FORECASTERS’ FORUM

Using Percentiles to Communicate Snowfall Uncertainty

DAVID R. NOVAK, KEITH F. BRILL, AND WALLACE A. HOGSETT*

NOAA/NWS/NCEP/Weather Prediction Center, College Park, Maryland

(Manuscript received 11 February 2014, in final form 16 June 2014)

ABSTRACT

An objective technique to determine forecast snowfall ranges consistent with the risk tolerance of users is demonstrated. The forecast snowfall ranges are based on percentiles from probability distribution functions that are assumed to be perfectly calibrated. A key feature of the technique is that the snowfall range varies dynamically, with the resultant ranges varying based on the spread of ensemble forecasts at a given forecast projection, for a particular case, for a particular location. Furthermore, this technique allows users to choose their risk tolerance, quantified in terms of the expected false alarm ratio for forecasts of snowfall range. The technique is applied to the 4–7 March 2013 snowstorm at two different locations (Chicago, Illinois, and Washington, D.C.) to illustrate its use in different locations with different forecast uncertainties. The snowfall range derived from the Weather Prediction Center Probabilistic Winter Precipitation Forecast suite is found to be statistically reliable for the day 1 forecast during the 2013/14 season, providing confidence in the practical applicability of the technique.

1. Introduction

Quantifying and communicating uncertainty has challenged the weather enterprise (Hirschberg et al. 2011). This is particularly true for snowfall, where uncertainties in precipitation amount, precipitation type (e.g., Kain et al. 2000; Lackmann et al. 2002), and snow-to-liquid ratio (SLR; e.g., Roebber et al. 2003; Alcott and Steenburgh 2010) are simultaneous factors affecting snowfall predictability. Further, regional features such as topography and proximity to moisture sources may affect predictability (e.g., Maglaras et al. 1995; Niziol et al. 1995; Schultz et al. 2002). A series of northeast U.S. snowstorms in the early 2000s (24–25 January 2000, 30 December 2000, and 4–5 March 2001) of varying forecasting success (Kocin and Uccellini 2004, 4–6) have highlighted the challenges of communicating snowfall uncertainty.

The U.S. weather enterprise communicates snowfall uncertainty in a variety of ways, including the probability of occurrence and probability of exceeding thresholds. Morss et al. (2008) found that the general U.S. public prefers ranges of information relative to deterministic information, and, indeed, the most common approach to communicating snowfall is a range (e.g., 2–4 in.). However, it is often unclear whether this range represents the variation of snowfall over an area, or the uncertainty in values at a point. In many cases, it is also unclear whether the ranges are based on quantitative measures of uncertainty. Similar public misinterpretations of probability of precipitation forecasts have been documented (e.g., Joslyn et al. 2009).

This article proposes an objective technique using percentiles from a probability distribution function (PDF) to determine forecast snowfall ranges consistent with the risk tolerance of users. This technique is dynamic, with the resultant ranges varying based on the spread of ensemble forecasts. Furthermore, this technique allows users to choose the risk tolerance, quantified as the false alarm ratio of their selected forecast range. Section 2 describes the methods used in this paper. The technique is applied to the 4–7 March 2013 snowstorm at two different locations in section 3 to illustrate its use. Section 4 presents
a statistical verification of the technique applied to the day 1 forecast during the 2013/14 season. The results are summarized in section 5.

2. Description of methods

The computation of forecast snowfall ranges is facilitated by the availability of a PDF, from which percentiles can be derived. The practice of applying PDFs is becoming more common with the maturity of ensemble datasets. The Weather Prediction Center (WPC) generates PDFs to populate its extensive suite of probabilistic winter precipitation forecasts (PWPFs; http://www.wpc.ncep.noaa.gov/pwpf/wwd_accum_probs.php). Although the WPC snowfall PDFs are not perfectly calibrated, the WPC dataset will be used to demonstrate the snowfall range method, and will be verified in section 4.

a. PWPF method

The PWPF method depends on a deterministic forecast and an estimate of uncertainty. The operational WPC Winter Weather Desk (WWD) forecaster generates 24-h snowfall accumulations, which serve as the deterministic forecast for the method. A multimodel ensemble with 6-h temporal resolution serves as an estimate of the forecast uncertainty. A 28-member multimodel ensemble was used for the example shown in section 3, and this configuration is described in Table 1. The multimodel ensemble was expanded to 32 members for the 2013/14 season (Table 1), and this configuration was used for the statistical verification presented in section 4.

Snowfall is calculated for each member of the multimodel ensemble. Precipitation type for National Centers for Environmental Prediction (NCEP) models is determined by applying the dominant-type algorithm (Manikin 2005). Precipitation type for non-NCEP models is determined by applying a simple decision tree algorithm using near-surface (2 m) temperature; and temperatures at the 925-, 850-, and 700-hPa mandatory isobaric levels. At each grid point where the precipitation type is diagnosed as snow, an SLR is applied. The SLR is an average of the value obtained using the Roebber et al. (2007) neural network algorithm (Rnna) applied to North American Mesoscale Model (NAM), the value from Rnna applied to GFS, a seasonal climatological value (Baxter et al. 2005), and the constant value of 11.0. Otherwise, where the precipitation type is diagnosed as ice pellets, an SLR of 2:1 is applied, and the amount is included in the snowfall accumulation.

A binormal PDF (Toth and Szentimrey 1990), which allows skewness, is used to create a distribution of values around the WPC deterministic snowfall forecast at each grid point. The fitting of the binormal distribution is a method-of-moments approach, as described by Wilks (2006, p. 80). The WPC deterministic snowfall forecast is treated as the mode of the distribution. The placement of the WPC forecast in the ensemble order statistics determines the skewness of the distribution. The variance of the distribution is matched to the variance of the multimodel ensemble. The WPC deterministic forecast is included as an additional member of the ensemble for the computation of the variance. This fit is done at each grid point; so, the PDF varies from grid point to grid point. Percentiles are derived from the final forecast PDF.

b. Snowfall range method

Given percentile information, users may generate their own forecast ranges derived from the PDF described above based on their own risk tolerance. We propose a simple method, which proceeds as follows:

1) Choose a false alarm ratio $X$ based on user-specific risk tolerance. This is the percentage of forecasts for which the user is willing to accept an observed snowfall amount outside the forecast range.

2) Apply the appropriate percentile range as
For example, assume a user has a low risk tolerance and can only accept a false alarm ratio of 10%. In this case, \( PL = 10/2 = 5 \) and \( PU = 100 - (10/2) = 95 \). Thus, the 5th and 95th percentiles are used to create the range. Assume a different user has a high risk tolerance and can accept a false alarm ratio of 50%. Thus, \( PL = 50/2 = 25 \) and the \( PU = 100 - (50/2) = 75 \), dictating use of the 25th and 75th percentiles to create the range.

In general, for a static false alarm ratio, the greater the forecast uncertainty, the larger the range. Thus, we expect the snowfall range to vary on a case-to-case basis (larger ranges for more uncertain events), vary by forecast projection (narrowing as the forecast projection decreases), and vary across a spatial area given an independent range is calculated at each grid point. An idealized example of the range at a single point varying with forecast projection for users with false alarm ratios of 20% and 50% is shown in Fig. 1.

### 3. Case example

The 4–7 March 2013 snowstorm is used to illustrate application of the technique to real data. This storm brought more than 25 cm (10 in.) of snow across 11 states from Montana to Massachusetts. The storm brought the largest snowfall accumulation in nearly two years to Chicago, Illinois, with 23.4 cm (9.2 in.) recorded at O’Hare International Airport (ORD). This snowfall was well forecast, and occurred in a cold air mass along an inverted trough extending from the parent low (Fig. 2a).

![FIG. 1. Idealized example of how the snowfall range may vary by forecast projection for a given forecast point for two different chosen false alarm ratios. The ranges for the 20% (50%) false alarm ratio are shown in solid (dashed). The corresponding percentiles are labeled in the legend. The idealized observation point is shown by a solid box.](image)

![FIG. 2. Surface evolution of the early March 2013 snowstorm valid at (a) 1200 UTC 5 Mar and (b) 1200 UTC 6 Mar 2013 showing isobars every 4 mb (solid), the 0° and −18°C isotherms (dashed blue), precipitation (green shading), and fronts and pressure systems. [Adapted from the NOAA Daily Weather Maps.]](image)

However, as the low moved east, rapid cyclogenesis occurred along the East Coast (Fig. 2b). The short-range forecast for the Washington, D.C., metropolitan area highlighted a snowfall range of 10–20 cm (4–8 in.), prompting schools and the federal government to close. However, less than an inch was observed in the city and points east, as the boundary layer temperatures remained just above freezing. Thus, the snowfall range forecast technique is applied to Chicago and Washington in this case to illustrate its use in different locations with different forecast uncertainties.

For purposes of this demonstration, a risk tolerance associated with a 20% false alarm ratio is assumed. Thus, \( PL = 10 \) and \( PU = 90 \), and users can expect that the observed snowfall will be within this range in 80% of the cases. In the Chicago region the day 3 forecast of the 10th and 90th percentiles showed a large spread, with the 10th percentile showing nearly no snow over the region, while the 90th percentile showed over 30 cm
Applying the risk tolerance to a specific point, the day 3 forecast range for ORD would be 0–35 cm (0–14 in.; see Fig. 4). As the event drew closer, the forecast snowfall range derived using $P_L$ and $P_U$ narrowed to 10–28 cm (4–11 in.) by day 2 (Fig. 4), and remained nearly the same for day 1. This relatively large range communicates to users that considerable meteorological uncertainty exists, even at a short day 1 lead time. The observed snowfall amount at ORD [23.4 cm (9.2 in.)] did fall within the forecast range, as should be expected in 80% of cases given the 20% false alarm ratio.

The forecast for the Washington region was even more uncertain. The day 3 forecast of the 10th and 90th percentiles shows an incredibly large spread over the mid-Atlantic, with the 10th percentile showing the possibility of no snow over most of the region, while the 90th percentile showed a crippling snowstorm for Maryland and Virginia (Fig. 5). Applying a risk tolerance to a specific point, the day 3 forecast range associated with a 20% false alarm ratio at Ronald Reagan Washington National Airport (DCA) was 0–50 cm (~0–20 in.; see Fig. 6). As the event drew closer, the range at DCA tightened to 2–38 cm (~1–15 in.; see Fig. 6). This is obviously still a very large range, which signals to users a very uncertain forecast just hours before the first precipitation fell. The observed amount was nearly outside of the forecast range, but users aware of the large range would not have been surprised.

4. Verification

The data on which the snowfall range method is based must have statistical reliability to be useful to decision makers. Thus, an assessment of the statistical reliability of the snowfall range forecasts derived from the PWPF
was conducted for the day 1 forecast during the 2013/14 season. A 20% false alarm ratio was assumed such that $P_L = 10$ and $P_U = 90$. Percentages of observed snowfall values falling below, within, and above the forecast ranges were computed using station observations of 24-h snowfall (ending at 1200 UTC each day) obtained from the National Climatic Data Center (NCDC) Cooperative Observer (COOP) Summary of the Day collection. Only stations that consistently appeared in the COOP collection on every date beginning 1 December 2013 and ending 31 March 2014 were used. The objective selection criterion yields 29 stations east of 105°W, 2 stations west of that longitude, and 1 station with obviously erroneous reports. The latter three stations are not included. The remaining 29 stations were used in the verification and are shown in Fig. 7.

Once the station locations are established, the PWPF database is scanned to obtain the values for the 10th and 90th percentile 24-h forecasts of 24-h snowfall accumulations valid at 1200 UTC, matching the observation valid time. The values are interpolated bilinearly from the 20-km-resolution forecast grid to the point locations. The data are further thinned by removing the trivial cases for which the observed snowfall and the 10th and 90th percentile forecasts are zero. The remaining 2693 cases are used to determine the fraction of observations falling below, within, and above the forecast range defined by the 10th and 90th percentile values.

Results of the analysis are shown in Table 2. The PWPF was remarkably statistically reliable for the day 1 forecast during the 2013/14 season, with 83% of the

1 The snowfall range forecasts are fixed probability central credible interval forecasts for which Winkler’s score is appropriate (Wilks 2006, 303–304); however, Winkler’s score is challenging to interpret in the absence of a suitable benchmark forecast. Winkler’s score is not used.
observations falling within the range (perfect is 80%). This indicates a slightly overdispersive PDF, with a slight bias toward incorporating too many (more than 10%) observed amounts above the 10th percentile of the forecast distribution (Table 2). Although the verification was only performed for select percentiles, limited geographical areas, and forecast projections, the results lend confidence that the PWPF provides statistically reliable PDFs for the snowfall range approach proposed in this paper.

5. Discussion and summary

This article describes an objective technique developed to determine forecast snowfall ranges consistent with the risk tolerance of users. The technique uses percentiles from a PDF and is dynamic, with the resultant ranges varying according to the forecast uncertainty at a given place, time, and forecast situation. Furthermore, this technique allows users to choose the risk tolerance, quantified as the false alarm ratio of their selected range forecast. The technique is demonstrated using an early March snowstorm in two different metropolitan areas with different levels of uncertainty. The proposed technique is offered as an approach to communicating snowfall forecasts in a meaningful, quantitative way, and may motivate a more flexible approach to users’ preparation for snowfall.

The use of the percentile accumulations in the manner described assumes that the PDF upon which they are based is well calibrated. Verification of the WPC PWPF dataset for the day 1 forecast for the 2013/14 season shows that it was reasonably well calibrated for this purpose.

The above verification results provide confidence in the practical applicability of the technique. However, the spatial resolution of the underlying gridded data may not be sufficient to resolve large snowfall gradients (such as may be found in areas of complex terrain). Further, one may question whether the public is ready for wide and/or varying ranges. For example, how will users respond to a forecast of 0–20 in.? Will users accept uncommon ranges, such as 3–9 in.? Future work exploring such questions, as well as the technique’s applicability to other meteorological variables (i.e., temperature), is encouraged.

Acknowledgments. Discussions with Dr. James Hoke motivated this work on the use of snowfall percentiles to communicate uncertainty. This work further benefited from insightful discussions with staff from the Baltimore/Washington Weather Forecast Office. Two anonymous reviewers provided constructive comments leading to improvements in the presentation of this work. The views expressed are those of the authors and do not necessarily represent a NOAA/NWS position.
REFERENCES


