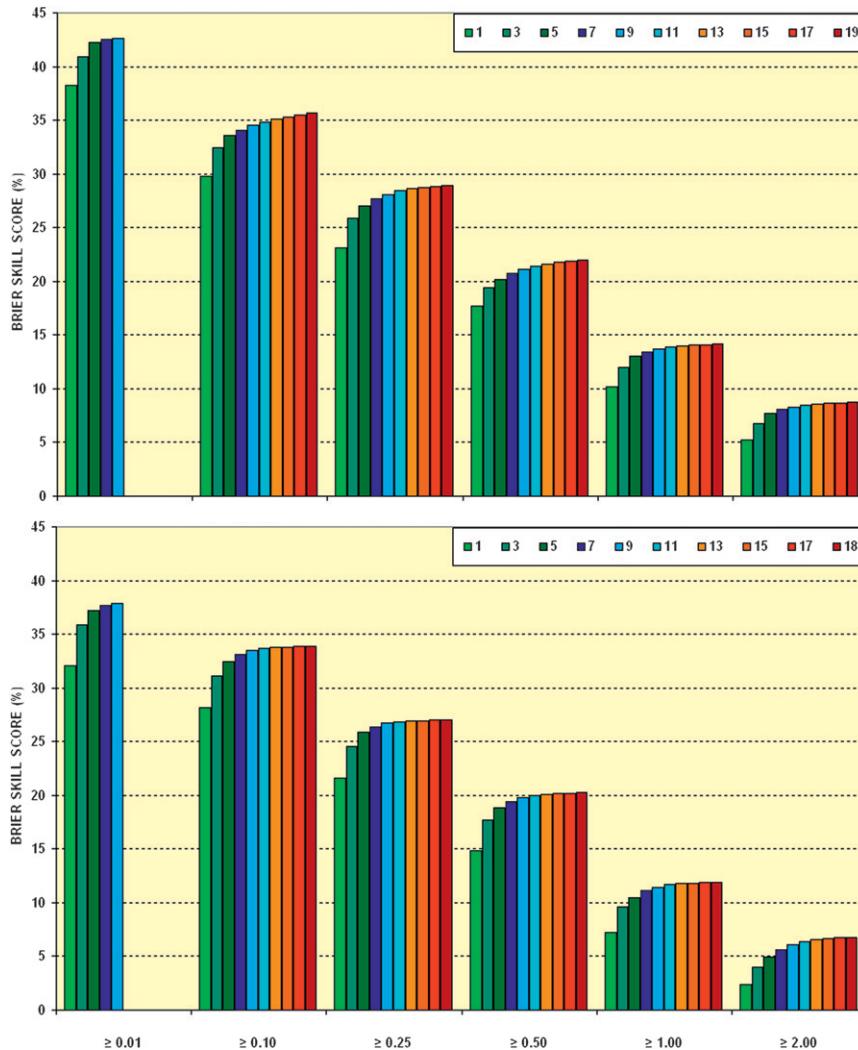
FIG. 10. BSSs for *unconditional* PQPFs vs precipitation threshold (in.) for cool season regression equations with a varying number of predictors (the number is indicated in the legends). The top chart applies to the 12-h projection and the bottom chart to the 114-h projection. The verification sample is for the CONUS domain and the 2008/09 cool season. The ≥0.01-in. threshold applies to the POP equations where the predictor maximum is 9; all other thresholds apply to the CPQPF equations (see text) where the predictor maximum is 19 (18) for the 12-h (114 h) projection. The unconditional PQPF was obtained by taking the product of POP and CPQPF for equations with identical numbers of predictors, but where the number of CPQPF predictors exceeded 9, the nine-term POP equations were linked with CPQPF equations with 11, 13, etc. predictors.

a short forecast projection (12 h) and a long projection (114 h) for the cool season (Fig. 10). The monotonic increase of BSS with the increasing number of predictors seen here does not indicate overfitting of the developmental data (Wilks 2006). Also, the flattening of the BSS profiles for high predictor numbers indicates that inclusion of additional predictors would not benefit PQPF performance.

Assessments of the more important predictors in the various regression equations were conducted by ranking

and summarizing predictors across groups of equations (to avoid being misled by possible rank order anomalies in individual equations[8]). The predictor ranking took into account the order of selection in individual equations

---

[8] Rank order anomalies can occur because of high intercorrelation among the predictors. High intercorrelations are due to many factors, which include the use of physically similar variables, multiple grid binary thresholds, multiple forecast projections, and multiple GFS cycles (section 3d).

TABLE 3. Nine highest-ranked predictors, which are listed in the order of decreasing rank, in cool season 6-h (a) POP and (b) CPQPF (see text) regression equations for all regions and four day 1 forecast projections (12-, 18-, 24-, and 30-h) combined (1200 UTC cycle). All references to precipitation apply to 6-h periods, and nonstandard abbreviations are as in Tables 1 and 2. Most of the stand-alone GFS predictors are in GB form, but GB thresholds are shown only for the interactive predictors.

(a) POP
  1) GFS tot precip (GB form with most precip thresholds less than 0.10)
  2) GFS mean layer RH (1–0.44 sigma)
  3) (Seasonal RF precip ≥0.10 in.) × [GFS tot precip. ≥0.10 in. (GB)]
  4) [Seasonal RF precip (mostly 0.01-in. threshold)]
  5) GFS 700-mb vert. vel.
  6) GFS 1000–500-mb mean RH
  7) [TVV (upward; GFS 10-m u, v comp.)] × (GFS 1000–500-mb mean RH)
  8) (Monthly RF precip ≥0.10 in.) × [GFS tot precip ≥0.10 in. (GB)]
  9) GFS K index
(b) CPQPF
  1) GFS tot precip (GB form with most precip thresholds 0.10 in. and greater)
  2) (Seasonal RF precip ≥0.25 in.) × [GFS tot precip ≥0.25 in. (GB)]
  3) (Seasonal RF precip ≥0.75 in.) × [GFS tot precip ≥0.65 in. (GB)]
  4) GFS convective precip (GB form with most precip thresholds 0.10 in and greater)
  5) (Seasonal RF precip ≥0.50 in.) × [GFS tot precip ≥0.40 in. (GB)]
  6) (Seasonal RF precip ≥0.10 in.) × [GFS tot precip ≥0.10 in. (GB)]
  7) [GFS 700-mb vert. vel. (upward)] × [GFS K index ≥30.0 K (GB)]
  8) GFS precipitable water (most GB thresholds 1.4 in. and greater)
  9) [GFS 700-mb vert. vel. (upward)] × (GFS precipitable water)

and the frequency of selection across multiple equations. With forward-selection screening, the earlier a predictor is selected the greater its contribution to the reduction of predictand variance (the higher its rank). Also, the higher the frequency of selection of a given predictor for an equation group, the greater its predictive contribution. For a given equation group, the ranking was performed for the top nine predictors; the arbitrary nine-predictor cutoff is about the same as the average number of predictors in a POP equation.

The ranked predictor sets showed marked contrasts across the two predictand types (POP versus CPQPF), geographical regions, and forecast projections. For example, Table 3 shows the ranked predictors for 6-h POP and CPQPF equations for the 1200 UTC cycle during the cool season, all geographical regions, and four day 1 forecast projections (12, 18, 24, and 30 h). Note that

highly ranked predictors for POP (Table 3a) are variables with information about whether precipitation will occur (GFS total precipitation together with precipitation climatology for *light thresholds*, layer mean relative humidity in the lower atmosphere, and 700-mb vertical velocity). The corresponding CPQPF predictors (Table 3b) are slanted more toward the amount of precipitation and the convective nature of heavy amounts [GFS total/convective precipitation and precipitation climatology for *thresholds of 0.10 in. and greater*, 700-mb vertical velocity together with the K stability index, and precipitable water for moderate and large grid binary (GB) thresholds]. These predictor contrasts are consistent with improved forecast skill with the CND vis-à-vis the UC PQPF approach shown in section 5a.

As an example of predictor variations across diverse geographical regions, Table 4 shows the ranked day 1 cool season CPQPF predictors for region 1, which has rugged mountain terrain, and region 14, which has flat terrain (Figs. 3 and 7). Note that strong predictors in region 1 (Table 4a) include precipitation climatology, terrain-forced vertical velocity, and terrain elevation. Contrasting predictors for region 14 (Table 4b) include GFS forecasts of convective precipitation and several convective stability indices, which is consistent with the subtropical climate of this region. Other unique predictors for this region are GFS forecasts of sea level pressure and the u-wind component. Finally, note that the predictor uniqueness for these regions is mostly obscured in the corresponding predictor set summarized over the entire CONUS domain (Table 3b).

The impacts of regional predictor variations on PQPF skill are indicated in Fig. 11. Here, the BSS for regionalized regression equations (REG) is compared with the BSS for nonregionalized equations (NON-REG) for four day 1 forecast projections over the 2008/09 cool season. The t test indicated that the CONUS skill superiority of REG over NON-REG (Fig. 11a) is statistically significant with 99% confidence for all precipitation thresholds in the 0.01–0.50 range and with 90% confidence for ≥1.00 in. (the tiny BSS increase for ≥2.00 is not significant). Figure 11b shows the corresponding BSSs for individual regions at the ≥0.25-in. threshold. Here, we see that the largest skill improvements for REG are for regions 1–5 in the mountainous western United States (Figs. 3 and 7); in the east, where terrain effects are small or nonexistent, the improvement of REG on NON-REG is small.

d. Impact of the topoclimatic interactive predictors

In section 3e we described attractive properties of the TCIA predictor variables and in the previous subsection we saw that these predictors had high CONUS-wide

TABLE 4. As in Table 3b, but for two contrasting regions.

(a) Region 1
  1) (Seasonal RF precip ≥0.50 in.) × [GFS tot precip ≥0.40 in. (GB)]
  2) GFS tot precip (GB form with most precip thresholds 0.10 in. and greater)
  3) [TVV (upward; GFS 850-mb u, v comp.)] × [GFS tot precip ≥0.04 in.(GB)]
  4) (Seasonal RF precip ≥0.75 in.) × [GFS tot precip ≥0.65 in. (GB)]
  5) [TVV (upward; GFS 10-m u, v comp.)] × (GFS 1000–500-mb mean RH)
  6) (Seasonal RF precip ≥0.25 in.) × [GFS tot precip ≥0.25 in. (GB)]
  7) [GFS 700-mb vert. vel. (upward)] × [GFS 850-mb moisture divergence (neg.)]
  8) (Seasonal RF precip ≥0.10 in.) × [GFS tot precip ≥0.10 in. (GB)]
  9) Terrain elevation
(b) Region 14
  1) GFS tot precip (GB form with most precip thresholds 0.10 in. and greater)
  2) GFS convective precip (GB form with most precip thresholds 0.10 in. and greater)
  3) [GFS 700-mb vert. vel. (upward)] × [GFS K index ≥30.0 K (GB)]
  4) (Monthly RF precip ≥0.50 in.) × [GFS convective precip ≥0.16 in. (GB)]
  5) GFS mean SLP
  6) GFS best lifted index
  7) (Seasonal RF precip ≥0.10 in.) × [GFS tot precip ≥0.10 in. (GB)]
  8) GFS u-wind components (850- and 500-mb)
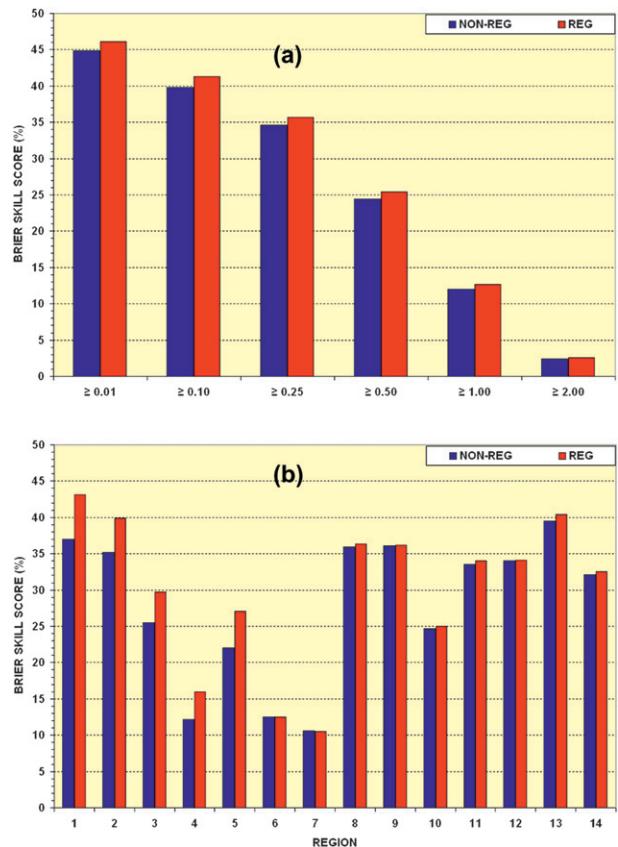 9) GFS total totals index

FIG. 11. BSSs (%) of 6-h PQPFs with NON-REG and REG regression equations for (a) the CONUS domain for multiple precipitation thresholds (in.) and (b) the individual regions for the ≥0.25-in. threshold. The scores are averaged over four day 1 forecast projections (12, 18, 24, and 30 h) from the 1200 UTC cycle for the 2008/09 cool season.

ranking, especially for the CPQPF equations. A test was conducted to quantify the impact of these predictors by withholding them in ''NO TCIA'' regression equations and then comparing the resulting PQPF skill against the skill with ''TCIA'' equations. Since the component variables of the TCIA predictors were retained in the candidate predictor lists for the NO TCIA equations, the experiment was designed to specifically quantify the contribution of the interactive effects on forecast performance.

Figure 12 shows a case of 36-h PQPFs for the ≥0.25-in. threshold with the NO TCIA and TCIA equations for the 6-h period ending 0000 UTC 5 January 2008. The maps show that the TCIA predictors yielded higher peak probabilities and improved spatial resolution in this case. Figure 13a shows the corresponding comparative BSS averaged for the day 2 period (36-, 42-, 48-, and 54-h forecast projections) over the CONUS domain for the 2007/08 cool season. In this chart the BSSs for the TCIA equations are significantly higher (at the 99% confidence level) than those for the NO TCIA equations for all precipitation thresholds. The strong contribution of the TCIA variables to forecast skill evident here is

consistent with a demonstration in section 3e of their powerful predictive effects.[9]

The strong predictive role of the TCIA predictors shown so far applies to the shorter forecast projections (through day 2). Examination of the ranked POP and CPQPF predictors as a function of increasing projection reveals that the contribution of the TCIA predictors gradually decreases with time. For example, Table 5 shows that the contribution of these predictors for day 5 (108-, 114-, 120-, and 126-h projections) was clearly lower than for day 1 (Table 3b), and, instead, simple predictors (GFS total/convective precipitation and v-wind

---

[9] A ''similar test'' of the topoclimatic predictors reported in CS (Figs. 10 and 11 of that article) contains an important distinction from the test conducted here; all topoclimatic predictors were withheld in that test, whereas here only those that were interacted with the GFS total precipitation forecast were withheld. This difference accounts for the higher impact of these predictors in CS.
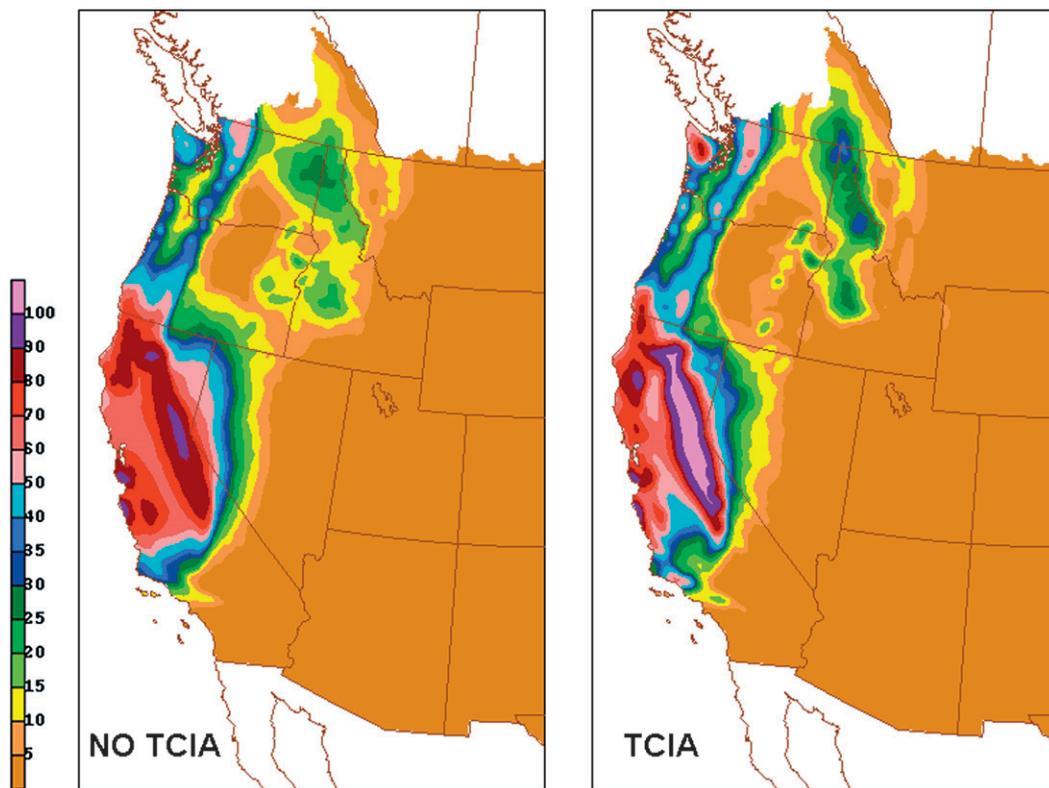
FIG. 12. The 36-h probability (%) of ≥ 0.25 in. for the 6-h period ending 0000 UTC 5 Jan 2008, where topoclimatic interactive predictors were excluded (NO TCIA) and included (TCIA).

components) assume an increased presence. Also, "stand alone" seasonal and monthly precipitation climatology variables appear for the first time in the day 5 list. Further evidence of the reduced role of the TCIA variables for day 5 is shown in Fig. 13b; the skill differences between the TCIA and NO TCIA PQPFs are noticeably smaller than for day 2 (Fig. 13a). Still, the top ranking of the GFS total and convective precipitation forecasts is maintained to day 5 (Table 5) despite a presumed increase in GFS forecast error with time. If the primary predictive effect of the TCIA variables is to account for the small-scale bias error in the GFS precipitation forecasts (as suggested in section 3e), these day 5 findings suggest that this mechanism is relatively ineffective as GFS forecast error increases.

e. Experiments with logistic regression

The multiple linear regression method is used virtually exclusively at MDL to produce a broad suite of centralized MOS weather guidance products for use at the NWS (Glahn and Lowry 1972; Carter et al. 1989; Glahn et al. 2009). The long history of success with linear regression is owed to its mathematical simplicity and effectiveness in producing useful forecast guidance

for operational settings. Its low computational cost allows for a large number of diverse potential predictors, including those formulated to account for nonlinear predictand–predictor relationships (see discussion of such HRMOS predictors in sections 3e, 5b, and 5d).

Nonlinear statistical models inherently account for predictand–predictor nonlinearities, and especially where the applications involve dichotomous (e.g., binary) predictands, they should provide a better fit of the predictand–predictor data than linear regression. The logistic regression method may be especially well suited for binary predictand events with 0 and 1 values as the nonlinear logistic function is constrained to the 0 to 1 range (Wilks 2006). In a PQPF application involving 24-h precipitation for light thresholds (≥ 0.01, ≥ 0.05, and ≥ 0.10 in.), Applequist et al. (2002) found that logistic forecast skill was significantly better than that for linear regression. High-quality 24-h PQPFs with the logistic method have subsequently been obtained by Hamill and Whitaker (2006), Sloughter et al. (2007), and Wilks and Hamill (2007), among others.

A significant drawback with the logistic method is that the constant and regression coefficients must be approximated with an iterative method, which involves
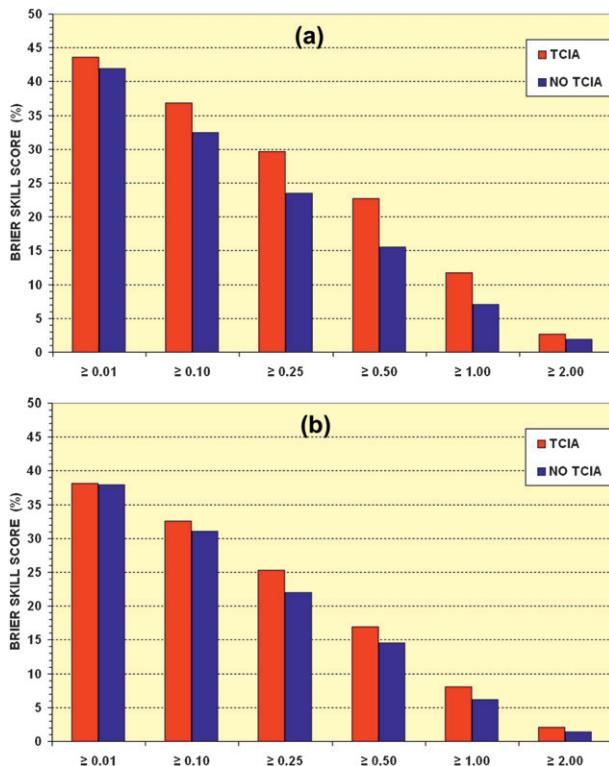
FIG. 13. BSSs (%) of 6-h PQPFs [for multiple precipitation thresholds (in.)] with and without topoclimatic interactive predictors (TCIA and NO TCIA, respectively; see text) for the CONUS domain averaged over (a) four day 2 projections (36, 42, 48, and 54 h) and (b) four day 5 projections (108, 114, 120, and 126 h) from the 1200 UTC cycle for the 2007/08 cool season.

TABLE 5. As in Table 3b, except for day 5.

1) GFS tot precip (GB form with most precip thresholds 0.10 in. and greater)
2) GFS convective precip (GB form with most precip thresholds 0.10 in. and greater)
3) GFS precipitable water (most GB thresholds 1.4 in. and greater)
4) V-wind components (850 and 500 mb)
5) (Seasonal RF precip ≥0.25 in.) × [GFS tot precip ≥0.25 in. (GB)]
6) (Seasonal RF precip ≥0.75 in.) × [GFS tot precip ≥0.65 in. (GB)]
7) Seasonal RF precip (mostly ≥0.25 in. and larger thresholds)
8) Monthly RF precip (all thresholds ≥0.25 in. and larger)
9) GFS 850-mb moisture flux divergence

substantial additional computational cost. With HRMOS PQPF, computational cost is an important consideration as this application involves very large developmental samples (up to 100 million cases) and many candidate predictors (around 150). Thus, applying the logistic regression involves trade-offs between potentially improved forecast performance and increased development cost. Here, we discuss logistic versus linear regression tests to address aspects of two questions: 1) To what degree is HRMOS PQPF performance improved by using logistic regression? 2) To what degree is it feasible to apply logistic regression in the HRMOS model context?

The logistic program used is "freeware" (available online at http://jblevins.org/mirror/amiller/#logit), where predictors to be fitted must be supplied (there is no predictor screening capability as there is with the linear regression program). We could use up to 84 million cases (composing 6.5 cool seasons) and four predictors, beyond which the logistic program failed (probably because available memory was exceeded) on the NCEP CCS. Since these restrictions do not apply to the MDL linear regression program, eight full cool seasons and all candidate predictors were screened to develop the comparative linear regression equations.

Because logistic regression tests with fewer than four predictors yielded lower BSSs (on an independent sample) than those tests with four (not shown), four predictors were used in comparative tests against linear regression. The four predictors in the POP and CPQPF logistic (LOG4) equations were determined by maximizing PQPF skill on an independent sample based on various candidate four-predictor sets. The variables in the candidate sets were selected, largely by physical reasoning, from those in Tables 1 and 2. Also considered were continuous forms of the GFS total and convective precipitation as well as the cube root transformation of each, as such model precipitation variables were used in previous logistic PQPF applications (Hamill and Whitaker 2006; Sloughter et al. 2007; Wilks and Hamill 2007; Hamill et al. 2008). The 36-h POP and CPQPF predictors that produced the highest BSS on an independent sample (and are used in the comparative tests against linear regression) are listed in Table 6. Note that these predictors appear among highly ranked day 1 predictors in the HRMOS linear regression equations (see Table 3).

Two sets of linear regression equations were matched with the LOG4 equations. One set consisted of "four predictor" test equations (LIN4), where a maximum of four predictors were screened from the full candidate lists (Tables 1 and 2). The second set consisted of the HRMOS model equations (LINOP), where the number of predictors was optimized as described in section 5c.

Figure 14 shows BSSs for unconditional PQPFs with LOG4, LIN4, and LINOP for four day 2 forecast projections (36, 42, 48, and 54 h) from 1200 UTC during the full 2009/10 cool season. Note that the BSSs for LINOP were slightly higher than those for both LIN4 and LOG4 across all thresholds (the LINOP improvement on LOG4 is significant at the 95% level for all thresholds but ≥2.00 in.). Also, the BSSs for LOG4 and LIN4 are

TABLE 6. Four POP and CPQPF predictors in cool season logistic regression equations that yielded the best (unconditional) PQPF skill (see text) over multiple precipitation thresholds at the 36-h forecast projection from 1200 UTC. All references to precipitation apply to 6-h periods, and nonstandard abbreviations are as in Tables 1 and 2.

POP
  GFS tot precip $\geq$0.01 in. (GB)
  GFS mean RH (lower atmosphere layer)
  [TVV (upward; GFS 10-m u, v comp.)] $\times$ (GFS 1000–500-mb
    mean RH)
  (Seasonal RF precip $\geq$0.10 in.) $\times$ [GFS tot precip $\geq$0.10 in. (GB)]
CPQPF
  GFS tot precip
  GFS convective precip
  [GFS 700-mb vert. vel. (upward)] $\times$ (GFS precipitable water)
  (Seasonal RF precip $\geq$0.25 in.) $\times$ [GFS tot precip $\geq$0.25 in.
    (GB)]

quite similar to one another for all but the $\geq$0.01-in. threshold where the LOG4 BSS is notably lower.

The corresponding PQPF reliability and sharpness for the three methods are shown for the $\geq$0.01- and $\geq$1.00-in. thresholds in Fig. 15. For $\geq$0.01 in., the LIN4 and LINOP probabilities are slightly more reliable than those for LOG4 [the plotted points are closer to the perfect reliability (diagonal) line]. For LOG4, the pronounced overforecasting (forecast probabilities higher than observed relative frequencies) together with the high frequency of issuance of probabilities in the 5%–30% range may be responsible for the low relative skill at this threshold. For $\geq$1.00 in., the reliability properties of LINOP and LOG4 are mostly similar to one another, except LOG4 again shows notable overforecasting of low probabilities. Also, LIN4 shows pronounced underforecasting and a more limited probability range than LOG4.

The comparative PQPF sharpness for the three models is indicated in the probability histograms in Fig. 15. The most notable differences between LOG4 and the linear methods are that LOG4 has lower frequencies of probabilities across the central span (15%–75%) of the full probability range at $\geq$0.01 in. and higher frequencies of upper probabilities at $\geq$1.00 in. As the frequency differences are mostly small otherwise, this suggests slightly better sharpness with the logistic method. Finally, between LIN4 and LINOP evidence of improved sharpness for $\geq$1.00 in. with the latter is the finding of higher frequencies of probabilities above 35%.

To summarize, this limited test shows that linear regression equations with an optimal number of predictors performed somewhat better than logistic equations with four predetermined predictors. Considering the enhanced mathematical appropriateness of logistic over linear regression for the PQPF predictand and the superior PQPF
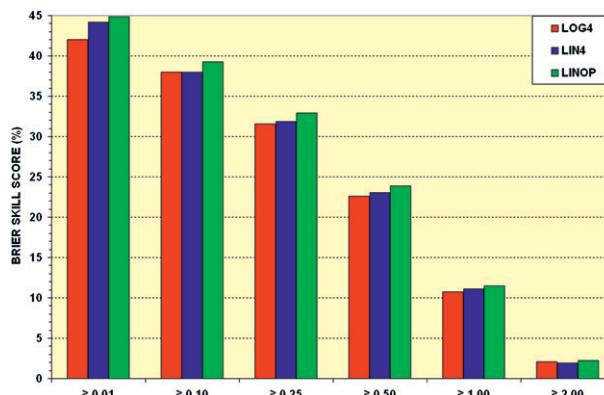

FIG. 14. CONUS BSS vs precipitation threshold (in.) for 6-h PQPFs averaged over four day 2 forecast projections from 1200 UTC during the 2009/10 cool season. The legend notations denote PQPF regression equations, where LOG4 is for the logistic method with four predetermined predictors (see text), LIN4 is for screening linear regression with a maximum of four predictors, and LINOP is for screening linear regression with an optimal number of predictors. The precipitation thresholds are in units of in.

performance with logistic regression reported by Applequist et al. (2002) (see previous citation), this (perhaps unexpected) finding is likely due to impediments encountered with the logistic method. Most notably, the predictor number and specific predictors in the logistic equations were not optimized to the degree applicable to the linear regression equations in part because the logistic program did not have a predictor screening capability. The essence of this limitation is the substantial computational cost associated with predictor screening, which is far less with linear regression. Another constraint with the logistic method is the increased computer memory requirement, which necessitated reducing the developmental sample from 8 seasons with linear regression to 6.5 seasons.

A factor that may have reduced the potential advantage of the nonlinear attribute of the logistic method in the test conducted here is the use of techniques to build nonlinear relationships into the linear regression equations. Recall that two such techniques were used: 1) the use of interactive predictors discussed in sections 3e, 5c, and 5d, and 2) the use of the conditional predictand approach discussed in section 5b. Recall, also, that each of these techniques improved PQPF skill in the linear regression equations.

6. Derived QPF elements[10]

Presently, the sole QPF element produced in NDFD is a continuously varying 6-h precipitation amount forecast,

---

[10] The generic term QPF element refers to one of three forecast precipitation quantities derived from the PQPFs (the basic QPF element). The derived QPF elements are described in this section.
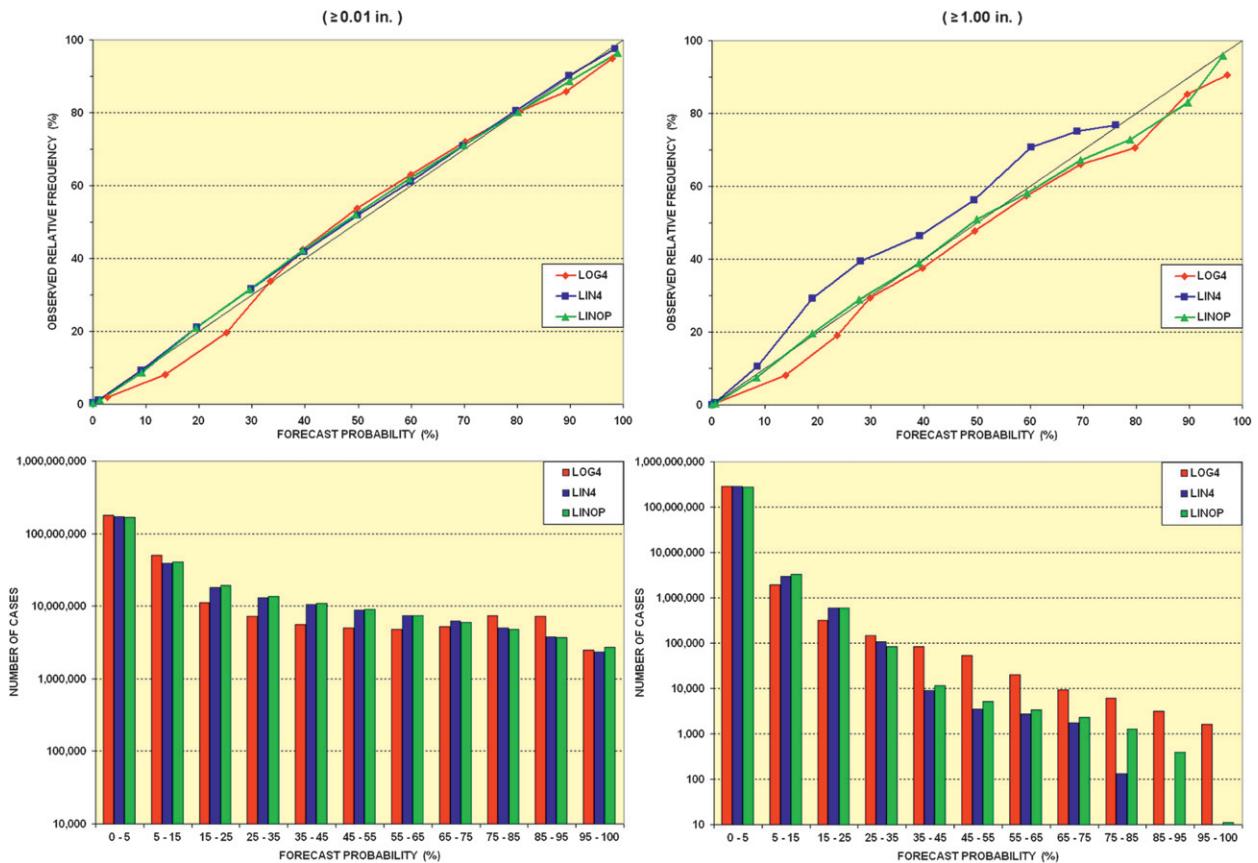
FIG. 15. (top) PQPF reliability diagrams and (bottom) probability histograms (with logarithmic ordinate scale) corresponding to Fig. 14 for the (left) ≥0.01- and (right) ≥1.00-in. thresholds. Forecast probabilities in the reliability diagrams are mean values for intervals shown in the histograms.

which is also the QPF element currently ingested into RFC streamflow prediction models. A similar HRMOS QPF element was derived from the 6-h PQPFs. The derivation consists of two steps, the first of which involves computing a categorical QPF element; in the second step the continuous QPF element is computed from the latter.

a. Categorical QPF

The formulation of the categorical QPF consists of three tasks. First, we compute a threshold probability constant, which is separate for each region, precipitation threshold, and forecast projection. The second step consists of checking whether or not a PQPF value at a grid point equaled or exceeded the corresponding threshold value and then setting a "yes" or "no" flag accordingly. When the test is performed over all thresholds, it yields a set of eight yes–no flags for the point. In the third step, a categorical QPF quantity, commonly called the best precipitation category (BC), is specified from the flag set. Important details concerning these steps are included in appendix A.

b. Continuous QPF

The derivation of the HRMOS continuous precipitation (CP) element from BC involves "interpolating" between the (discontinuous) precipitation categories and "extrapolating" beyond the unbounded 2.0-in. peak category. Details of the computation appear in appendix B.[11] Not surprisingly, raw CP fields exhibited undesirable stair-stepping as a result of interpolating into the discontinuous BC field. To treat the problem, a specially formulated smoothing scheme was devised, a key aspect of which involves incorporation of yet another derived HRMOS QPF element: expected precipitation.

---

[11] The techniques for computing categorical and continuous QPFs from the PQPFs are very efficient computationally, which is important considering the large data volume at hand. Despite the simplicity of the methods, in section 7c we show that the derived QPF elements perform at a high level relative to other model QPF products. Thus, more elaborate QPF modeling methods, such as those involving ensemble calibration (e.g., Hamill and Colucci 1998) and Bayesian model averaging (Sloughter et al. 2007), were not considered in part out of concern over high computation cost.

c. Expected precipitation

Given the probability p_i for precipitation exceedance threshold i (PQPF$_i$) at a grid point, the probability p_{i+1} for the next higher exceedance threshold, the conditional climatic mean precipitation μ_i between threshold i and $i + 1$, and p_N (and μ_N) are corresponding quantities for the peak exceedance threshold, the expected precipitation (EP) is given by

$$EP = \sum_{i=1}^{N-1} \mu_i(p_i - p_{i+1}) + \mu_N p_N \qquad (3)$$

(Wilks 2006), where $i = 1$ corresponds to ≥ 0.01 in. (in the case of p) and $i = N$ corresponds to ≥ 2.00 in. The μ's (constants) were computed by combining precipitation data from the entire developmental dataset, as earlier tests indicated that the variability of the μ's as a function of geography, season, and time of day is quite small. Note from (3) that an appropriate alternate term for EP is probability-weighted average precipitation [Wilks (2006, p. 82)].

Since both EP and CP are continuously varying precipitation quantities, it is important to note the distinctive properties of each element. Note that EP is computed from forecast probabilities for multiple threshold precipitation amounts. Since the probabilities for the heavy thresholds are generally quite low, EP will strongly underforecast heavy precipitation events. Also, EP will greatly overforecast very light precipitation amounts because (3) yields a nonzero value whenever the probability (however small) of any precipitation threshold is nonzero. Because of these properties, it quite unlikely a given EP field will closely resemble the corresponding observed precipitation field.

Underlying the definition of CP, on the other hand, is a conversion from the probability of a precipitation threshold to a corresponding categorical forecast, which involves applying a threshold probability (section 6b). Since the threshold probabilities for the various precipitation thresholds are specified to yield an appropriate bias for the categorical forecasts, a given CP field may closely resemble the corresponding observed precipitation field. Still, an inherent weakness of CP is the abrupt change in its value where the probability for a given precipitation threshold falls just below or just meets the threshold probability value.

The application of conventional grid smoothing to treat stair-stepping in the raw CP fields (noted in section 6b) was not fruitful, as the level of required smoothing resulted in an unacceptable level of erosion of valid CP detail. In search of an alternative, we noted that since EP fields have characteristics of highly smoothed CP fields, the former

could serve as a control for "smoothing" the latter. With this premise, a smoothing operator was defined as

$$CP_i = EP_i f_i \quad \text{and}$$

$$f_i = \frac{\sum_{k=1}^{N} CP_{k,i}}{\sum_{k=1}^{N} EP_{k,i}}, \qquad (4)$$

where the smoothed CP value at point i (CP$_i$) is obtained by scaling EP$_i$ with the factor f_i. The summation in (4) is over all grid points N within a circle of about 20-km radius, centered on point i, and f_i is set to 0.0, where CP$_i$ = 0.0. This simple operator adequately mitigated the stair-step artifact, as indicated by the CP examples in the following section.

7. Properties and performance of the basic and derived QPF elements

In this section, predictive properties of the HRMOS QPF elements are described for two selected heavy precipitation cases, one for the cool season and the other for the warm season. Also, the forecast performance of the CP element is examined through limited comparative scoring. The aim is to provide a concise assessment of the quality of the QPF elements.

a. Cool season case

Figure 16 shows the observed 6-h precipitation and 60-h QPF elements valid at 0000 UTC 5 January 2008. The observed precipitation field features an extensive area of precipitation inland of the U.S. West Coast, where relatively high amounts are focused on the western (windward) slopes of the Coastal Mountains and Sierra Nevada in California with peak amounts in the 2–4-in. range. Note that the finescale distribution of observed precipitation matches up well with the finescale patterns in the four HRMOS QPF elements (Figs. 16b–e). Thus, this case exemplifies the high-spatial-resolution property of the QPFs, which is most prominent in the mountainous western United States. The enhanced spatial detail in the West, which also appears for warm season cases (not shown), reflects the predictive effectiveness of the TCIA predictors (see Figs. 12 and 13). Note from Fig. 11b that cool season PQPF skill for the two West Coast regions equals or exceeds the highest skill anywhere over the CONUS.

Close inspection of the probability map of 6-h precipitation ≥ 0.50 in. (Fig. 16b) yields several findings. One feature is the strong probability sharpness, as the
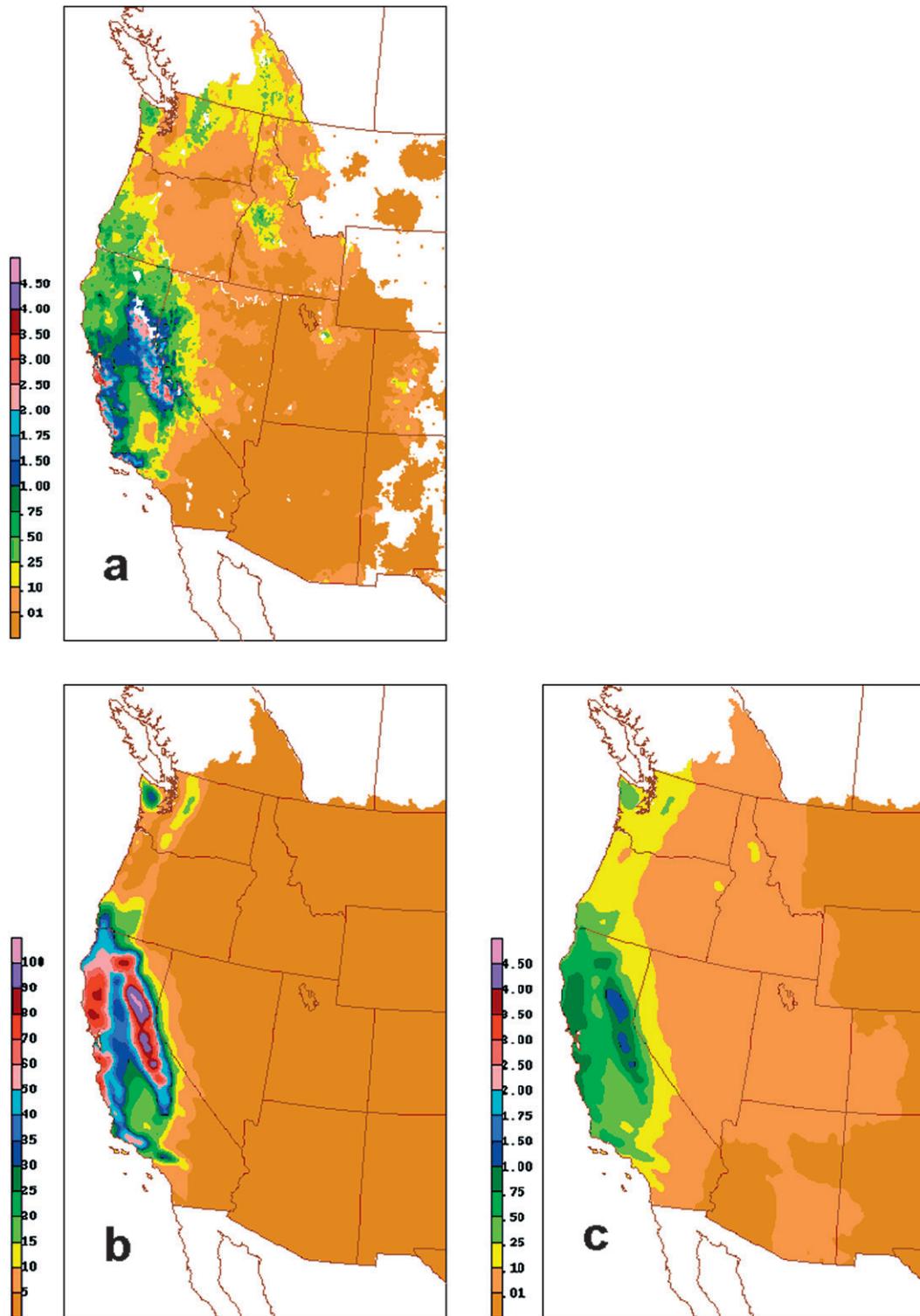
FIG. 16. (a) QCed stage IV precipitation analysis (in.; white areas indicate missing values) and 60-h QPF elements for the 6-h period ending 0000 UTC 5 Jan 2008 (cool season case). The QPF elements are (b) probability (%) of ≥0.50 in., and (c) EP, (d) BC, and (e) CP, all in units of in. Note that the units in (a),(c),(d), and (e) are identical, but the color bar in (d) is unique.
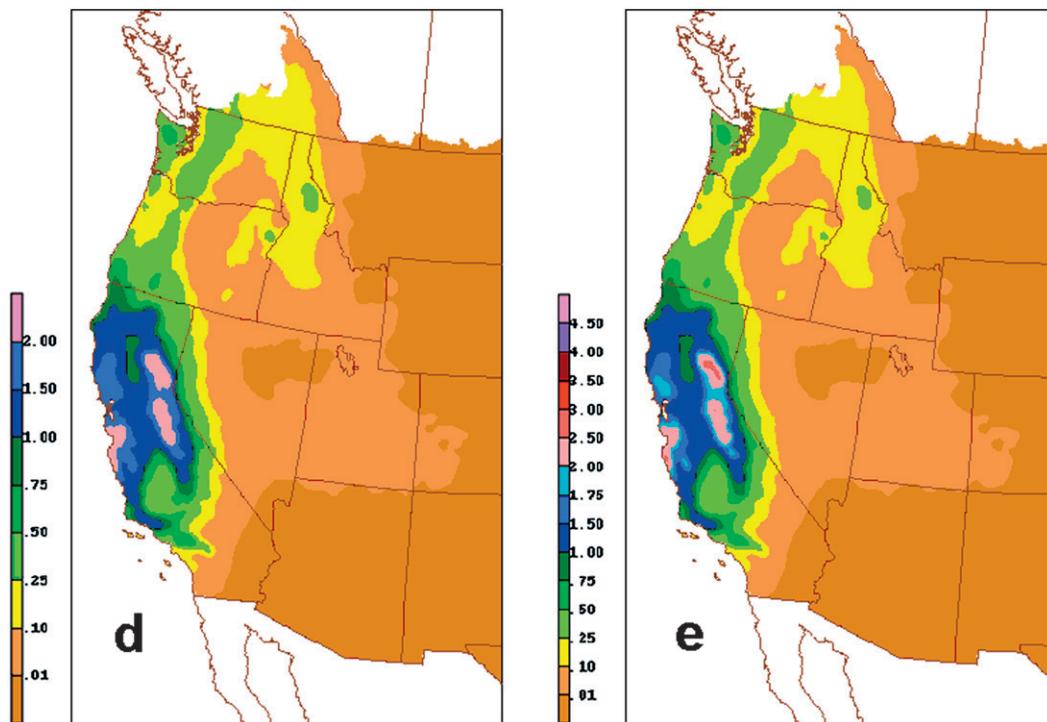
FIG. 16. (*Continued*)

full range of probabilities (0%–100%) is present. Note also that the coverage of ≥0.50-in. probabilities of about 20% and higher closely matches the area where such precipitation amounts were observed (Figs. 16a and 16b). Also, the spatial pattern of the probabilities closely matches the patterns of the upper values of the EP, BC, and CP fields (Figs. 16c, 16d, and 16e, respectively), which reflects the derivation of these elements from the probabilities. Finally, the finding of the expanded geographical coverage of the light precipitation in these derived QPF elements reflects the contribution of probabilities for precipitation thresholds below ≥0.50 in (not shown).

Careful inspection of the derived BC and CP elements (Figs. 16d and 16e) reveals both common and unique features between them. The strong similarity of the BC and CP patterns stands out, which is expected since CP is essentially a continuous precipitation replica of the discontinuous BC field (appendixes A and B). Note that the BC field contains the full range of categorical precipitation amounts (Fig. 16d), and CP values vary continuously from 0.0 to a peak of about 2.5 in. Recall that CP values above the BC peak precipitation threshold of ≥2.00 in. are extrapolated (section 6b); here, the peak extrapolated values are only slightly greater than this peak threshold.

The EP field (Fig. 16c) is similar to CP, as it also consists of continuously varying precipitation amounts; however, EP contains distinguishing properties. Specifically, note that peak EP values are much lower than peak CP values, as EP incorporates the uncertainty associated with the CP amounts (section 6c). Note also that the geographical coverage of nonzero EP values extends beyond the coverage of nonzero CP values, which reflects the precipitation occurrence uncertainty near the zero lines of the latter. Thus, the forecaster-guidance utility of EP should complement that for CP, as the former provides warning for low threat precipitation events. Here, the extension of small EP values eastward of corresponding CP values provides a warning of the scattered light observed precipitation from western Montana to western Colorado to southeastern Arizona. Finally, note that an additional potential practical benefit of EP is that, since it combines the contributions of the PQPFs for multiple precipitation thresholds into one map, it could alleviate the map examination workload for time-pressed line forecasters.

b. Warm season case

The selected warm season case is for the 6-h period ending 1200 UTC 13 September 2008, as Hurricane Ike was coming ashore over Galveston Bay (Texas Gulf coast). As before, Fig. 17 contains maps of the observed 6-h precipitation and a 48-h prediction of each QPF element. The observed precipitation field (Fig. 17a) shows a small area of extremely heavy precipitation in the Galveston Bay area of Texas, where many values exceeded 4.5 in. and the peak value was 10.8 in. In addition,
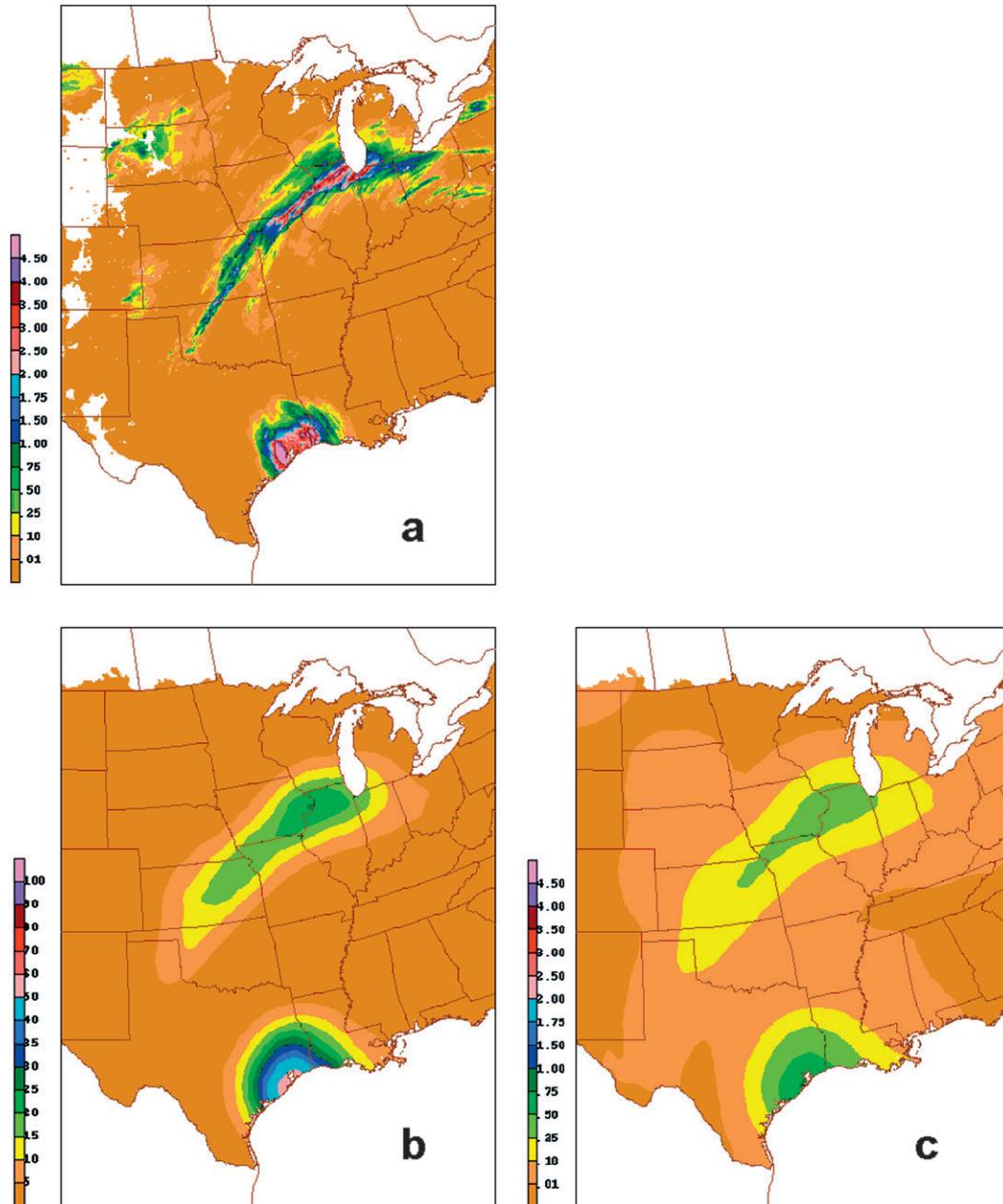
FIG. 17. As in Fig. 16, but for the 6-h period ending 1200 UTC 13 Sep 2008 (warm season case), where the QPF elements are for the 48-h projection from 1200 UTC 11 Sep 2008.

a narrow swath of heavy precipitation from Oklahoma to Ohio also contains small peaks with over 4.5 in.

Each of the 6-h QPF elements (Figs. 17b–e) accurately predicted the locations of the two heavy precipitation areas. Further, CP (Fig. 17e) accurately predicted the intensity of the extremely heavy precipitation around Galveston Bay, which provides a striking example of the high precipitation intensity resolution attribute of the HRMOS model. On the other hand, CP did not accurately predict the intensity of the similarly heavy

precipitation from northeast Missouri to northwest Indiana. Since the BC field (Fig. 17d) contains areas of the peak ≥2.00-in. precipitation threshold for both precipitation maxima, the difference in peak CP values between them results from contrasting extrapolated precipitation amounts (appendix B).

It is also noteworthy that the forecast probabilities are lower for this warm season event than for the previous cool season case despite the occurrence of much higher precipitation amounts. Note that peak ≥0.50-in.
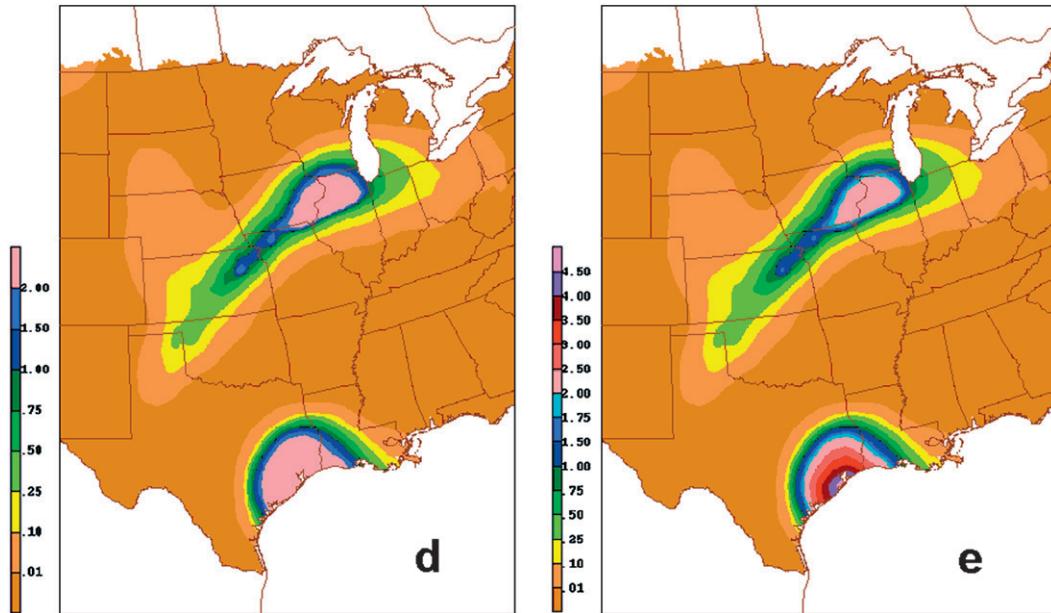
FIG. 17. (*Continued*)

probability values in Fig. 17b are just above 50%, while they reach 100% for the cool season case (Fig. 16b). Also, the comparatively low EP values in Fig. 17c indicate that the probabilities are lower across all precipitation thresholds. The lower probabilities here are probably due to relatively weak synoptic-scale forcing and the absence of resolvable localized precipitation forcing. The very broad areal coverage of small EP values in Fig. 17c is consistent with this assessment. In contrast, the QPF predictability for the western cool season case undoubtedly benefited from the combination of strong localized topographic forcing and strong synoptic-scale forcing, both of which are well represented in the HRMOS predictor database.

c. Comparative scoring of CP

Here, the forecast performance of the 6-h CP element is examined by comparatively scoring it with two operational model QPFs: the GFS (also used as an HRMOS QPF predictor) and GMOS. In this limited forecast performance assessment, each model QPF was "categorized" into a set of multiple binary variables, where a binary variable denotes the predicted occurrence (value of 1) or nonoccurrence (value of 0) of a 6-h precipitation exceedance threshold. The scoring was conducted on the 4-km HRAP grid where the QCed QPE data were used for validation. This necessitated interpolating the GFS and GMOS QPFs to the HRAP grid. Since the GFS QPF was obtained from an MDL archive with an 80-km grid, two interpolations relative to the GFS native grid are involved in the rendering to the HRAP grid. The GMOS

QPF was interpolated from the 5-km NDFD grid. All interpolations were performed with a precipitation-area-preserving method (National Weather Service 1974). Note that CP did not involve grid interpolation, and the verifying data were the same as those used in HRMOS model development. Since it could be argued that the scoring framework favors CP, the performance assessment should be regarded as preliminary.

The CONUS threat score [TS; same as the critical success index (Schaefer 1990)] and bias (where unbiased forecasts have a 1.0 bias) for the categorized GFS, GMOS, and CP (denoted HRMOS here) QPFs are shown in Fig. 18 for the day 1 and day 3 forecast periods during the cool and warm seasons. Figures 18a and 18b show that the TSs for HRMOS are clearly higher (indicating better forecast accuracy) than those for GFS and GMOS; the charts also show that the HRMOS TS improvement increases with increasing precipitation threshold. (All HRMOS TS improvements in Figs. 18a and 18b have high statistical significance.) Figures 18c and 18d show that the HRMOS bias was just above 1.0 over all thresholds; HRMOS bias scores for other forecast projections (not shown) are about the same as these for all forecast projections to 156 h. Contrastingly, the GFS and GMOS QPFs show poor bias; the GFS overforecasting of light precipitation thresholds and underforecasting of heavy thresholds is quite pronounced.

Several comments concerning the scores in Fig. 18 are noteworthy. 1) The superior scores for HRMOS are likely due primarily to TCIA predictors and secondarily to geographical regionalization, which were used to
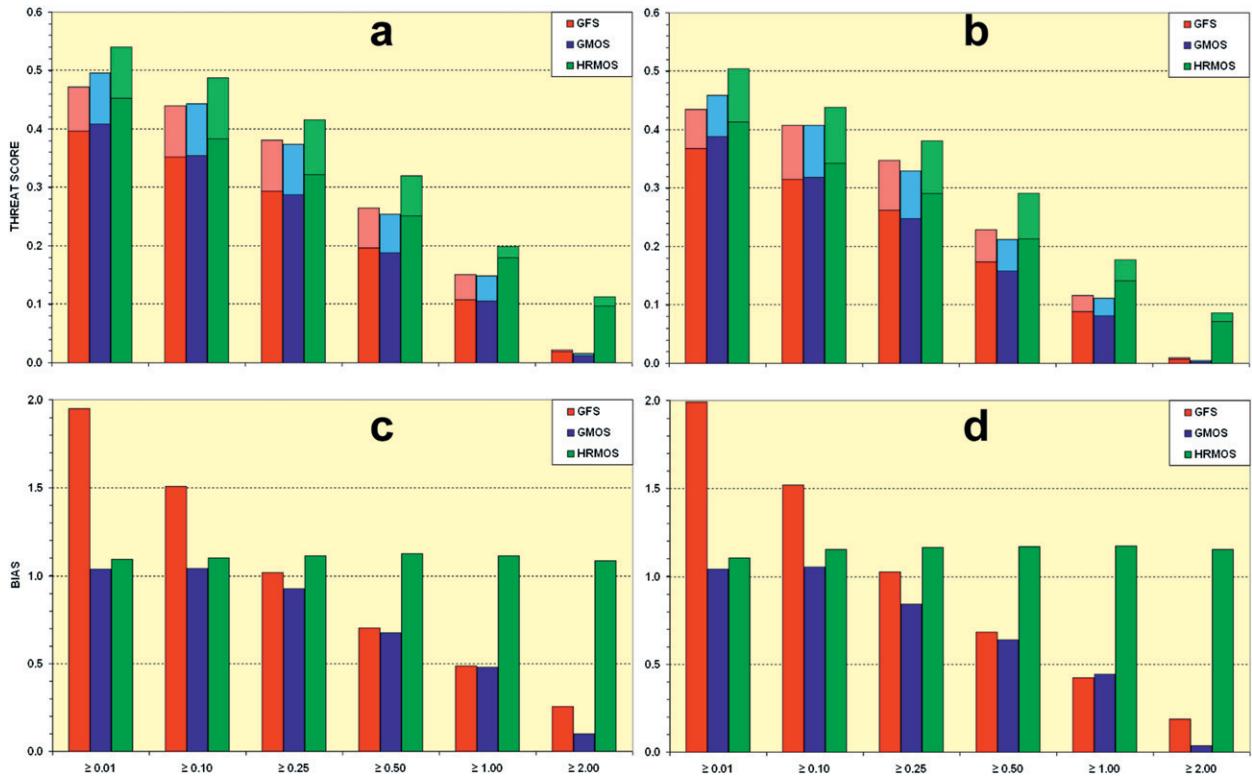
FIG. 18. (a),(b) CONUS threat score and (c),(d) bias versus precipitation threshold (in.) for three model QPFs. The threat score in (a) is for day 1 (light shades) and day 3 (dark shades) for the cool season (1 Oct 2008–31 Mar 2009); (b) as in (a), but for the warm season (1 Apr–30 Sep 2009). The biases in (c) and (d) are for the cool season and warm season, respectively, and they are shown only for day 3 (day 1 bias scores were similar).

obtain high-quality PQPFs from which the CP element is derived. An additional contribution of regionalization may result from its use in each of the two postprocessing steps used to derive CP from the PQPFs. 2) The non-improvement of GMOS on the GFS raises a concern because the station-oriented QPF from which the corresponding GMOS element was derived (Glahn et al. 2009) uses the GFS as the driving model. 3) It is not known to what degree the double interpolation of the GFS QPF may have contributed to its poor bias scores. This aspect of the verification also deserves investigation. 4) The small seasonal differences in TS for all products are somewhat surprising, as larger warm season degradations in TS have been found in other studies (e.g., Olson et al. 1995). However, the unusually wet summer during 2009 over the central and northern plains may have contributed to relatively good warm season scores seen here (thus mitigating the usual warm season TS degradation), as a positive correlation between seasonal wetness and good QPF performance scores was reported in Olson et al. (1995).

8. Summary and comments

The HRMOS QPF model incorporates a number of new features into a fine-grid, GFS-based MOS application to produce QPF elements with enhanced spatial and intensity resolutions. Enhanced spatial resolution stems from the inclusion of multiple finescale precipitation and lightning climatologies along with detailed topography in the predictor database. Simple product variables, which combine finescale static information in topography and climatology data with large-scale dynamical information in GFS model QPFs, are shown to be quite effective predictors in the PQPF regression equations. Tests indicate the interactive property of these "topo-climatic interactive predictors" yields significantly increased PQPF skill; geographical regionalization of the PQPF linear regression equations and the derived QPF elements also contributes to enhanced forecast performance.

Limited tests with logistic versus linear regression to produce the 6-h PQPFs indicated that the use of the logistic method is feasible within the HRMOS model context. Also, the logistic method yielded comparable PQPF performance when a limited number of predictors were used with each method. However, the PQPF skill with linear regression was slightly higher than that with logistic regression, as many predictors selected with an objective screening procedure were used with the linear

regression method. A similar predictor screening capability was not available with the logistic method, which is a computational cost limitation.

High precipitation intensity resolution in the QPF elements is due fundamentally to the use of radar-based (gauge-based in the western United States) fine-grid QPE data, which supports the use of heavier precipitation exceedance thresholds than those used for the currently operational MOS QPF elements. Precipitation intensity resolution is also enhanced through derivation of a continuous precipitation element, where extrapolated amounts beyond the peak exceedance threshold can be large.

Limited comparative verification of several 6-h model QPFs in categorical form was presented. Preliminary results showed HRMOS QPF had better threat score and bias than corresponding GFS and GMOS QPFs. This finding suggests that the HRMOS QPFs may provide useful guidance for human forecasters. Also, additional comparative performance scoring (not discussed here) of the HRMOS QPF guidance against other model QPF guidance and human-produced QPFs support this finding. (An article on this subject for the formal literature is planned.) Thus, preparations are in progress to operationally implement the HRMOS model, whereby (1) GMOS QPF elements (currently operational over the CONUS) will be replaced by similar HRMOS QPF elements in early 2011, and 2) additional HRMOS QPF elements will be made available to the operational weather enterprise during late 2011.

The HRMOS QPF elements have been produced in an experimental mode since June 2008, and periodic model upgrades have been made to correct various QPF deficiencies discovered from daily monitoring. The performance of the PQPFs with the current model appears to be reliable, especially during "core months" of the warm and cool seasons. However, during season-transition months (March–April and September–October), we have noted instances of poor forecasts for highly unusual weather events, especially for the CP and BC elements. This suggests that extending the seasonal stratification in the model from two seasons to, say, four seasons could alleviate the problem.

Looking to the future of MOS QPF applications, we envision small but steady improvements in the coming years, as several opportunities show promise. These include 1) a refinement of the present HRMOS model to extend the seasonal stratification from two to four seasons, as noted above; 2) the potential use of a nonlinear regression method such as logistic; and 3) new MOS applications with predictors from multiple operational NWP models, ensemble NWP model output, and mesoscale NWP models, where convective precipitation is explicitly modeled. An important factor in the speed

that such new MOS QPF applications can be exploited is the availability of adequate historical samples from stable NWP model formulations. For QPF applications, where the heavy precipitation occurrences are rare events, at least several years of stable model output are needed.

APPENDIX A

BC Computation

A key component of the best precipitation category derivation is the computation of threshold probability constants. This one-time computation is performed with the dependent sample of PQPFs through an iterative procedure, whereby the threat score [same as the critical success index (CSI); Schaefer (1990)] for "trial" categorical forecasts is optimized within a prescribed bias range. (A trial categorical forecast denotes whether the PQPF value exceeded a trial threshold probability.) The prescribed bias range over all thresholds and forecast projections was 1.0–1.3, where unbiased forecasts have a 1.0 bias value. The slight overforecasting bias is considered a desirable attribute for QPF, especially for heavy amounts.

Another key component of the BC determination is the application of BC specification rules to the set of eight yes–no flags (section 6a). A long-standing practice at MDL has been to use a simplistic specification rule where BC is the highest precipitation threshold (category) with a yes value. For application here, this "top down" rule was modified to include the condition that all categories below the peak also have a yes value ("bottom up" rule). Preliminary tests revealed the bottom-up rule yielded coherent HRMOS BC patterns for well-behaved flag sets (where all flags are yes below the peak yes category and no otherwise). However, for rare occasions of ill-behaved flag sets, a spurious (unacceptable) hole or spike could appear in the BC map. Typical cases involve isolated convective precipitation, where POP values are low and PQPFs for higher precipitation categories flip-flop above or below the associated threshold values. Experimentation

with extended rules led to the formulation of a "decision tree" rule set for application to ill-behaved flag sets. The combination of the bottom-up and decision tree rules resulted in acceptable BC patterns.

Two additional problems with the BC specification required treatment. One was that BC bias was found to exceed the 1.3 bias upper bound used in the threshold probability computation. This "bias inflation" was most problematic for the highest thresholds and longest forecast projections, where forecast probability ranges are very small. The problem, which was found to be related to the diverse spatial scales of light versus heavy precipitation, was addressed by fine-tuning the bias range control used in the threshold computation. We found that by setting the bias range control to near 1.25 for light thresholds and short forecast projections and to near 1.10 for the highest thresholds and longest projections, the resulting BC bias fell within the targeted 1.0–1.3 range across all thresholds and projections.

Another problem that warranted treatment was that BC fields, initially, exhibited severe discontinuities along regional boundaries. The problem was due to the discontinuous spatial distribution of the threshold probabilities across the CONUS [i.e., since a probability threshold was fixed within a nonoverlapping region (Fig. 7), discontinuities in the threshold values arise at the regional boundaries]. The problem was treated by applying a threshold value to the associated overlapping region instead of the discrete region and then using the regional weighting technique to obtain smooth transitions of the thresholds between regions. This blending of the regional threshold constants is effective in preventing regional discontinuities in the BC grids.

APPENDIX B

CP Derivation

Continuous precipitation is computed from BC with an interpolation formula for the bounded precipitation categories and an extrapolation formula for the unbounded peak category. Given BC at a grid point, the interpolated CP is given by

$$\text{CP} = x_l + \alpha(x_u - x_l)\frac{(p - p_{\text{th}})}{(p_{\text{mxgrd}} - p_{\text{th}})} + \beta(x_u' - x_l')\frac{p'}{p_{\text{th}}'}, \quad \text{(B1)}$$

where

- x_l is the lower precipitation bound for the BC value,
- x_u is the upper precipitation bound for the BC value,
- p is the probability of occurrence of the BC exceedance threshold at the point,

- p_{mxgrd} is the maximum probability of occurrence of the BC exceedance threshold over the forecast domain,
- p_{th} is the threshold probability for the BC exceedance threshold at the point,

and the "prime" notation for the same variables in the third term indicates they apply to the next higher (not forecasted) BC value. Thus, the third term provides a supplemental contribution to CP from the next higher precipitation category, where the maximum contribution occurs as the forecast probability p' approaches the corresponding threshold probability p_{th}' (from below). With values of 0.50 and 0.38 for the empirical parameters α and β, respectively, CP comes close to x_u for $p = p_{\text{mxgrd}}$ and $p' \rightarrow p_{\text{th}}'$. Values for the empirical parameters were deduced from trial and error testing, where the eye appeal of CP maps and CP forecast performance scores were driving considerations.

An extrapolated CP value, which pertains to the unbounded peak BC, is given by

$$\text{CP} = x_l + (x_u'' - x_l)\frac{(p - p_{\text{th}})}{(p_{\text{max}} - p_{\text{th}})}, \quad \text{(B2)}$$

where x_l, p, and p_{th} are as in (B1); x_u'' is an estimated upper precipitation bound; and p_{max} is the maximum attainable forecast probability for the category at the point. From the developmental sample, the upper precipitation bound was estimated separately for the 6- and 12-h valid periods, each region, and each season from an analysis of the conditional distribution of the peak precipitation category. For the 6-h valid period, peak x_u'' values (over all regions) of 4.7 and 4.6 in. for the warm and cool seasons, respectively, appeared in region 11 (Fig. 7); a minimum value of 2.6 (2.7) in. for the warm (cool) season appeared in region 1 (5). The maximum possible probability for the peak category was similarly estimated, but its stratification was extended to include the forecast projection and model cycle; thus, the stratification for this parameter is the same as that for the threshold probabilities (appendix A). Both x_u'' and p_{max} were initially specified as regional constants; then, each constant was assigned to all points in an overlapping region, and, finally, the regional weighting technique was used to obtain smooth transitions of the constants across the regions.

REFERENCES

Antolik, M. S., 2000: An overview of the National Weather Service's centralized statistical quantitative precipitation forecast. *J. Hydrol.*, **239**, 306–337.

——, and N. M. Baker, 2009: On the ability to develop MOS guidance with short dependent samples from an evolving numerical model. Preprints, *23rd Conf. on Weather Analysis and Forecasting/19th Conf. on Numerical Weather Prediction,*

Omaha, NE, Amer. Meteor. Soc., 6A.1. [Available online at http://ams.confex.com/ams/pdfpapers/154242.pdf.]

Applequist, S., G. E. Gahrs, and R. L. Pfeffer, 2002: Comparison of methodologies for probabilistic quantitative precipitation forecasting. *Wea. Forecasting,* **17,** 783–798.

Bocchieri, J. R., 1979: A new operational system for forecasting precipitation type. *Mon. Wea. Rev.,* **107,** 637–653.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.,* **78,** 1–3.

Carter, G. M., J. P. Dallavalle, and H. R. Glahn, 1989: Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. *Wea. Forecasting,* **4,** 401–412.

Charba, J. P., 1987: Features of an operational objective 0-6 and 3-9 h system for forecasting heavy precipitation amounts. Preprints, *Seventh Conf. on Hydrometeorology,* Edmonton, AB, Canada, Amer. Meteor. Soc., 137–142.

——, 1998: The LAMP QPF products. Part I: Model development. *Wea. Forecasting,* **13,** 934–965.

——, and F. G. Samplatsky, 2009: Operational thunderstorm guidance forecasts to 24 hours on a 20-km grid. Preprints, *23rd Conf. on Weather Analysis and Forecasting/19th Conf. on Numerical Weather Prediction,* Omaha, NE, Amer. Meteor. Soc., 15B.5. [Available online at http://ams.confex.com/ams/pdfpapers/154173.pdf.]

——, and ——, 2011: Regionalization in fine-grid GFS MOS 6-h quantitative precipitation forecasts. *Mon. Wea. Rev.,* **139,** 24–38.

——, Y. Lui, M. H. Hollar, B. E. Exley, and A. Belayachi, 1998: Gridded monthly climatic frequencies of precipitation amount for 1-, 3-, and 6-h periods over the conterminous United States. *Wea. Forecasting,* **13,** 25–57.

Clark, A. J., W. A. Gallus Jr., M. Xue, and F. Kong, 2009: A comparison of precipitation forecasting skill between small convection-allowing and large convection-parameterizing ensembles. *Wea. Forecasting,* **24,** 1121–1140.

Cummins, K. L., M. J. Murphy, E. A. Bardo, W. L. Hiscox, R. B. Pyle, and A. E. Pifer, 1998: A combined TOA/MDF technology upgrade of the U.S. National Lightning Detection Network. *J. Geophys. Res.,* **103,** 9035–9044.

Daley, 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.

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,** 955–965.

Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. Forecasting,* **13,** 377–395.

Gilbert, K. K., R. L. Cosgrove, K. L. Sheets, and G. A. Wagner, 2009: Gridded model output statistics: Improving and expanding. Preprints, *23rd Conf. on Weather Analysis and Forecasting/19th Conf. on Numerical Weather Prediction,* Omaha, NE, Amer. Meteor. Soc., 6A.6. [Available online at http://ams.confex.com/ams/pdfpapers/154285.pdf.]

Glahn, H. R., 1985: Statistical weather forecasting. *Probability, Statistics, and Decision Making in the Atmospheric Sciences,* A. H. Murphy and R. W. Katz, Eds., Westview, 289–335.

——, and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.,* **11,** 1203–1211.

——, and D. Ruth, 2003: The new digital forecast database of the National Weather Service. *Bull. Amer. Meteor. Soc.,* **84,** 195–201.

——, K. Gilbert, R. Cosgrove, D. P. Ruth, and K. Sheets, 2009: The gridding of MOS. *Wea. Forecasting,* **24,** 520–529.

Hamill, T. M., and S. J. Colucci, 1998: Evaluation of Eta–RSM ensemble probabilistic precipitation forecasts. *Mon. Wea. Rev.,* **126,** 711–724.

——, and J. S. Whitaker, 2006: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application. *Mon. Wea. Rev.,* **134,** 3209–3229.

——, R. H. Hagedorn, and J. S. Whitaker, 2008: Probabilistic forecast calibration using ECMWF and GFS ensemble reforecasts. Part II: Precipitation. *Mon. Wea. Rev.,* **136,** 2620–2632.

Henkel, A., and C. Peterson, 1996: Can deterministic quantitative precipitation forecasts in mountainous regions be specified in a rapid, climatologically consistent manner with Mountain Mapper functioning as the tool for mechanical specification? *Abstracts, Fifth National Heavy Precipitation Workshop,* State College, PA, NWS/NOAA, 31 pp. [Available from Office of Climate, Water, and Weather Services, W/OS, 1325 East-West Hwy., Silver Spring, MD 20910.]

Hughes, K. K., 1999: AVN-based statistical forecasts of thunderstorms and severe thunderstorms for the contiguous U.S. Preprints, *17th Conf. on Weather Analysis and Forecasting,* Denver, CO, Amer. Meteor. Soc., 201–205.

——, 2004: Probabilistic lightning forecast guidance for aviation. Preprints, *11th Conf. on Aviation, Range, and Aerospace/22th Conf. on Severe Local Storms,* Hyannis, MA, Amer. Meteor. Soc., 2.6. [Available online at http://ams.confex.com/ams/pdfpapers/81859.pdf.]

Iredell, M., and P. Caplan, 1997: Four-times-daily runs of the AVN model. NWS Tech. Procedures Bull. 442, 3 pp.

Jensenius, J. S., Jr., 1990: A statistical comparison of the forecasts produced by the NGM and LFM for the 1987/88 cool season. *Wea. Forecasting,* **5,** 116–127.

——, 1992: The use of grid-binary variables as predictors for statistical weather forecasting. Preprints, *12th Conf. on Probability and Statistics in the Atmospheric Sciences,* Toronto, ON, Canada, Amer. Meteor. Soc., 225–230.

Kanamitsu, M., 1989: Description of the NMC Global Data Assimilation and Forecast System. *Wea. Forecasting,* **4,** 335–342.

——, and Coauthors, 1991: Recent changes implemented into the Global Forecast System at NMC. *Wea. Forecasting,* **6,** 425–435.

Lowry, D. A., and H. R. Glahn, 1976: An operational model for forecasting probability of precipitation—PEATMOS PoP. *Mon. Wea. Rev.,* **104,** 221–232.

National Weather Service, 1974: Postprocessing the LFM forecasts. Tech. Procedures Bull. 174, 20 pp.

Olson, D. A., N. W. Junker, and B. Korty, 1995: Evaluation of 33 years of quantitative precipitation forecasting at NMC. *Wea. Forecasting,* **10,** 498–511.

Reap, R. M., and D. S. Foster, 1979: Automated 12–36 hour probability forecasts of thunderstorms and severe local storms. *J. Appl. Meteor.,* **18,** 1304–1315.

Rodgers, E., and Coauthors, 2005: The NCEP North American Mesoscale modeling system: Final Eta Model/analysis changes and preliminary experiments using the WRF-NMM. Preprints, *21st Conf. on Weather Analysis and Forecasting,* Washington, DC, Amer. Meteor. Soc., 4B.5. [Available online at http://ams.confex.com/ams/pdfpapers/94707.pdf.]

MONTHLY WEATHER REVIEW VOLUME 139

Ruth, D. P., B. Glahn, V. Dagastaro, and K. Gilbert, 2008: The performance of MOS in the digital age. *Wea. Forecasting,* **24,** 504–519.

Schaefer, T. J., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting,* **5,** 570–575.

Sloughter, J. M., A. E. Raftery, T. Gneiting, and C. Fraley, 2007: Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *Mon. Wea. Rev.,* **135,** 3209–3220.

Wier, S., M. Mathewson, and T. J. LeFebvre, 1998: Grid editing for the Interactive Forecast Preparation System. Preprints, *14th Int. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology,* Phoenix, AZ, Amer. Meteor. Soc., 469–473.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Academic Press, 627 pp.

——, and T. A. Hamill, 2007: Comparison of ensemble-MOS methods using GFS reforecasts. *Mon. Wea. Rev.,* **135,** 2379–2390.