



The 2017-2018 HMT-WPC Winter Weather Experiment



14 November 2017 – 09 March, 2018
Weather Prediction Center
College Park, MD

Findings and Results

Michael Bodner - NOAA/NWS/WPC, College Park, MD

Sarah Perfater - I.M. Systems Group, NOAA/NWS/WPC/HMT, College Park, MD

Benjamin Albright - Systems Research Group, NOAA/NWS/WPC/HMT, College Park, MD

Joshua Kastman - CIRES, NOAA/NWS/WPC/HMT, College Park, MD

Table of Contents

Introduction	2
Science and Operations Objectives	2
Experiment Operations	3
Executing a Remote Experiment	4
Cases	5
Verification	11
Findings and Research to Operations Recommendations	15
Micro-physical Post-Processed Method	15
<i>12-km NAM and 3-km NAM Nest</i>	15
<i>12-km GFDL/EMC FV3</i>	16
Model Explicit Method	19
<i>HRRRX Precipitation Type Forecasts</i>	19
<i>HRRRX Snowfall Accumulation Forecasts</i>	22
<i>HRRRE Precipitation Type Forecasts</i>	27
<i>HRRRE Snowfall Accumulation Forecasts</i>	30
Environmentally-Assessed Weighted Probabilities	33
<i>Weighted Percentages</i>	33
<i>Conditional Probabilities</i>	34
<i>NBMv3.1 Snowfall</i>	36
Environmentally-Assessed Ensemble Probabilities -- Fuzzy Clustering	39
<i>Daily Derivation of Clusters</i>	39
<i>Application to Forecasting Precipitation Type</i>	41
HREFv2 Point/EAS Precipitation Type Probability Methods	44
WWE Precipitation Type Forecasts	50
Round-Table Discussion Summaries	54
Acknowledgements	56
References	57
Appendix A: Featured Guidance and Tools for Experimental Forecasts	58
Appendix B: WPC MODE Settings for Objective Verification	66
Appendix C: Corrected Precipitation Type Algorithm for the NAM Nest	67

Updated May 16, 2018

Introduction

In an effort to support improvements in both WPC and WFO winter weather forecasts, the Hydrometeorology Testbed at WPC (HMT-WPC) conducted the 8th Annual Winter Weather Experiment (WWE). The experiment brought together members of the operational forecasting, research, and academic communities to address winter weather forecast challenges.

The 2017-18 WWE provided an opportunity for participants to evaluate a suite of both deterministic and probabilistic precipitation type algorithms in a real time setting to assign the best precipitation type for the thermal environment. Experimental data sets were used side by side with current operational methods.

The 2017-18 WWE was conducted both remotely and on location in College Park, MD this year utilizing the web-based distance communication software, Mikogo, paired with a telecon to encourage and promote interactivity and engagement with the participants. More on the remote aspect of the experiment can be found in the Remote Experiment section below.


Science and Operations Objectives

The main objectives of the 2017-18 Winter Weather Experiment were to:

- Explore the testing of multiple microphysical and probabilistic precipitation type methodologies to determine which methods best enhance the forecast process.
- Explore the applications of Fuzzy Clustering to the prediction of winter weather.
- Enhance collaboration among NCEP centers, WFOs, and NOAA research labs on winter weather forecast challenges.

Fast findings for science and operations goals can be found in Table 1.

Table 1. Transition Metrics for select experimental guidance and techniques



FY18 Transition Metrics

WPC-HMT

- **Winter Weather Experiment**

Major Tests Conducted	Transitioned to Operations	Recommended for Transition to Operations	Recommended for Further Development & Testing	Rejected for Further Testing	Funding Source
NBMv3 Conditional Prob Snow	X		X*		MDL
NBMv3 Conditional Prob Ice	X		X*		MDL
NBMv3.1 Snowfall Forecast		X	X*		MDL
HRRRE Prob of precip type		X			ESRL/GSD
Variable density snowfall accumulation for application in the HRRR/HRRRE			X		ESRL/GSD
HREF EAS method for prob of ptype		X			WPC/EMC/GSD USWRP
Stony Brook Fuzzy Clustering for Precipitation Type			X		WPC/ Stony Brook U.
Totals	2	2	5	0	

**although currently or soon to be operational, HMT recommends continued development and testing*

Experiment Operations

The experiment was conducted weekly over the **full winter season** beginning Tuesday, November 14, 2017 and ending Friday, March 9th, 2018. This year, the WWE was executed **remotely** from the WPC-OPC Collaboration Room at the NOAA Center for Weather and Climate Prediction (NCWCP) in College Park, MD. Experiment participants and field representatives were asked to join remotely on “Forecast Tuesdays” (10:30 am-12:30 pm EST) and “Verification Wednesdays” (10:30 am-12:30 EST) throughout the season. Experimental data and tools were explored and used to create experimental precipitation type forecasts on Tuesdays. Forecast and data set verification was presented and discussed and subjective feedback collected on Wednesdays. Note that the experiment was not conducted over the weeks of November 20th, December 25th, January 1st, and January 7th due to holiday scheduling and the annual AMS meeting.

Each Tuesday, the experiment participants were exposed to a suite of experimental precipitation type model guidance and asked to create experimental precipitation type forecasts for snow, ice and sleet, in four 6-hour time stamps valid over 24 hours for the **Day 2 (12 – 12 UTC) or Day 3 (12 -12 UTC) time periods**. If a winter event in the Day 1 period proved to be a more interesting event to explore, **Day 1 (18 - 12 UTC)** could also be evaluated using the guidance, however only three 6-hour time stamps were prepared in the forecast. Participants were called upon to contour areas of precipitation type (snow, freezing rain) for each of the 6-hour times stamps for the selected forecast day. The 5km WPC QPF was used as a component of the snow and ice forecast. This was an essential piece of areal guidance in preparing the placement of the precipitation type forecasts. During the first half of the experiment season, the 5km WPC was used in conjunction with the WPC snow-to-liquid ratio (SLR), the ice-to-liquid ratio (ILR), and the experiment precipitation type forecast to generate 6 and 24 hour snow/ice accumulation forecasts. In order to focus the verification session to be more evaluation of the experimental precipitation type tools, the generation of snow/ice forecasts was eliminated for the second half of the experiment.

On Wednesdays, participants had the opportunity to objectively analyze and subjectively evaluate the performance of both the experimental forecasts and the experimental model guidance over the previous week. Retrospective cases were analyzed to fill periods lacking significant winter weather or to capitalize on interesting events.

In addition to the remote experiment, there were two weeks of residence experiments at NCWCP which included contributing developers, research to operations (R2O) partners, and NWS WFO Science Operations Officers (SOOs). The two weeks for the in-house experiment at NCWCP were:

Week 1: February 12th - 16th, 2018

Week 2: March 5th - 9th, 2018

Executing a Remote Experiment

Driving factors, such as reduced travel funds and missing major winter events when being limited to a 4-week residence-only experiment, were reasons to explore executing the WWE remotely this year. WPC utilized Mikogo (<https://go.mikogo.com>), a screen sharing technology, which can be used easily in both Linux and Windows environments. When multiple users were connected to the same Mikogo session, presenter controls could be exchanged during the session to allow for multiple members of the testbed staff and expert presenters to share experiment data and information. Each experiment day, the teleconference number and new Mikogo session ID was posted in a document that was shared with all experiment participants.



Figure 1. *On the left, a remote screen with the Mikogo panel present and data. On the right, participants and leaders in the Collaboration Room at WPC discussing the data along with those on the teleconference call.*

In addition to the Mikogo document, a Google Drive folder shared with the experiment participants contained an attendance document, a Google Forms survey through which scores and comments were recorded during verification, the experient orientation, all available documentation related to the experimental model guidance, verification session slides, and the presentations given by experts throughout the experiment. Participants and contributors had access to all experiment documents to explore and inquire on their own time and collaboratively with the group.

Feedback about the remote experiment was positive. Participation and number of winter weather cases were markedly improved. The remote experiment enabled participation and the infusion of subject experts from a distance with ease and enriched the program. The primary drawback was the inability to efficiently engage those who were participating remotely in an interactive process. WPC-HMT will explore ways to improve efficiencies in remote participation for following experiments.

Cases

Table 2 gives an overview of the specific cases that were verified each week of the remote experiment and each day during the 2 in-house weeks over the approximate 4 month experiment. A total of 17 cases were verified over the 9 remote sessions and the 2 in-house experiment weeks. The cases are arranged in chronological order based on the start and end dates of the cases verified. This section will give a broad synoptic overview of the 4 month time period as well as highlight some of the more interesting and challenging cases that the forecasters and participants faced during the experiment.

Table 2. *Chronological list based on the start and end date of the 17 total cases verified over the course of the 2017/18 WWE.*

Start Date	Start Time	End Date	End Time	Experiment Date (Week Number-(Day) - Date)
11/16/17	12Z	11/17/17	12Z	2 - 11/29/17
12/1/17	12Z	12/2/17	12Z	3 - 12/6/17
12/6/17	12Z	12/7/17	12Z	4 - 12/13/17
12/13/17	12Z	12/14/17	12Z	5 - 12/20/17
12/20/17	12Z	12/21/17	12Z	6 - 1/17/18
1/3/18	12Z	1/4/18	12Z	10-Th - 2/15/18
1/8/18	12Z	1/9/18	12Z	10-W - 2/14/18
1/16/18	12Z	1/17/18	12Z	7 - 1/24/18
1/23/18	12Z	1/24/18	12Z	8 - 1/31/18
2/1/18	12Z	2/2/18	12Z	9 - 2/7/18
2/4/18	12Z	2/5/18	12Z	10-F - 2/16/18
2/7/18	12Z	2/8/18	12Z	10-Tu - 2/13/18
**Note Week 11 had no forecast/verification exercise due to technical problems.				
2/17/18	12Z	2/18/18	12Z	12 - 2/28/18
2/22/18	12Z	2/23/18	12Z	13-W - 3/7/18
2/25/18	12Z	2/26/18	12Z	13-F - 3/9/18
2/28/18	12Z	3/1/18	12Z	13-Tu - 3/6/18
3/2/18	12Z	3/3/18	12Z	13-Th - 3/8/18
Key	Remote Verification Weeks		In-House Verification Days	

Figure 2 shows the average 500 hPa heights over the United States during select periods over the course of the entire experiment.

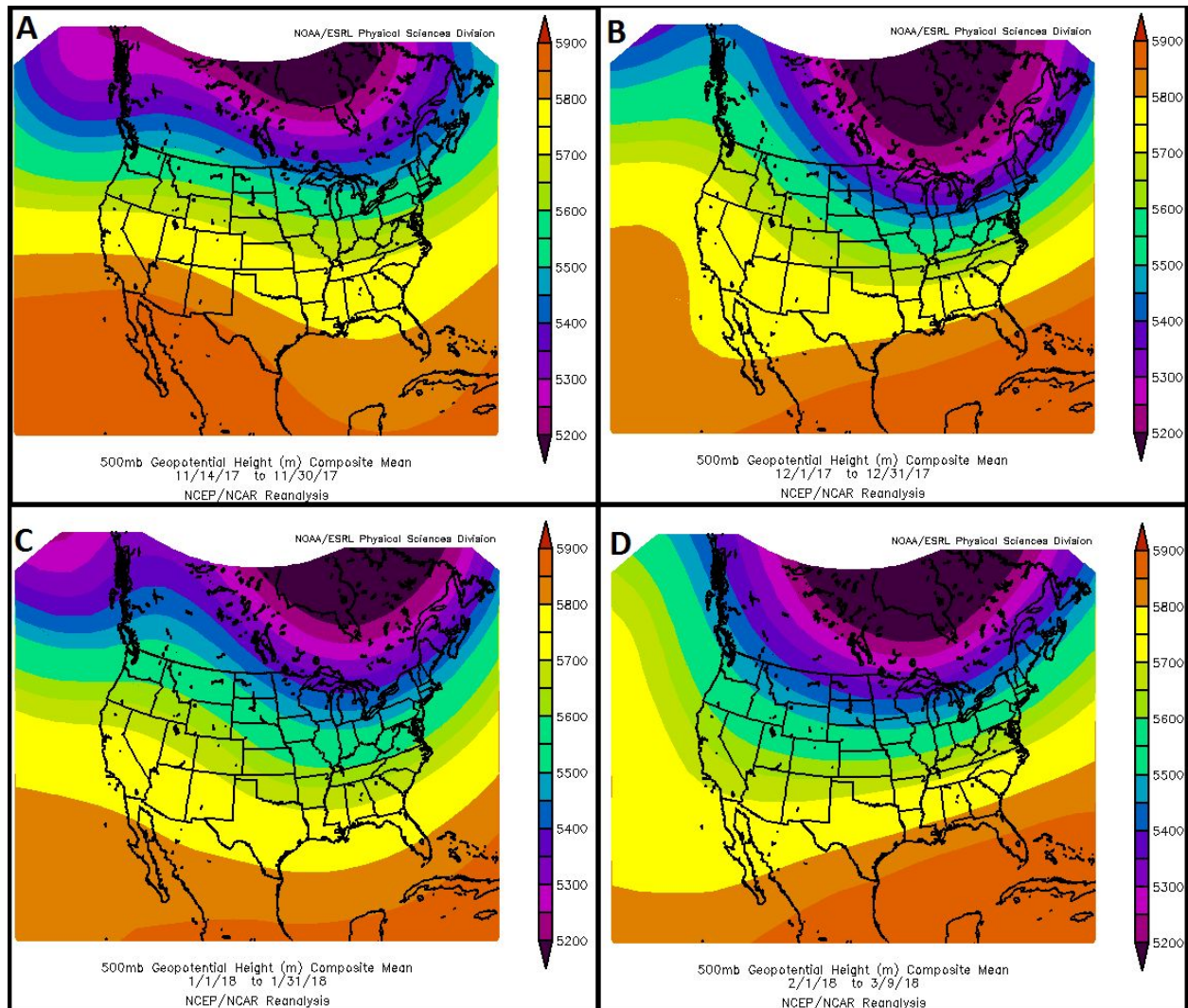


Figure 2. Average 500 hPa heights for the periods of (A) November 14 - 30, 2017, (B) December 1 - 30, 2017, (C) January 1 - 31, 2018, and (D) February 1 - March 9, 2018. Images generated from the NCEP/NCAR Reanalysis provided by NOAA/ESRL/PSD (<http://www.esrl.noaa.gov/psd/data/composites/day/>).

The last half of November (Figure 2A) featured a deep trough over the eastern third of the United States (U.S) with ridging in the Western U.S. In December, (Figure 2B) the trough flattened in the East and the axis shifted west while the pattern amplified in the Western U.S. with a ridge just off the coast and a trough in the Southwest. During January (Figure 2C), the pattern was similar to the latter half of November featuring a trough in the eastern half of the U.S. and a broad ridge over the western states. Finally, in February and the first week of March (Figure 2D), the 500 hPa pattern became more zonal in the central and southern latitudes with a trough dominating the northern part of the U.S. A much more amplified pattern existed off

each respective coast in the Pacific and Atlantic Oceans. Figure 3 shows the associated 500 hPa anomalies for the abovementioned periods.

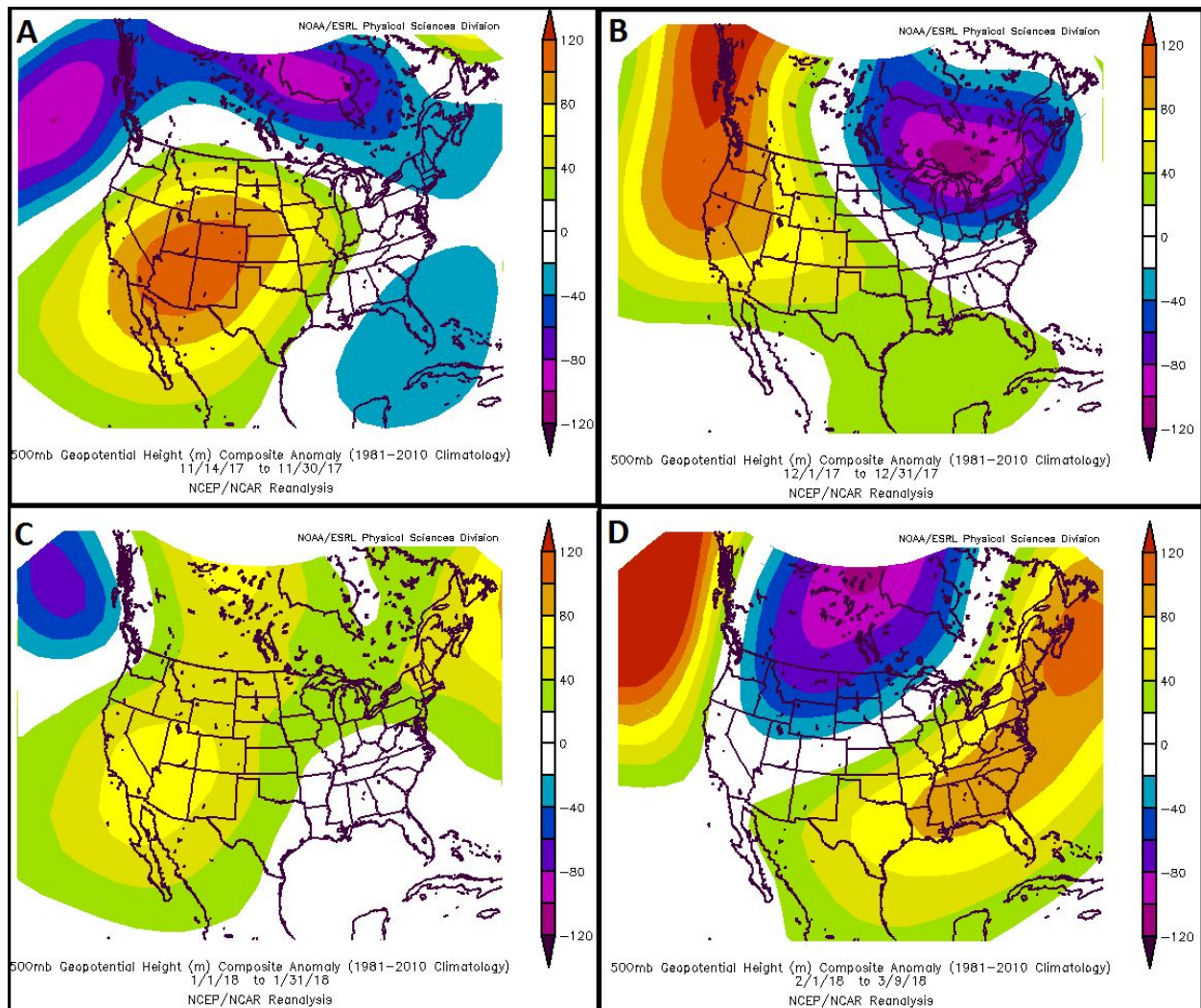


Figure 3. Average 500 hPa height anomalies for the periods of (A) November 14 - 30, 2017, (B) December 1 - 30, 2017, (C) January 1 - 31, 2018, and (D) February 1 - March 9, 2018. Images generated from the NCEP/NCAR Reanalysis provided by NOAA/ESRL/PSD (<http://www.esrl.noaa.gov/psd/data/composites/day/>).

Average surface temperature anomalies are shown in Figure 4. The first three periods from the middle of November to the end of January were predominantly warmer in the Western United States and in portions of the Central Plains. In December (Figure 4B), colder than normal temperatures existed over much of New England as well as southeastern Texas. Much of the eastern part of the country was colder than normal in January (Figure 4C), especially areas in the Southeast where temperatures were up to 4 degrees below normal. The pattern shifted in

February and the first week of March (Figure 4D) with most areas east of the Mississippi River experiencing above normal temperatures as high as 4 degrees. Much below normal temperatures were located in the northern Central Plains and the Northwest U.S.

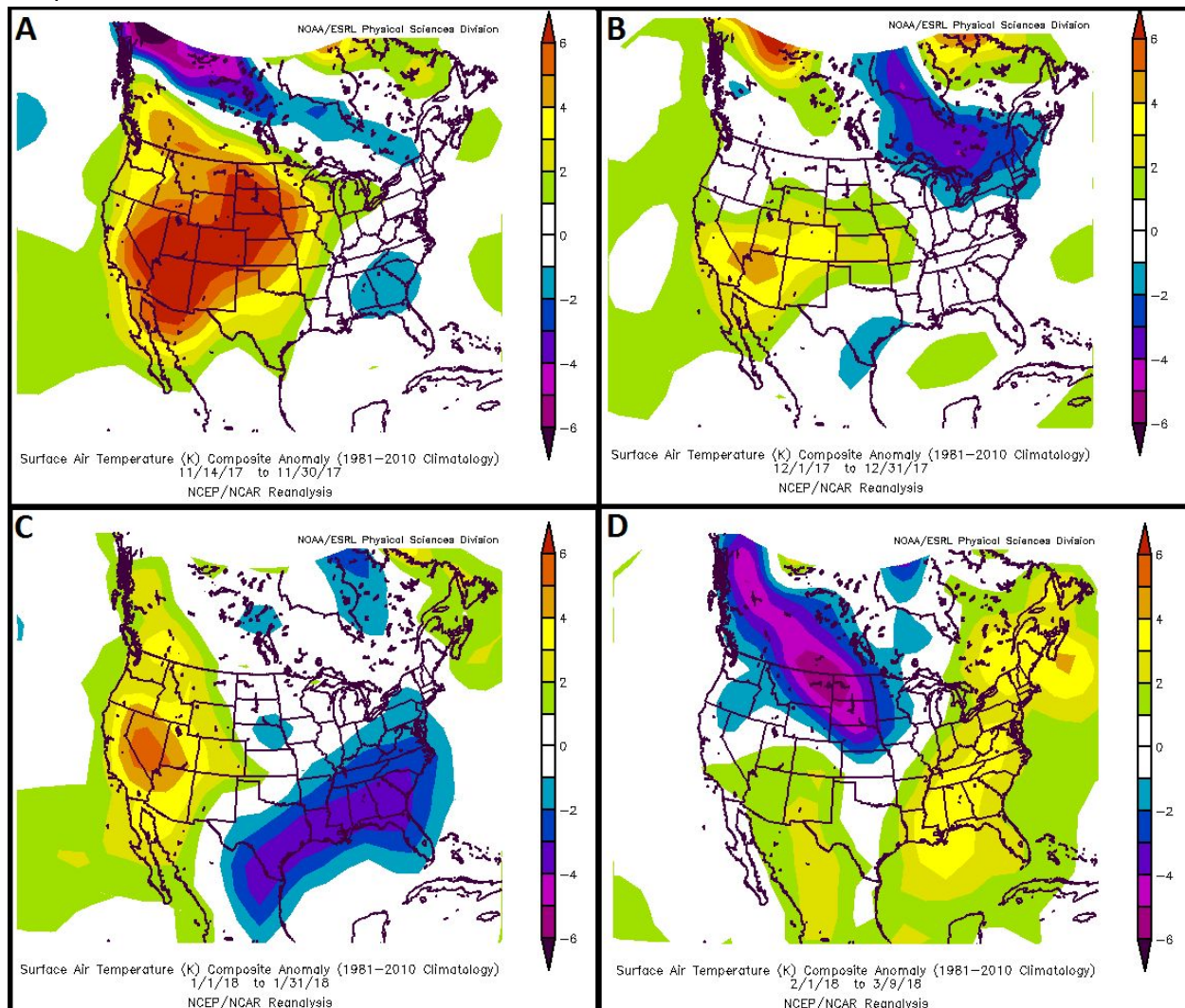


Figure 4. Average surface temperature anomalies for the periods of (A) November 14 - 30, 2017, (B) December 1 - 30, 2017, (C) January 1 - 31, 2018, and (D) February 1 - March 9, 2018. Images generated from the NCEP/NCAR Reanalysis provided by NOAA/ESRL/PSD (<http://www.esrl.noaa.gov/psd/data/composites/day/>).

The Southeastern U.S. had a particularly snowy winter season. Figure 5 shows two notable events that occurred during the experiment and were verified as cases by participants. The first event (Figure 5 left) produced measurable snow well into South Texas and along the western Gulf Coast with heavy amounts throughout northern Georgia and the western Carolinas. During the second event (Figure 5 right), snow fell over much of the southeastern coast with very heavy amounts in southeast Virginia. Freezing rain was also observed along the coast in

eastern South and North Carolina.

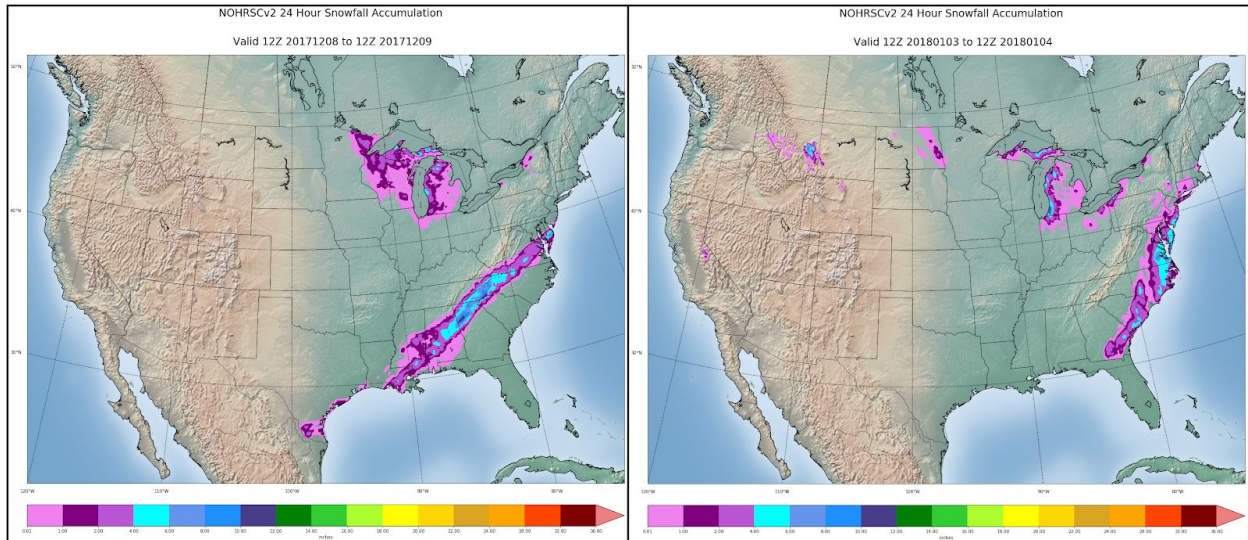


Figure 5. (Left) National Operational Hydrologic Remote Sensing Center Snowfall Analysis version 2 (NOHRSCv2) 24 hour snow analysis valid from 12Z 12/8 - 12Z 12/9/2017 and (right) valid from 12Z 1/3 - 12Z 1/4/2018. Scale = 0.01;1;2;4;6;8;10;12;14;16;18;20;22;24;28;32;36 inches)

Despite the below-normal surface temperatures shown previously in the months of December and January in the Mid-Atlantic states and New England, some of the more impactful events in those regions came in February and early March. Figure 6 shows some of the more notable events that impacted those regions as well as a few major events in the Western U.S. The western and central portion of Montana especially saw significant snowfall in these and other events. The event shown in the top right of Figure 6 valid 12Z 2/7/18 - 12Z 2/8/18 was notable not only for the snowfall amounts over the Northeast, but also the mixed precipitation in the form of freezing rain in the southern Mid-Atlantic states. Overall, extending the WWE over the majority of the winter season provided unique opportunities and challenges to utilize the tools to forecast and verify a wide variety of cases.

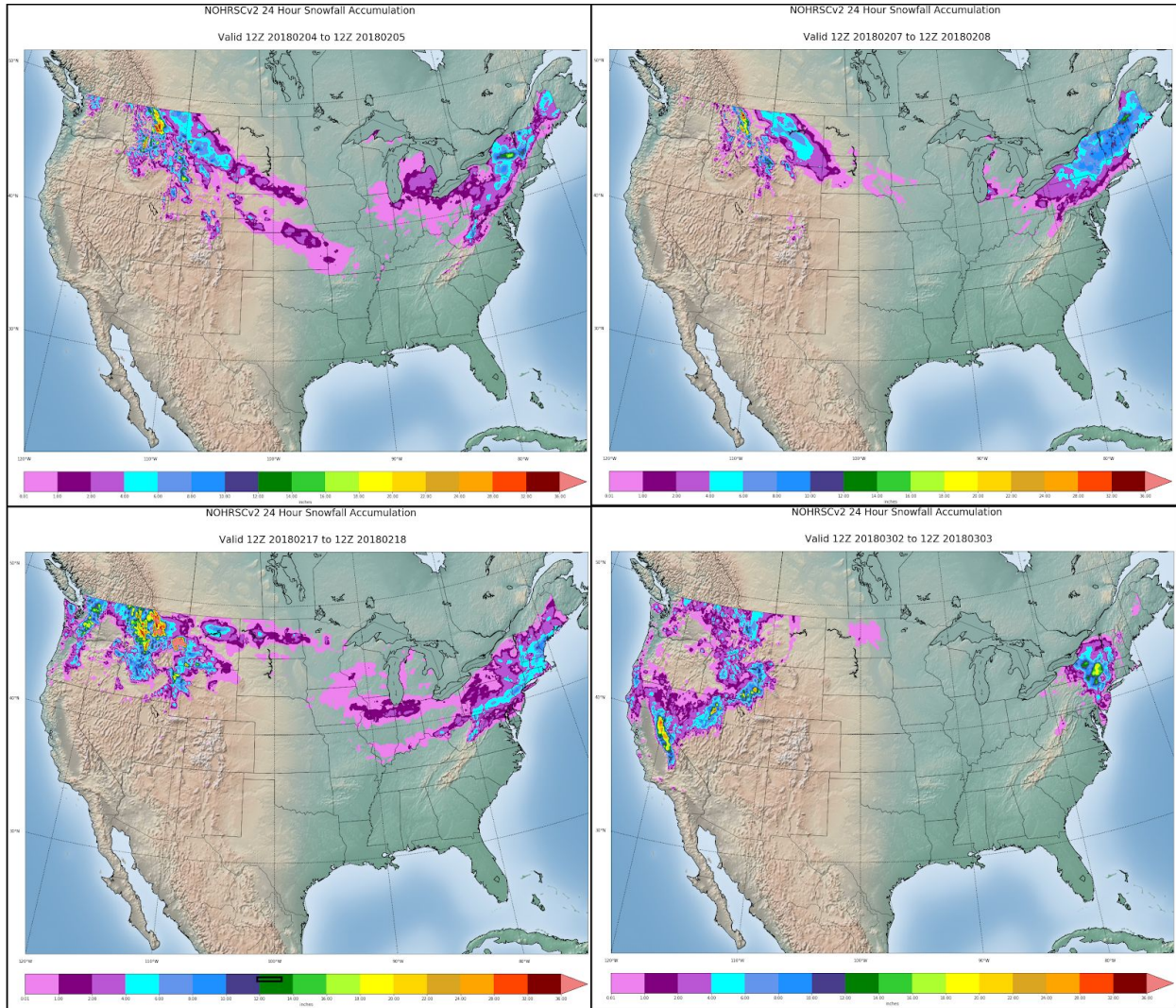


Figure 6. (Top Left) NOHRSCv2 24 hour snow analysis valid from 12Z 2/4 - 12Z 2/5/2018, (Top Right) valid from 12Z 2/7 - 12Z 02/8/2018, (Bottom Left) valid from 12Z 2/17 - 12Z 02/18/2018, and (Bottom Right) valid from 12Z 3/2 - 12Z 3/3/2018. Scale = 0.01;1;2;4;6;8;10;12;14;16;18;20;22;24;28;32;36 inches)

Verification

For this year's remote experiment, verification was conducted over a two-hour session each Wednesday morning from 10:30 AM - 12:30 PM EST. Users would follow along remotely by viewing a slideshow shared through the Mikogo software and fill out an online Google Form survey individually. Participants from EMC were often in the room during the remote sessions and provided feedback in person. During the in-house weeks, a similar slideshow was shown and individual scores and comments were collected from those in the room. Online, remote participation was still possible and encouraged during the two in-house weeks.

The National Operational Hydrologic Remote Sensing Center Snowfall Analysis Version 2 (NOHRSCv2; available online at <https://www.nohrsc.noaa.gov/snowfall/>) was used to verify model forecasted snowfall amounts as well as to help verify snow precipitation type forecasts throughout the experiment. A 24-hour analysis is available at 00Z and 12Z each day and a 6-hour analysis is available at 00Z, 06Z, 12Z, and 18Z. NOHRSCv2 is a synthesis of numerical weather prediction (NWP) short-term forecasts that are bias- and error-corrected with River Forecast Centers (RFC) precipitation analyses that are converted to snowfall using a climatological, gridded snow-to-liquid ratio, which is all augmented and enhanced with surface snowfall observations.

For the experiment, the 12Z 24 hour analysis and the four 6 hour analyses were used, typically at least two days beyond the valid time to ensure greater quality control. The 6 hour NOHRSCv2 analysis was used for the first time during the WWE for verifying snowfall precipitation type in six hourly timesteps, therefore the actual snowfall amounts from the six hour analyses were not evaluated. The 0.01 inch contour was plotted to estimate the overall snowfall extent to help verify snowfall precipitation type.

To verify hourly precipitation, the components of the main verification methodology are quantitative precipitation estimate (QPE) data from Stage IV (Lin and Mitchell, 2005) or the Multi-Radar Multi-Sensor Gauge Corrected (MRMS-GC) QPE and hourly initialization fields from the 13 km Rapid Refresh (RAP) Model. The precipitation type is derived by employing a WPC algorithm which scans for freezing temperatures at 925 hPa, 850 hPa, and 700 hPa to estimate the depth of the cold layer as well as any melting and refreezing layers to identify a sleet or freezing rain environment. Two-meter temperatures are used to distinguish freezing rain versus rain. If sleet, freezing rain, or rain is ruled out, snow is identified as the precipitation type where 850 hPa temperatures are less than 0 degrees Celsius (-2 degrees Celsius over higher terrain). The precipitation types derived from this methodology are conditional based on the environment, therefore the 0.01 inch Stage IV or MRMS-GC QPE contour was overlaid. "Truth" in verification was determined to be areas over which the QPE contours overlapped an identified conditional precipitation type probability from the RAP analysis. Figure 7 shows an example of the RAP/MRMS-GC precipitation type/QPE analysis. For select examples, the 3 km experimental High-Resolution Rapid Refresh (HRRR) was used in place of the RAP to determine precipitation type in an effort to evaluate the WPC algorithm using a higher resolution model.

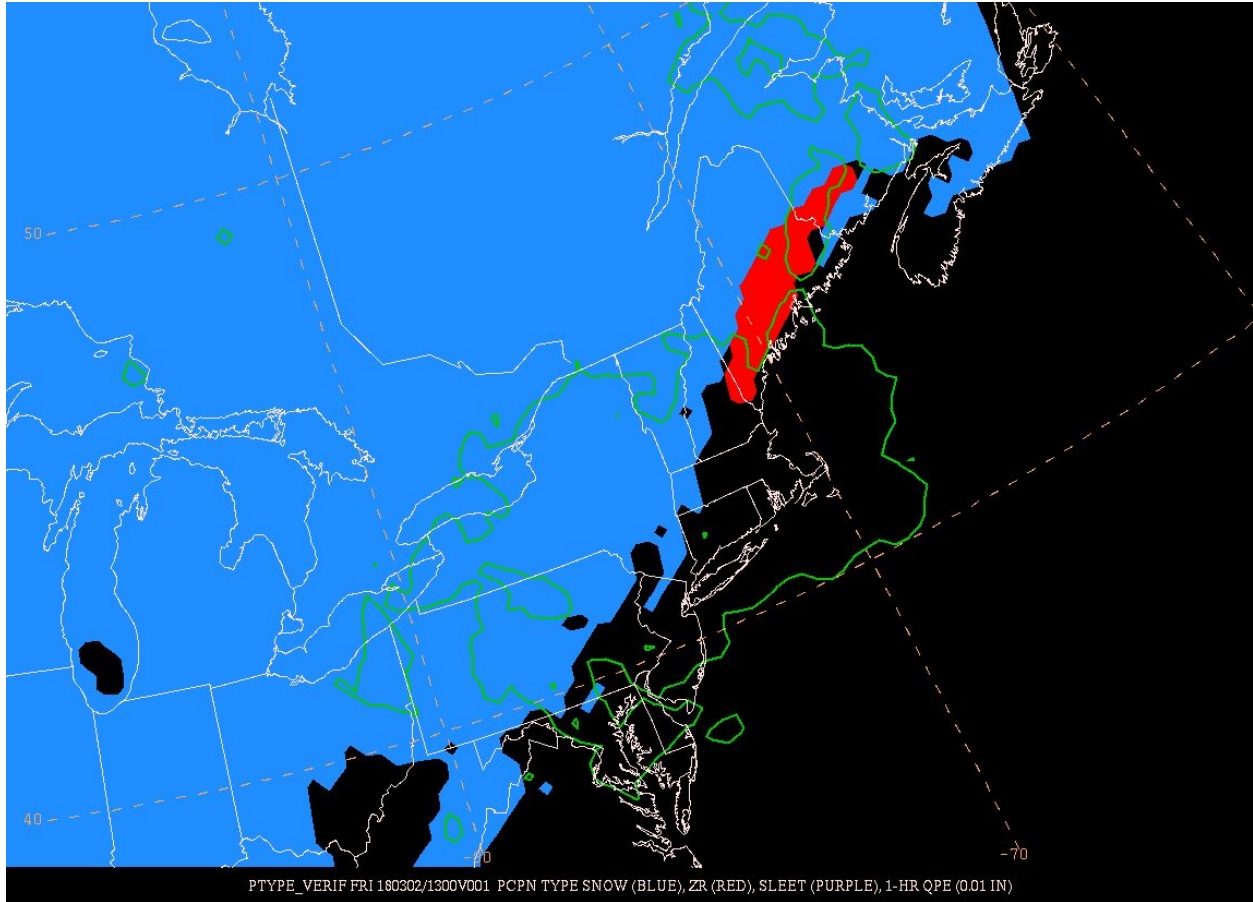


Figure 7. An example of the RAP/MRMS-GC precipitation type verification valid at 13Z 3/2/18. Areas of blue indicate snow conditional precipitation type, red freezing rain conditional precipitation type, and the green contours denote the 0.01 inch QPE from MRMS-GC.

Surface observations were also used to verify precipitation type. Local storm reports (LSRs) issued from local WFOs were gathered over the 24 hour verification period and displayed in six hour increments that matched the periods of the precipitation type forecasts. Any reports for snow or freezing rain were used, as those were the two main precipitation types that were sought. One of the major issues with using LSRs for precipitation type verification is the timing of the reports. There were many instances in which LSRs were issued for precipitation that occurred for a much earlier period than the issuance time. For example, at 12Z there would be LSRs for snowfall that occurred the previous day and in areas confirmed to have little to no precipitation during the first 12Z-18Z period of verification. Figure 8 shows an example of a large group of LSRs in a region that experienced snow the previous day. The area circled in the left image of Figure 8 still has many LSRs for the 12Z - 18Z period of 2/25/18 despite them being located south of the 6 hour NOHRSCv2 0.01 inch snowfall contour which is indicated as the thick blue line on the left. This indicates that the LSRs are reporting the previous day's snowfall, which is shown on the right side of the figure.

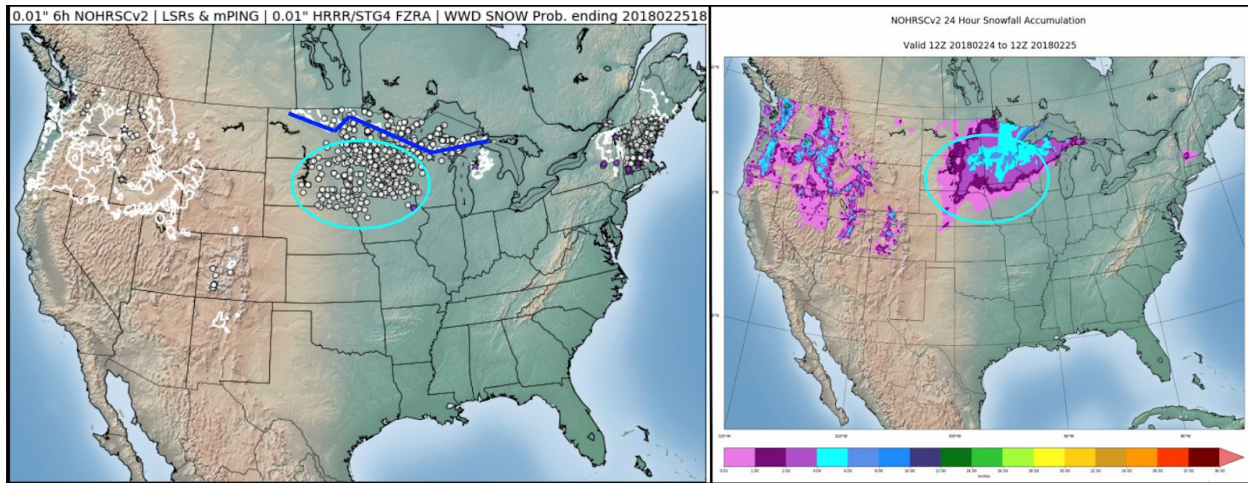


Figure 8. Example showing LSRs on the left valid from 12Z - 18Z 2/25/2018 and NOHRSCv2 24-hour snowfall valid 12Z 2/24/2018 to 12Z 2/25/2018. The dark blue line signals the southward extent of the 0.01 inch snowfall contour during the six hour, 12Z - 18Z 2/25/2018 period of interest. The cyan circle (left) highlights LSRs still being reported for snowfall that fell the previous 24 hours, as highlighted by the cyan circle (right).

Meteorological Phenomena Identification Near the Ground (mPING) reports for snow, graupel, freezing rain, and freezing drizzle were collected in six hour intervals and displayed as another source of ground verification. A drawback of mPING reports is that they are not quality controlled. Therefore at times seemingly erroneous reports would be located in areas far away from the area of interest that experienced precipitation.

Finally, roughly halfway through the experiment, select automated surface observing stations (ASOS) that include freezing rain sensors were also included in verification. The observations were collected at six hour intervals. Because precipitation type was the main verification interest, the amount of freezing rain measured at each station was less important than whether any amount was measured. One problem that became apparent with the six hour ASOS freezing rain measurements was low amounts were often recorded in areas of freezing fog. Therefore, when plotting the stations that had any non-zero amount of freezing rain, many false alarm observations would be plotted in areas of freezing fog where there was no measurable precipitation.

In addition to subjective verification, the Method for Object-Based Diagnostic Evaluation (MODE) was used to compare various snowfall accumulations forecasts at select thresholds from the HRRRX and HRRRE to the NOHRSCv2 24 hour snowfall analysis (see Appendix B for WPC MODE settings). MODE outputs various statistics comparing the forecast objects (model snowfall accumulation) to the observed objects (NORHSCv2 snowfall analysis) including centroid distance, angle, and intersection area. In addition, MODE outputs contingency table

statistics. The Gilbert Skill Score (GSS) and critical success index (CSI), commonly referred to as Equitable Threat Score (ETS) and Threat Score (TS) respectively, were also computed over the whole domain for the two models and featured during the verification sessions.

Findings and Research-to-Operations Recommendations

Micro-physical Post Processed Precipitation Type Method

Precipitation type data used in the experiment that was derived from microphysics parameterization in the models was classified as *micro-physical post processed precipitation type*. These methodologies leveraged elements from the respective microphysics schemes, and the quantitative precipitation forecasts (QPF) from the model to estimate an instantaneous classification of precipitation.

12km NAM (NAM12) and NAM CONUS Nest

The National Center for Environmental Prediction (NCEP) North American Model (NAM) suite is run with the Ferrier-Aligo (FA) microphysics scheme. The FA scheme features a variable density graupel parameter, also known as the rime factor. The rime factor is a ratio of the growth of snow by liquid accretion plus vapor deposition divided by vapor deposition. Thus the higher the liquid accretion, the higher the resultant rime factor. A rime factor for the lowest sigma layer of the atmosphere is provided in the post processed fields for both the NAM12 and NAM CONUS Nest (See Appendix A for more details). Five years of evaluating the rime factor parameter in precipitation type transition zones has demonstrated that rime factors over 10 typically align with increased graupel in the lowest sigma layer (HMT-WPC 2015), and therefore a higher likelihood of sleet being the instantaneous precipitation type for that time.

The fraction of frozen precipitation parameter is also heavily utilized in winter weather forecasting and the identification of precipitation type. This parameter estimates the percentage of the model QPF that will be frozen upon reaching the ground. Using the fraction of frozen precipitation field to estimate precipitation type works directly with the QPF and thermal profile, and alleviates the need to empirically estimate warm layers in the forecast model soundings.

During the experiment, WPC tested the forecasting of precipitation types by using an adjustment of the current EMC algorithm used in the NAM suite. The specific goal of the WPC adjustment was to better identify areas of sleet. Widespread and impactful sleet events occurred in the mid Atlantic region in March 2017 (Novak, 2017), and in January 2011 (Hamrick, 2011). The sleet in both cases was not well forecast. Therefore the WPC algorithm was adjusted to put a greater emphasis on sleet in identifying potential precipitation type. The algorithm logic is listed below, and microphysics parameters were checked at each time stamp:

If Fraction of Frozen Precipitation > 90, P-type is Snow
If Fraction of Frozen Precipitation <= 89, and >= 50, P-type is Sleet
If Fraction of Frozen Precipitation <= 49,
P-type is Rain if 2-meter temperature > 32F
P-type is Freezing Rain if 2-meter temperature <= 32F

A final check step was done to highlight sleet:

If Rime Factor is > 20, then P-type is Sleet thereby overwriting any previously assigned P-type

Findings

The WPC precipitation type methodology for the NAM12 and NAM CONUS Nest was available for the duration of the experiment. Subjective evaluation showed that the identification of areas of snow was good in most cases using this methodology, however rain and freezing rain were too often not forecast. Upon review with EMC developers, it was determined that the sleet overwrite step resulted in over forecasting of sleet as the precipitation type at the expense of the other non-snow precipitation types. Moreover, the sleet overwrite step resulted in several mis-assignments and mis-placements of snow, freezing rain, and rain. Given the forecast inconsistencies in this finding, participant scores will not be included in this report.

Recommendation

The recommendation from the WPC-HMT team was to rework the sleet overwrite component of the algorithm which checks the microphysics parameters. Testing where these algorithm changes was deployed in retrospective cases and will be added to the NAM12 and NAM CONUS NEST for future use. Upon collaboration with EMC, corrections to the WPC algorithm calculations were applied after the WWE was completed. **Appendix C** highlights a retrospective case example in which the old and new calculations can be compared.

FV3 (13km)

A beta version of the next generation global prediction system (NGGPS), the Finite Volume Cubed-Sphere (FV3) with a Lagrangian adaptive vertical grid and 13 km horizontal resolution was provided by EMC for the second half of the experiment. This beta run featured the GFDL microphysics scheme which is currently scheduled to be deployed with the forthcoming operational NGGPS-FV3. A six-hour average precipitation type parameter was made available, and was useful when assigning a 6-hour precipitation type during the experiment forecast exercises. The six-hour average is determined by tallying all precipitation types assigned in each time step in the model run history, then the most dominant is assigned as the deterministic precipitation type over the 6-hour period. This is the same methodology applied to the current operational GFS forecast for precipitation type, where the most hazardous species is assigned if two or more types have equal weighting. The hazard preference is freezing rain, snow, sleet, and rain.

Findings

The FV3 data was available for the experiment from January 17 through March 9. The FV3 precipitation type fields were implicit to the model QPF, therefore the performance of the 6-hour average precipitation type was often correlated to the timing and placement of the FV3 QPF. The box plot shown in Figure 9 shows the spread of the scores between a low of 2 and a high of 9. Most scores fell between 4 and 6 out of 10 as shown in Figure 10. The average participant score of the FV3 precipitation type over the experiment was 6.06 out of 10.

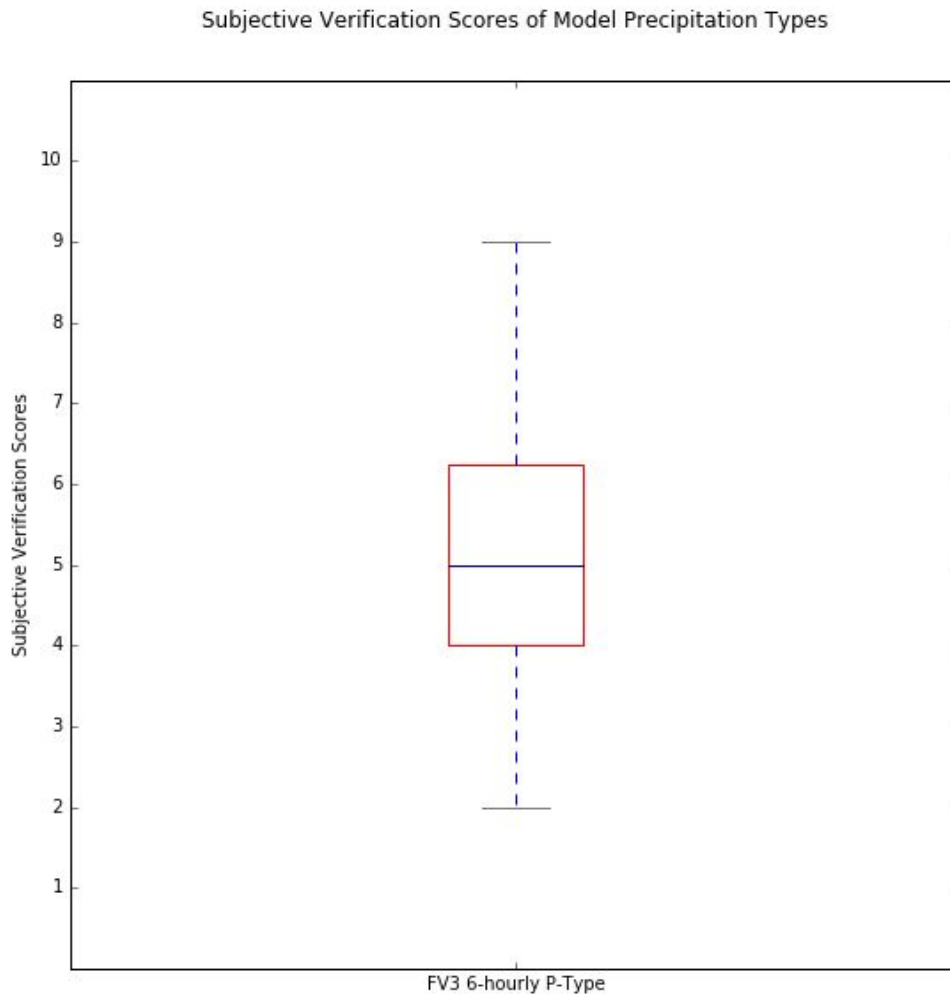


Figure 9. Box plot of the subjective scores for the FV3 6-hourly precipitation type forecasts.

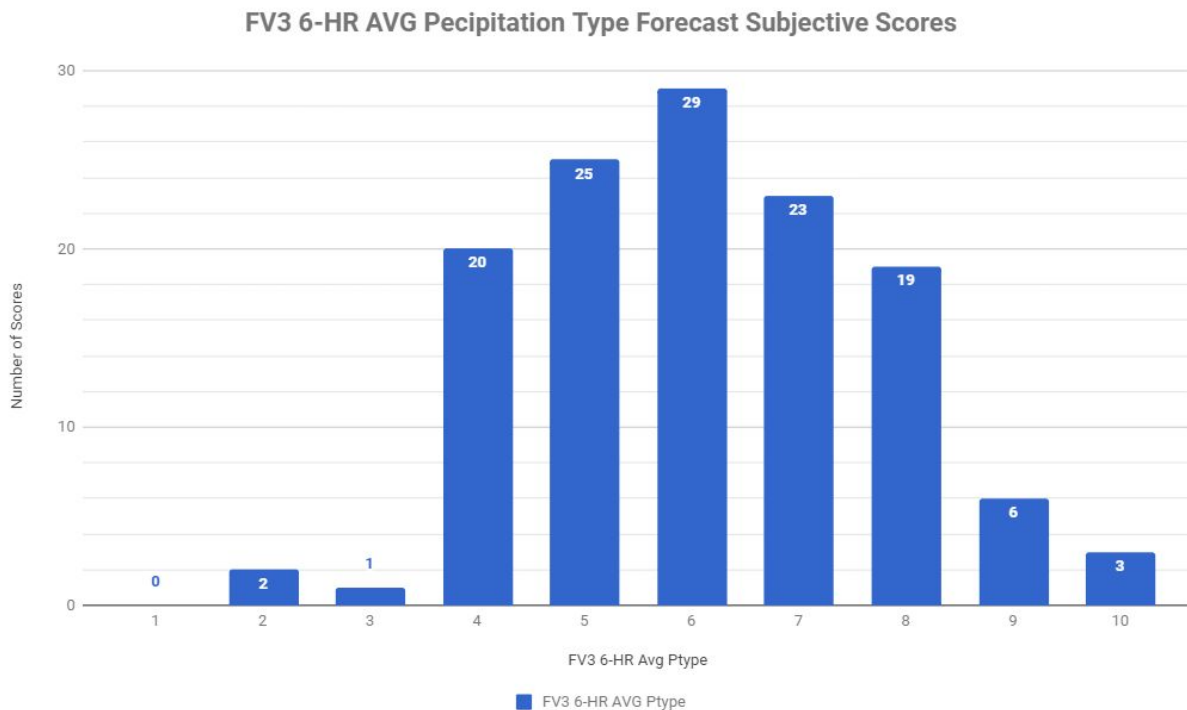


Figure 10. Individual subjective evaluation score distribution over the course of the entire experiment for the FV3 6-hourly precipitation type forecasts.

Participants were often satisfied with the identification of the precipitation type transition zones identified by the FV3. However, there still was a tendency to not forecast enough sleet and freezing rain in precipitation type transition zones. As shown in Figure 11, timing of the cold air was a struggle at times resulting in a solution that was too warm and producing rain when snow or ice was observed.

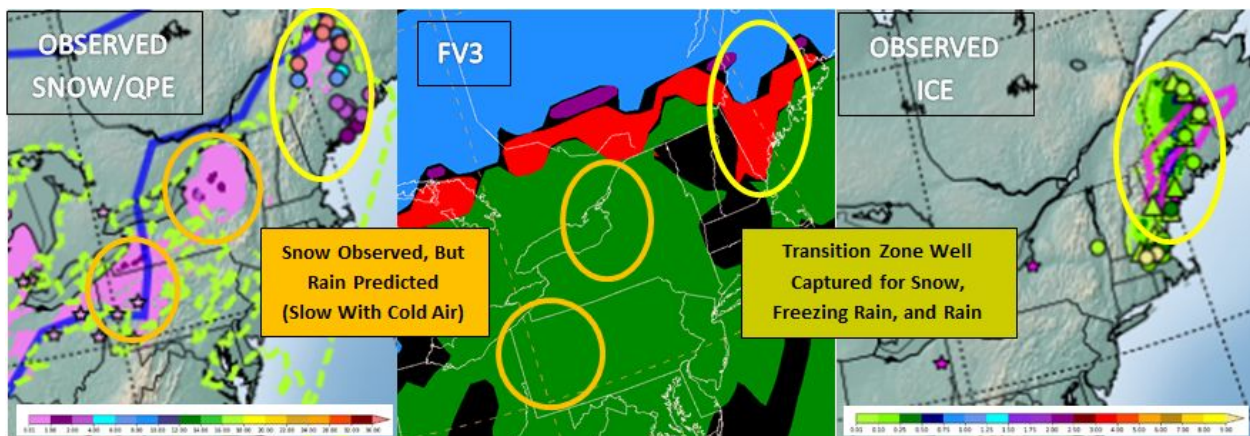


Figure 11. 6-hour observed snowfall (NOHRSCv2) in pink and QPE in green dashes on the left,

6-hour observed ice in green (right), and the 6-hour FV3 precipitation forecast in the center all valid from 18 UTC January 23 to 00 UTC January 24, 2018. The orange circles identify areas where snow was observed but rain was predicted in the FV3. The yellow circles identify a transition zone that correctly predicted snow, freezing rain, and rain based on the observations.

Recommendation

EMC indicated that the six-hour average precipitation type field will not be available after this winter. Therefore other methodologies will need to be tested in future versions of FV3. In FY19, a density graupel parameter, or other parameter similar to the FA rime factor field, is scheduled to be delivered in the FV3. If this field becomes available, a precipitation type methodology similar to that employed in the NAM can be implemented and tested. Testing the application of snow, ice, and rain mixing ratios found in the high resolution structure of the FV3 to precipitation type identification will also be conducted.

Model Explicit Method

The HRRRX and HRRRE both use a model explicit microphysics based algorithm to predict instantaneous precipitation type (snow, sleet, rain and/or freezing rain). Up to 3 precipitation types may be forecast simultaneously, with the exception being that the algorithm will choose between rain OR freezing rain as they can not occur together. The algorithm uses 2-m temperature, 3-D hydrometeor mixing ratios and fall rates to provide a first guess of precipitation type reaching the ground.

HRRRX Precipitation Type Forecasts

The HRRRX precipitation type algorithm was evaluated during the second half of the experiment. For each verification session a case was chosen, usually the previous week's forecast or another event that had occurred, and participants would evaluate the precipitation type hourly in six hour blocks over the course of 24 hours. To verify precipitation type, the RAP analysis precipitation type algorithm and MRMS-GC QPE were used. Anywhere the MRMS-GC QPE 0.01" contour overlapped a specific precipitation type, then that would be considered a "hit." Participants ranked each six hour forecast block on a scale of 1 (very poor) to 10 (very good) based on how well the HRRRX precipitation type algorithm matched with the verification over those six hours.

Figure 12 is a box plot for the subjective evaluation scores for the hourly HRRRX precipitation type forecast over all the individual six hour periods and Figure 13 shows the distribution of all the individual scores given by the participants. The average subjective score for the HRRRX precipitation type algorithm performance was 6.60 with a standard deviation of 1.52.

Subjective Verification Scores of Model Precipitation Types

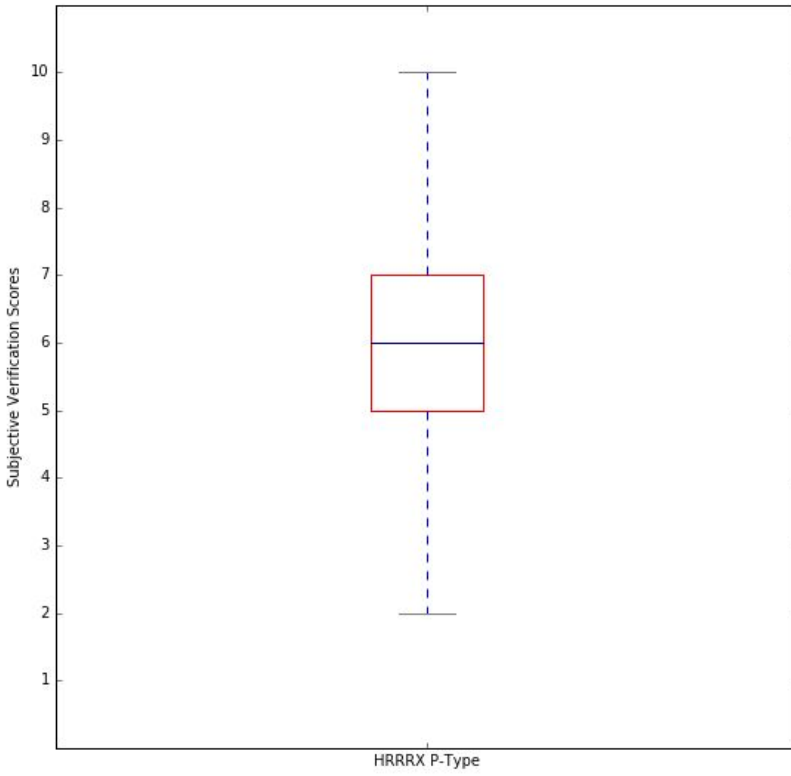


Figure 12. Box plot of the subjective scores for the HRRRX hourly precipitation type forecasts.

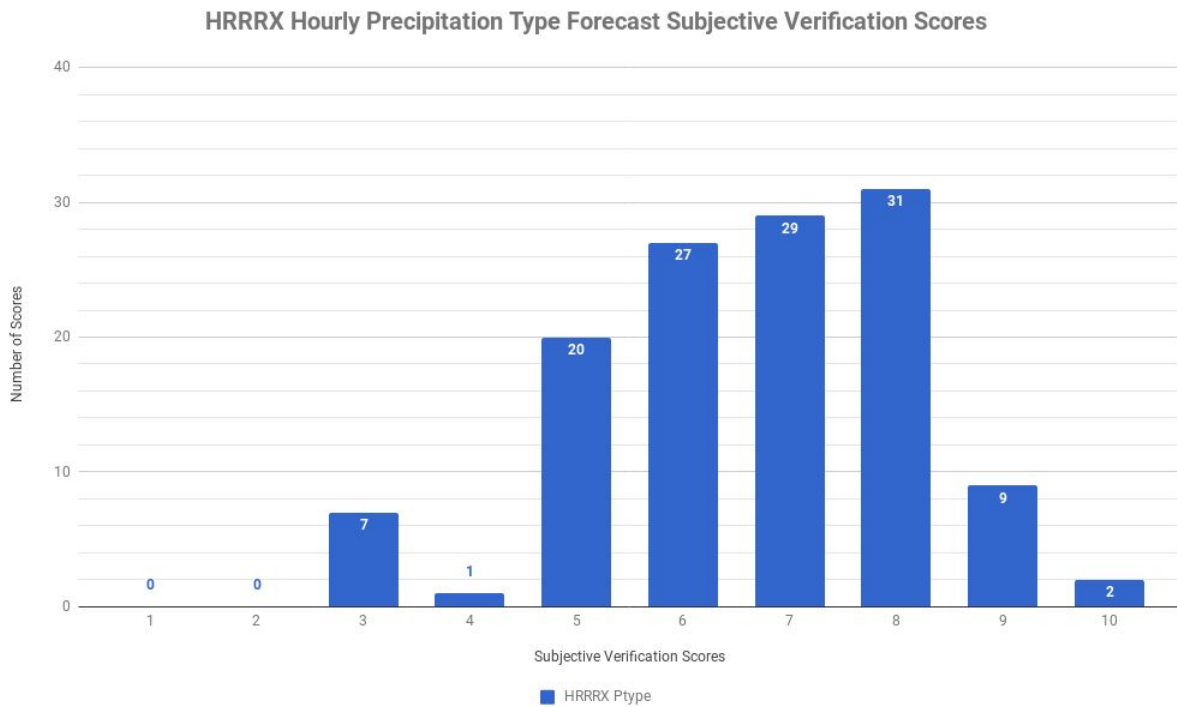


Figure 13. Individual subjective evaluation score distribution over the course of the entire experiment for the HRRRX hourly precipitation type forecasts.

Feedback

As the high average score indicates, participant feedback on the HRRRX hourly precipitation type forecasts was largely positive. One problem area that was noted by several participants was that the HRRRX precipitation type was too warm at times on the eastern side of storm systems or in cold air damming situations. Figure 14 shows an example where the HRRRX (right) has rain and mixed precipitation too far west when compared to the RAP/MRMS-GC verification (left).

Recommendation

WPC-HMT recommends continued testing of the HRRRX precipitation type algorithm to further evaluate whether there is a problem on the eastern extent of some storm systems in which the precipitation type is too warm or if the problems can be attributed to other issues such as overall timing of the storm system. Cold air damming situations should also be investigated further.

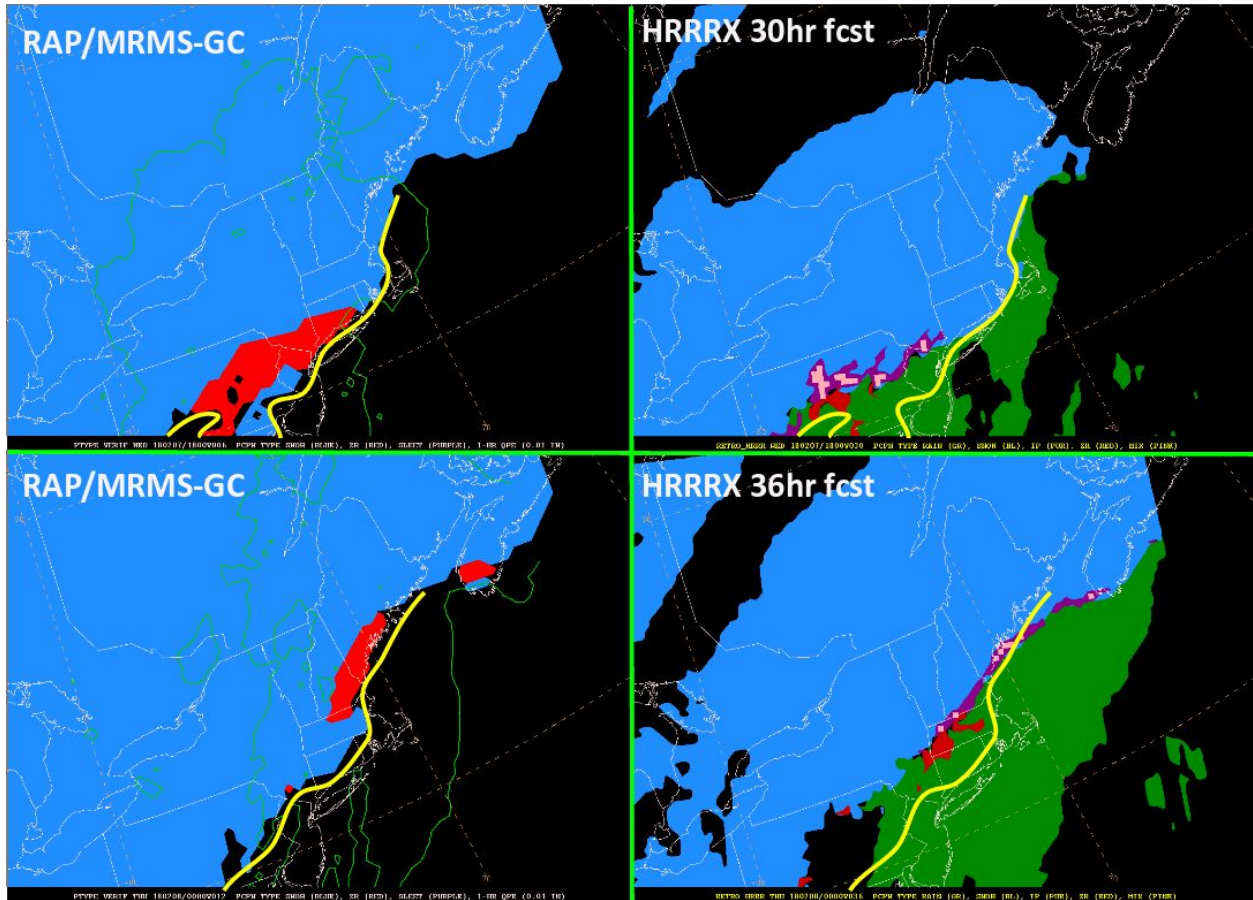


Figure 14. (Top left) RAP/MRMS-GC precipitation type verification valid at 18Z 2/7/2018 and (top right) the HRRRX 30 hour forecast for precipitation type valid at same time. (Bottom left) RAP/MRMS-GC precipitation type verification valid at 00Z 2/8/2018 and (bottom right) the HRRRX 36 hour forecast for precipitation type valid at same time. Yellow lines in all images indicate rough transition zones between snow/mixed precipitation and rain based on the verification.

HRRRX Snowfall Accumulation Forecasts

In addition to examining precipitation type from the HRRRX, the variable density snow algorithm applied to the snowfall accumulation forecast was evaluated which calculates snow and graupel depth for every model timestep and accumulates it in an hourly bucket.

For each verification session, either a Day 1 or Day 2 snowfall accumulation forecast for the HRRRX (24 hour or 48 hour forecast) was chosen based on data availability and what was determined to be a more useful case. The case was shown to the participants and compared against the NOHRSCv2 24 hour snowfall analysis for verification. Participants would subjectively rank each forecast on a scale of 1 (very poor) to 10 (very good) based on how well

the probabilistic values represented what happened and comment on the utility of each probabilistic scheme.

The HRRRX average subjective evaluation score over the course of the WWE was 6.92, with a standard deviation of 1.48. Figure 15 is a box plot for the subjective evaluation scores collected during experiment and Figure 16 shows the distribution of the individual scores given to the HRRRX by the participants. Most scores fell between 6 and 8 out of 10.

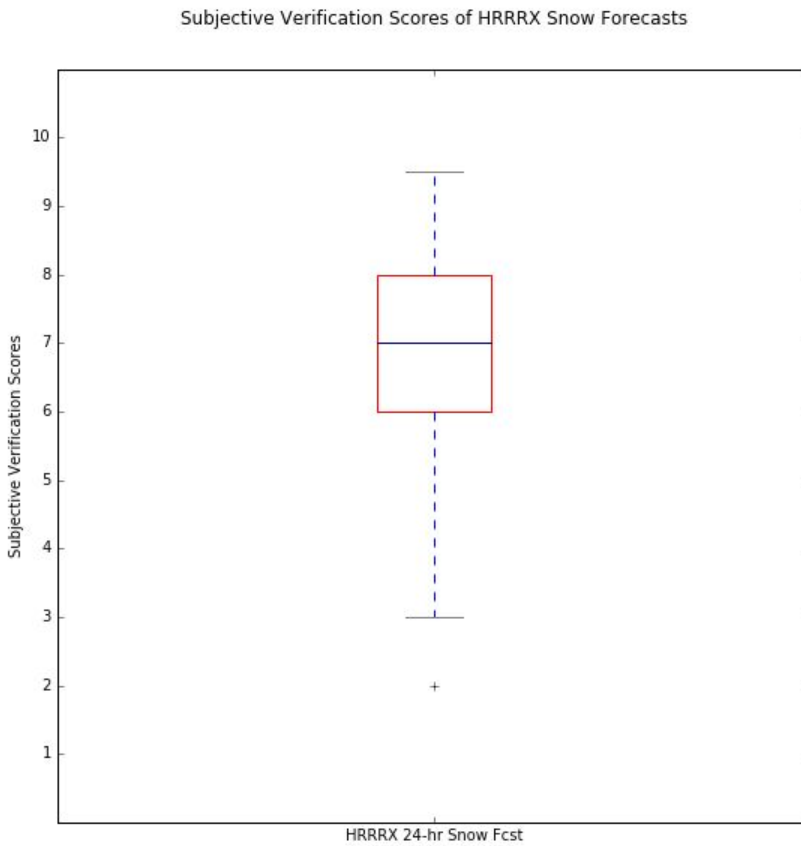


Figure 15. Box plot of the subjective scores for the HRRRX 24 hr Day 1/Day 2 deterministic snowfall using the variable density approach.

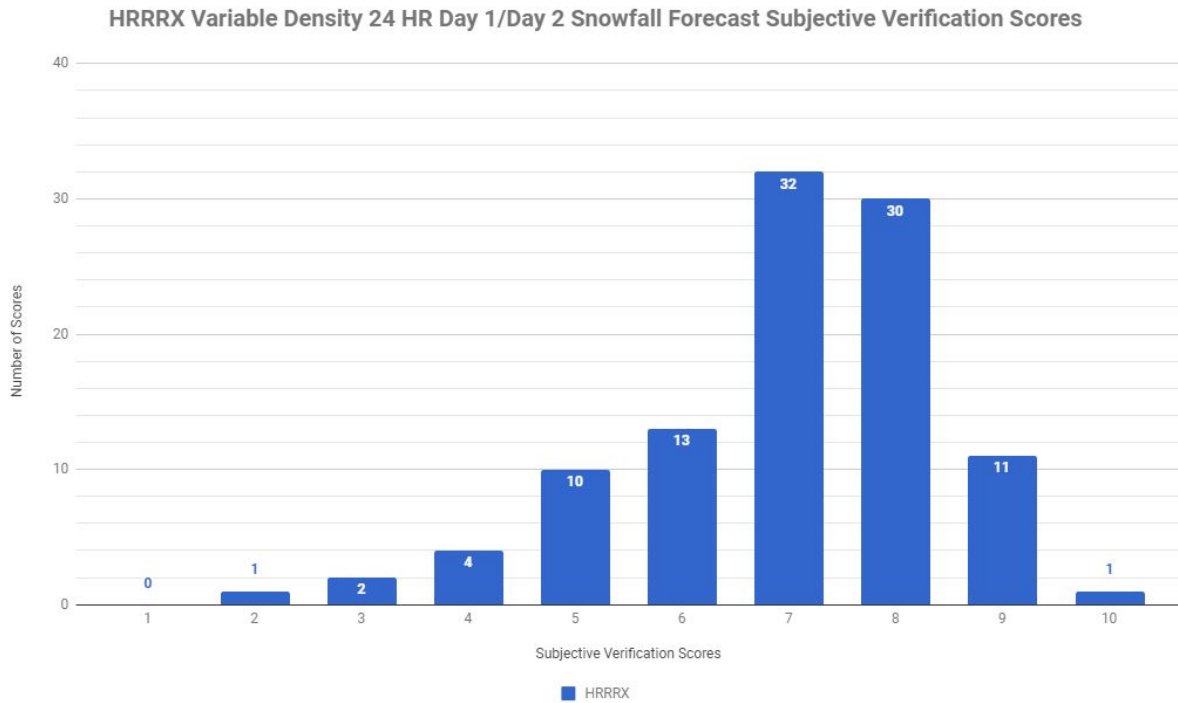


Figure 16. Individual subjective evaluation score distribution over the course of the entire experiment for the HRRRX 24 hr Day 1/Day 2 deterministic snowfall using the variable density approach.

Feedback

Participants overall rated the HRRRX very favorably as it often handled both the snowfall areal coverage and amounts associated with larger scale systems quite well. There were a couple of common issues, however. The HRRRX struggled with small-scale lake effect snow bands. Very often these features would be under-forecast and in some cases entirely missed. Figure 17 shows an example of a poor lake effect snow forecast. The coverage of the bands coming off of Lake Erie and Lake Ontario in western New York and the magnitude of the two main bands are both underdone in the HRRRX Day 2 48 hour forecast on the right of the figure. Observed snowfall off of Lake Erie totaled 6-8 inches and off of Lake Ontario 12-14 inches, however the HRRRX does not have totals higher than 4-6 inches in either band. In some instances the problems may be linked to the variable density snowfall algorithm being capped by a 13:1 SLR, shown in Figure 18. Often times these lake effect snow events can have very high SLRs which would be underforecast by the HRRRX if the SLR is higher than 13:1.

Another issue noted by participants with the HRRRX snowfall accumulation was during the late season nor'easters. Figure 19 shows the NOHRSCv2 24 hour snowfall accumulation (left) valid 12Z 3/2 to 12Z 3/3 compared with the Day 1 (24 hour forecast) 24 hour snowfall accumulation from the HRRRX (right). The HRRRX accumulations were noticeably underdone in southeast

Pennsylvania and central and southern New Jersey. They are also slightly underdone in the central New York state area as well. Surface temperatures during this system were a few degrees above freezing, especially in the the southeast Pennsylvania area. As Figure 18 shows, snow to liquid ratios within the variable density algorithm are capped at 4:1 when the lowest level temperatures are above 0°C. In this particular case, snow ratios in southeast Pennsylvania were much higher despite the above freezing temperatures. Despite the minor issues outlined above, the variable density snowfall accumulation algorithm in the HRRRX was rated favorably by the participants.

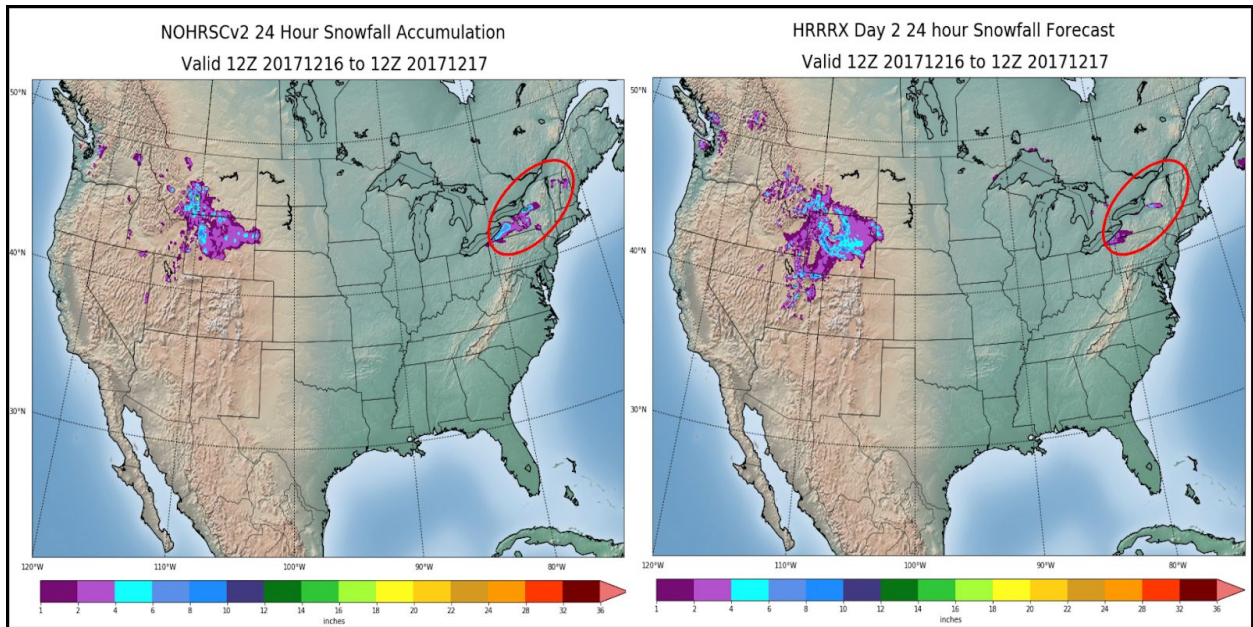


Figure 17. NOHRSCv2 24 hour snowfall accumulation valid 12Z 12/16 - 12Z 12/17/17 (left) and a Day 2 (48 hour) 24 hour snowfall accumulation forecast from the HRRRX valid over the same time period (right) with the area of western New York highlighted by the red circle.

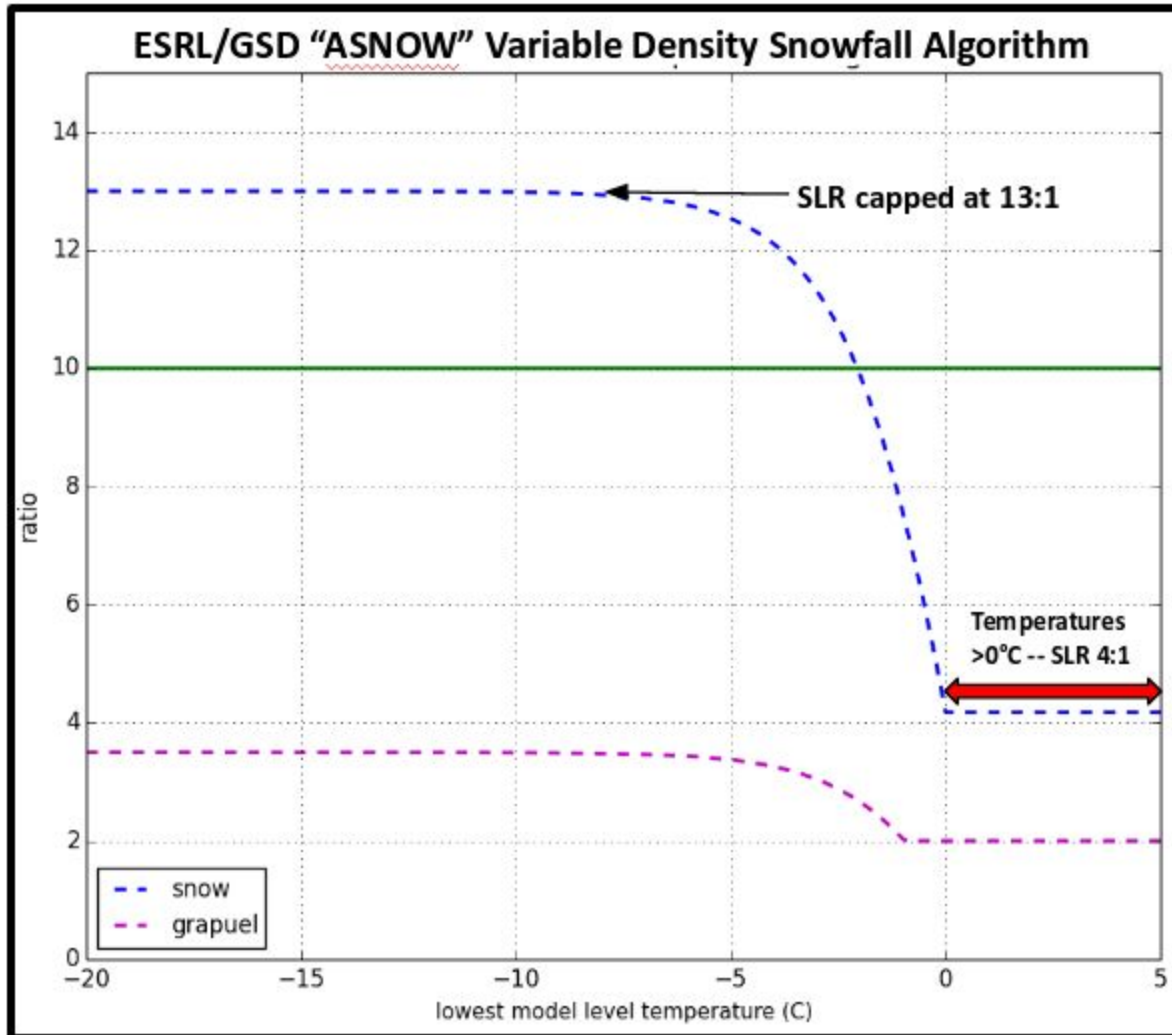


Figure 18. ESRL/GSD “ASNOW” variable density snowfall algorithm showing SLR (blue) and graupel-to-liquid ratio (purple) at various temperatures. Graph provided courtesy of ESRL/GSD. Text/arrows added by WPC-HMT.

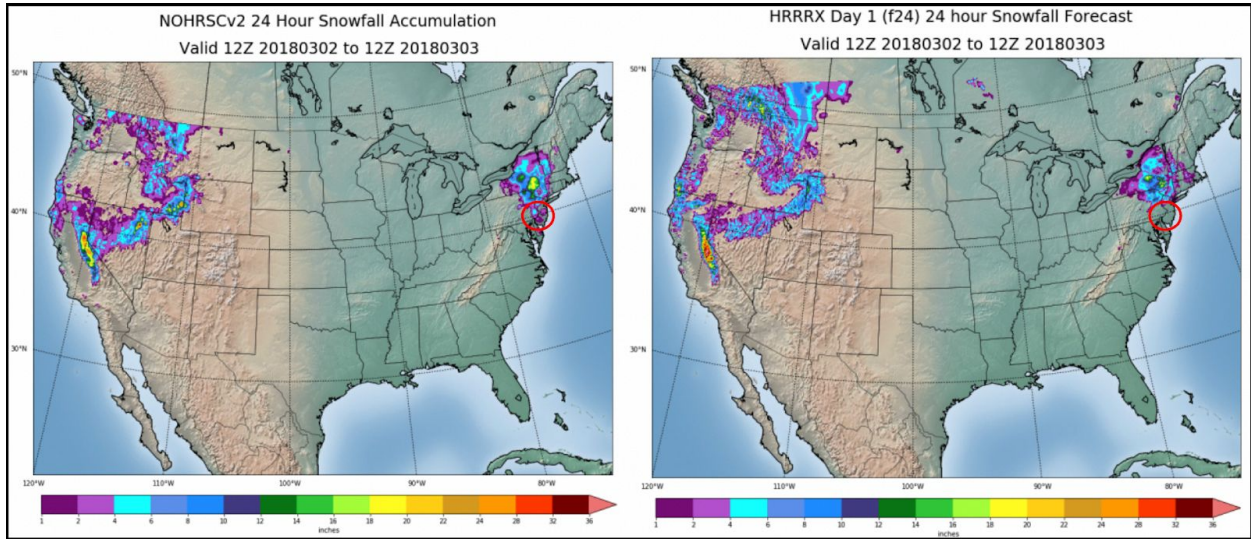


Figure 19. NOHRSCv2 24 hour snowfall accumulation valid 12Z 03/02 - 12Z 03/03/18 (left) and a Day 1 (24 hour) 24 hour snowfall accumulation forecast from the HRRRX valid over the same time period (right) with the area of southeast Pennsylvania highlighted by the red circle.

Recommendation

It is recommended moving forward that the algorithm be improved to try and capture lake effect snow events better by allowing higher SLRs than the current 13:1 maximum. Heavy snow situations in temperatures above 0°C also needs to be improved by allowing for higher SLRs than 4:1.

HRRRE Precipitation Type Forecasts

Hourly HRRRE precipitation type probability forecasts were evaluated over one six hour period. The combination RAP/MRMS-GC method was used as verification and participants were asked to only comment on the performance of the HRRRE and not give individual numeric scores, similar to how the other ensemble precipitation type methods were evaluated.

Feedback

During the first half of the experiment, spatial coverage of the precipitation type probabilities from the HRRRE was noticeably low, especially when comparing the outer extent of the lowest probability threshold to a deterministic solution from the HRRRX. Figure 20 shows an example of the HRRRX (left) precipitation type (rain (CRAIN), snow (CSNOW), freezing rain (CFRZR), and sleet (CICEP)) forecast that shows where the precipitation of each type fell during the past hour. The field on the right from the HRRRE is an instantaneous precipitation type probability field computed from CRAIN, CSNOW, CFRZR, and CICEP, plus the instantaneous precipitation rate (PRATE) for each ensemble member. Based on the feedback and examples shown during the first part of the WWE, the developers of the HRRRE (ESRL/GSD) decided to eliminate the PRATE condition from the HRRRE precipitation type calculations. This change allowed for greater areal

coverage in the HRRRE precipitation type probabilities that looked more similar to the HRRRX deterministic solution. Figure 21 shows a comparison between snow probabilities from the HRRRE with and without the PRATE threshold.

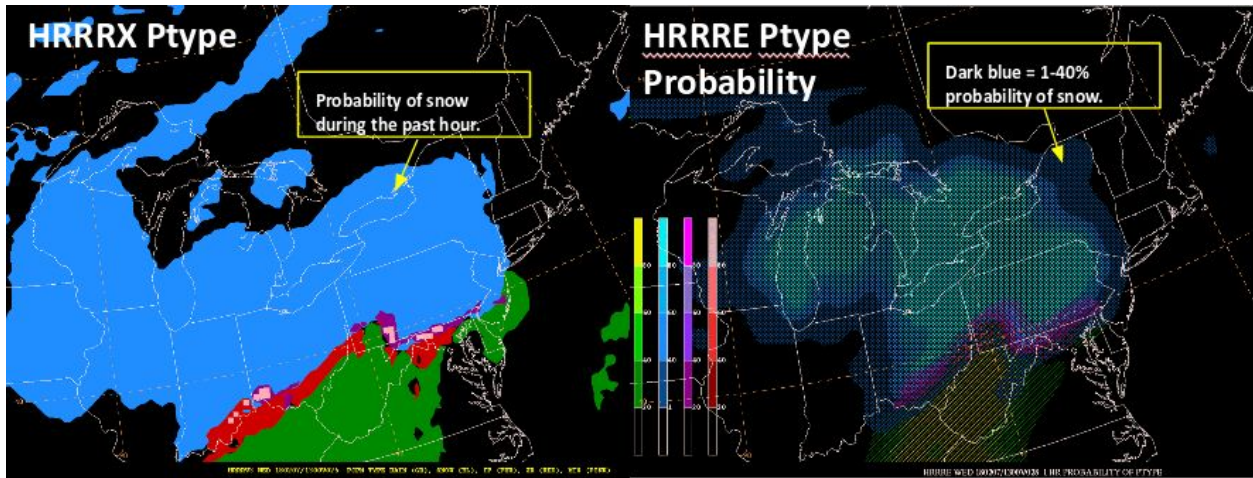


Figure 20. HRRRX hourly precipitation type forecast valid 13Z 2/7/2018 (left) and HRRRE hourly precipitation type probability valid at the same time (right).

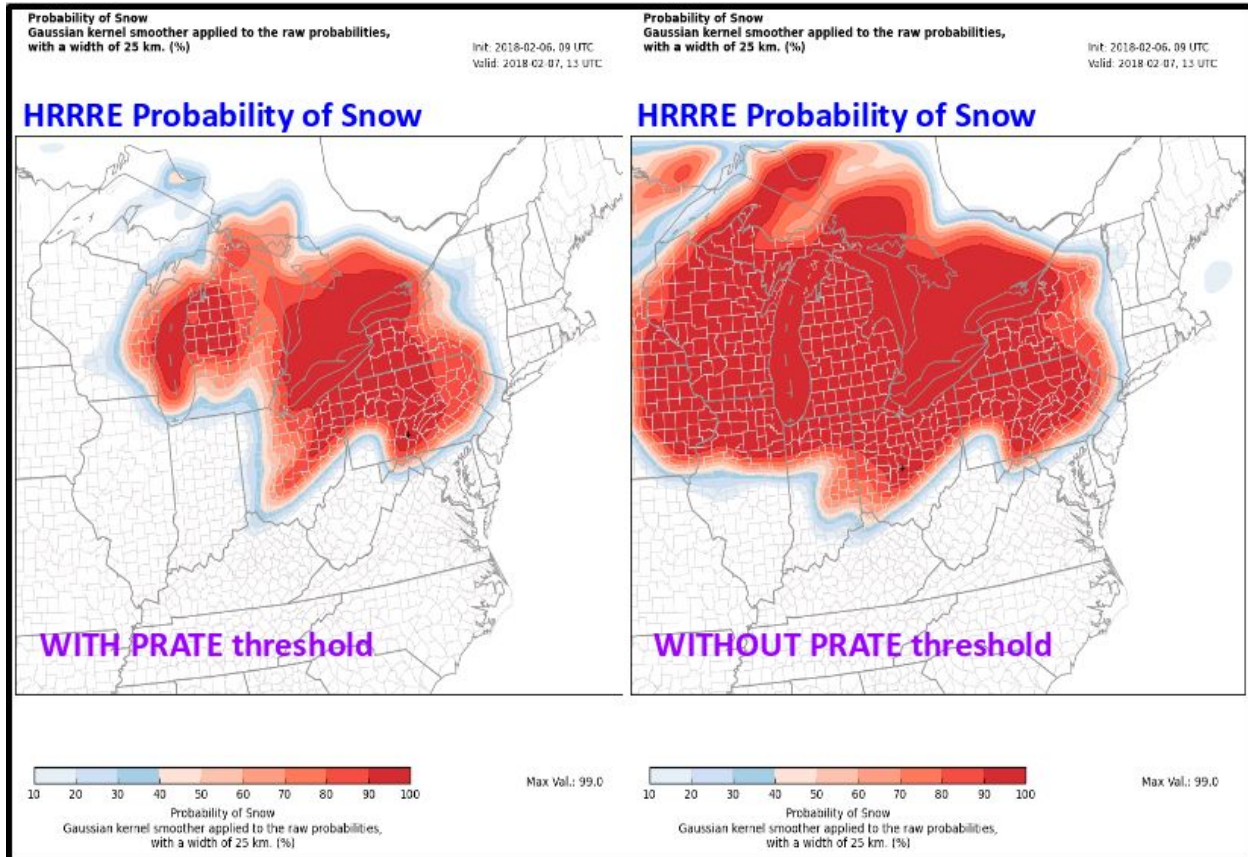


Figure 21. HRRRE probability of snow with the PRATE threshold implemented (left) and without the PRATE threshold (right) valid at 13Z 2/7/2018. Images courtesy of ESRL/GSD.

In addition to the extent of the spatial coverage, feedback indicated the HRRRE generally well-predicted probabilities of snow and rain but struggled at times with freezing rain and sleet. For some events, the HRRRE was too slow to develop high enough probabilities in areas of mixed precipitation or under-forecast probabilities which did not support the verification for sleet or freezing rain. Despite these issues, participants offered positive feedback. Forecasters favored having a probabilistic depiction of precipitation type and for most cases found the guidance to be useful to the forecast process.

Recommendation

It is recommended that the PRATE threshold not be reinstated in the precipitation type calculations as the coverage of the lower probabilities improved. Testing over the entire CONUS domain is recommended, as well as further testing and evaluation of the freezing rain and sleet probabilities.

HRRRE Snowfall Accumulation Forecasts

The HRRRE probability matched mean (PMM) snowfall accumulation forecast was evaluated. For each verification session, either a Day 1 or Day 2 snowfall accumulation forecast for the HRRRE (36 hour or 51 hour forecast) was chosen based on data availability and what was determined to be a more useful case. The case was shown to the participants and compared against the NOHRSCv2 24 hour snowfall analysis for verification. Participants would subjectively rank each forecast on a scale of 1 (very poor) to 10 (very good) based on how well the probabilistic values represented what happened and comment on the utility of each probabilistic scheme.

Figure 22 is a box plot for the subjective evaluation scores collected over the entire experiment and Figure 23 shows the distribution of the individual scores given to the HRRRE PMM snowfall forecast by the participants. The average score was 6.02 with a standard deviation of 1.59.

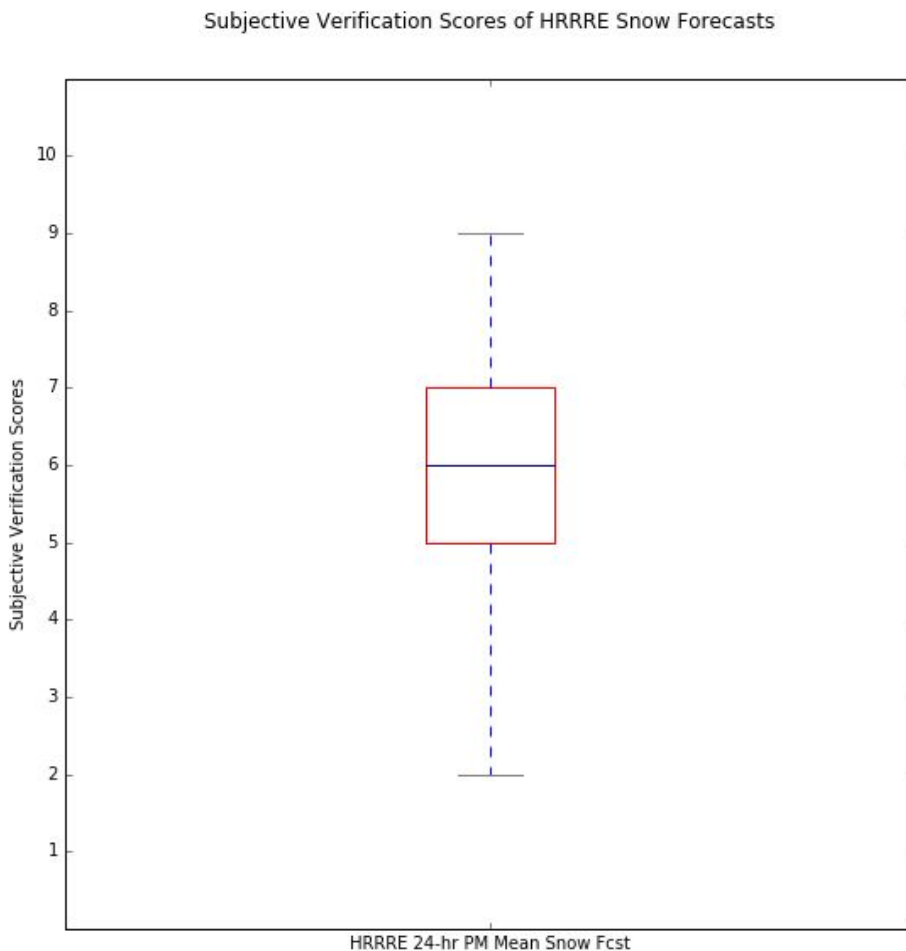


Figure 22. Box plot of the subjective scores for the HRRRE 24 hr Day 1/Day 2 deterministic

snowfall using the variable density approach.

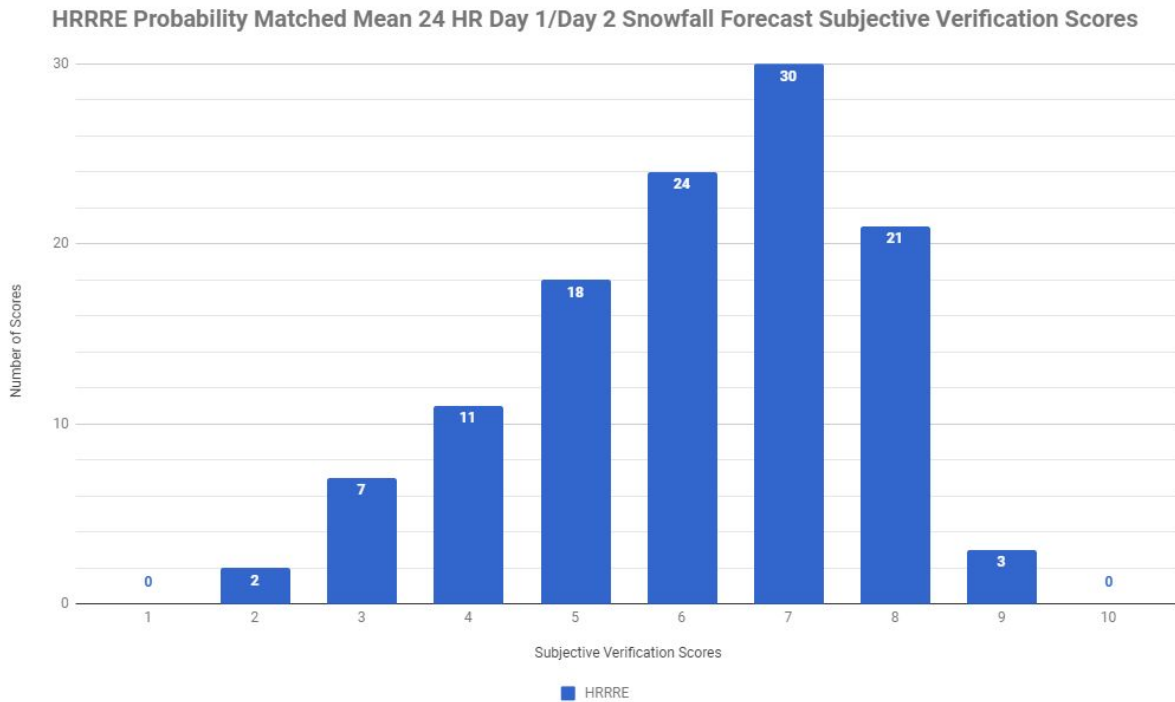


Figure 23. Individual subjective evaluation score distribution over the course of the entire experiment for the HRRRE 24 hr Day 1/Day 2 PMM snowfall forecast.

Feedback

As noted in Appendix A description of the HRRRE, the model’s domain covered roughly 55% of the CONUS. Therefore, no western mountain cases were evaluated during verification. Overall the participants had very similar feedback about the HRRRE PMM snowfall as they did about the deterministic HRRRX snowfall forecast. The HRRRE PMM snowfall was more consistently underdone when compared to the deterministic HRRRX. It also had similar problems with the lake effect snow bands as well as nor’easters that occurred with temperatures above freezing. Figure 24 is an example valid from 12Z 2/4/18 - 12Z 2/5/18 where the HRRRE PMM snowfall (right) was underdone in New England, in the Appalachian Mountains, and in the western U.S. Figure 25 uses MODE analysis to compare two HRRRE PMM snowfall forecasts to the NOHRSCv2 analysis objectively. The case on the left highlights an example where the HRRRE amounts were too light at the 4 inch threshold in Wisconsin. The case on the right highlights an example where the coverage of the 6 inch amounts from the HRRRE matched well with the analysis but the timing was off slightly, with the forecast displaced slightly east of the observed snowfall.

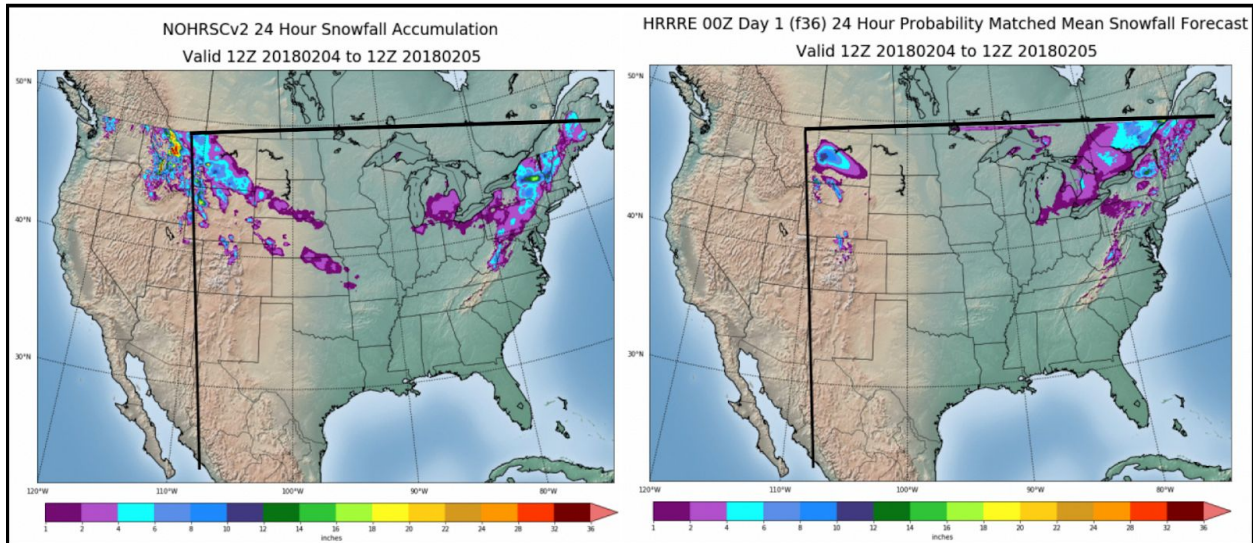


Figure 24. NOHRSCv2 24 hour snowfall accumulation valid 12Z 2/4 - 12Z 2/5/18 (left) and HRRRE Day 1 (36 hour model forecast) 24 hour PMM snowfall accumulation forecast (right) for the same time period. The black lines denote the western and northern extent of the HRRRE domain.

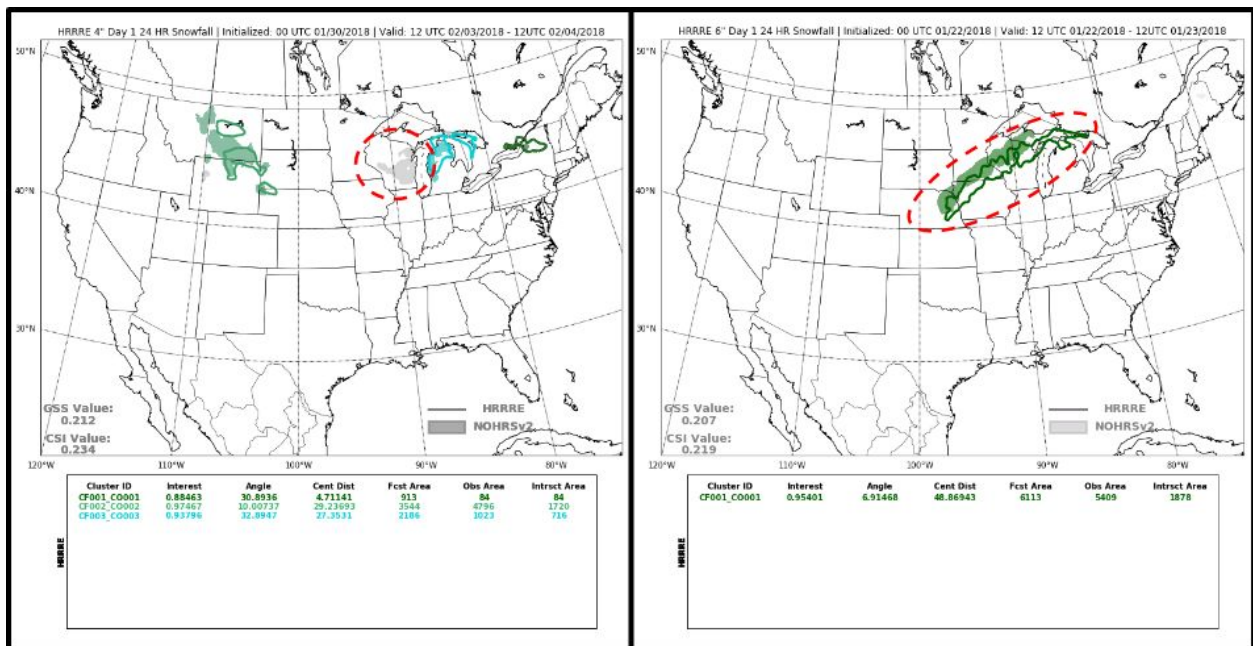


Figure 25. (Left) MODE analysis for 4 inches of snowfall with NOHRSCv2 (shaded) and HRRRE PMM snowfall (contoured) valid 12Z 2/3 - 12Z 2/4/2018. (Right) MODE analysis for 6 inches of snowfall with NOHRSCv2 (shaded) and HRRRE PMM snowfall (contoured) valid 12Z 1/22 - 12Z 1/23/2018.

Recommendation

It is recommended that the domain for the HRRRE be expanded to cover the entire CONUS in order to evaluate PMM snowfall over the more complex terrain. Improvements to the variable density snowfall algorithm recommended in the HRRRX section may benefit the HRRRE output as well.

Environmentally-Assessed Weighted Percentages and Probabilities

An instantaneous precipitation type can be derived from an environmental assessment of the atmosphere. An environmental precipitation type methodology assigns a precipitation type independent of where precipitation is being forecast. Two methodologies of environment assessed precipitation type were evaluated during the experiment. Plan view maps depicting forecasts of precipitation type using weighted percentages and conditional probabilities were used in the weekly forecast exercise, and critiqued in the verification sessions each week.

Both weighted percentages of precipitation type from the WPC Winter Weather Desk (WWD), and the conditional probabilities of precipitation type from the National Blend of Models (NBM) are derived from an average of selected multiple model and ensemble system forecast means. The resultant mean from this particular “blend” of models is also called a “blended mean.” The mean forecast solution from the blend of models is different from an ordinary ensemble mean in that certain members contribute more than others to the respective solution. A higher weight is assigned to models in the blend that are expected to help improve forecast performance. WPC forecasters vary the membership, and the weighting of the membership in each blend, at each forecast cycle. The weighting of model membership in the NBM is done algorithmically by employing a Mean Absolute Error (MAE) based weighting technique for each cycle.

Weighted Percentages

The precipitation type output from the WPC Winter Weather Desk (WWD) forecast is a weighted percentage (Figure 26). A precipitation type algorithm developed at WPC (WPC-PTA) is used to estimate an instantaneous precipitation type at each grid point on the forecast grid domain. The algorithm checks temperatures at 700 hPa, 850 hPa, and 925 hPa, as well as 2-meter temperatures to determine a species of precipitation type. A top-down vertical mask searches for the lowest level of the 0.5 degrees Celsius wet bulb temperature to assist precipitation type determination in complex terrain over the western U.S. by changing precipitation type from rain to snow where the snow level intersects the ground. This algorithm is applied to each model and member of ensemble systems selected by WPC forecasters to be part of the forecast blend for precipitation type. A weighted average is then computed for each precipitation species, and the resultant weighted percentages depicting the number of binary-determined affirmatives for each precipitation type in relation to the total

weighted ensemble for an assigned time stamp are plotted. A QPF mask is applied so that only non-zero probabilities are plotted in areas where precipitation is forecast.



Figure 26. An example of the WPC WWD weighted percentages of snow, freezing rain, and sleet valid at 00Z 3/17/2018.

Conditional Probabilities

A top down method used in the National Blend of Models (NBM) for assessing the environment for potential precipitation type. The first of 3 inputs, the maximum wet bulb temperature aloft (MaxTAloft), is determined at each grid point, and a probability of a precipitation type can be determined based on a statistical curve (Appendix A, Figure 45) derived from historically compiled sounding data to best determine the depth of the warm layer in freezing rain and sleet events (Roebber et al. 2001, Just 2017). The second input is extracted from a probability of refreeze to sleet curve, and the third input is the probability of cloud ice present in a layer based on the environmental temperature. These three inputs for assessing environmental potential for precipitation type are applied to each member of the model blend, then averaged over the domain of models to generate a forecast condition probabilities for precipitation type (Figure 27). No QPF mask is applied to the NBM precipitation type probabilities.



Figure 27. An example of the NBM conditional probabilities of snow, freezing rain, and sleet valid 00Z 3/17/2018.

Participants correctly discerned that the instantaneous conditional probabilities from the NBM, and the weighted percentages from the WPC WWD-PTA were not actual probabilities of

occurrence of a specific precipitation type. This often created ambiguity in applying the fields to the forecast exercise. The broad areal coverage of these probability fields and the lack of gradient (binary in appearance), particularly from the NBM, due to the conditional assessment of the environment, made it difficult to incorporate into the forecast process.

Additionally, the high likelihood of more than one precipitation type for the same time period from the NBM, which is typically the case in the physical atmosphere, was noted frequently by the participants. This overlap made it difficult to determine an instantaneous precipitation type, and assign an average type over the 6-hour period being forecast. This also required forecasters to look at other parameters such as MaxTAloft or fraction of frozen precipitation in order to assign a definitive precipitation type. The WPC WWD probabilities were typically preferred over the NBM due to greater precision in areal coverage and weighted probabilities more representative of the environment.

Feedback

The NBM conditional probabilities demonstrated utility in the Day 3 time range, however, in the shorter time scales, the probabilities tended to be too low and missed snow, freezing rain and sleet in areas over which it was observed. This was particularly the case in areas of complex terrain and near coastlines. It was noted that this could be a result of the decision algorithm in the NBM where snow is assigned for temperatures $< 34^{\circ}\text{F}$, an even mix of rain/snow for temperatures $34\text{--}37^{\circ}\text{F}$, and all rain for temperatures $> 37^{\circ}\text{F}$, which often failed in areas with variable terrain or the influence of bodies of water (Pacific northwest, Northeast coast, Great Lakes region).

In highly dynamic environments, this surface based algorithm may be negating the influence of cooling layers aloft and enhanced precipitation rates. Another potential contributor to precipitation type misses in the short term could be the struggle of high-resolution models used in the blend to correctly predict of the 2-meter temperature. The models struggle to retain the cold air or scour out the cold air appropriately at the surface. Areas of sleet and freezing rain tended to be too narrow, particularly sleet which exclusively presented along the periphery of the snow probabilities. Moreover, differences in precipitation rate not accounted for in the precipitation type methodology may be resulting in the wrong precipitation type being identified in thermal environments that could support either snow or freezing drizzle. Finally, dry air intrusion aloft which adjusts the wet bulb temperature may affect the NBM precipitation type algorithm resulting in “false alarm” areas.

Recommendations

It is recommended that MDL continue to work with WPC-HMT to improve some of the precipitation type predictors. The existing decision algorithms for wet bulb temperature aloft and surface temperature should be reevaluated for adjustment. The modified MaxTAloft (Appendix A, Figure 45) approach that is over-predicting sleet on the periphery of the predicted

snow should be explored, as well as false alarms for freezing rain related to 2-meter temperature errors in the higher-resolution models.

NBM3.1 Snowfall

NBM v3.1 snow is initiated from the hourly or 6-hourly QPF with a blended snow ratio applied. Snow level is calculated using a blend of 0.5°C wet bulb heights from a number of models. The heights are calculated first then blended with the NBM snowfall forecast.

The participants were presented with the NBM3.1 deterministic snowfall forecast. The average participant rating of the NBM snowfall, on a scale from 1 to 10, was 6 (Figure 28). As seen in the box plot of Figure 29, the highest score was 8 and the lowest was 3. Most of the scores over the experiment were between 5 and 7.

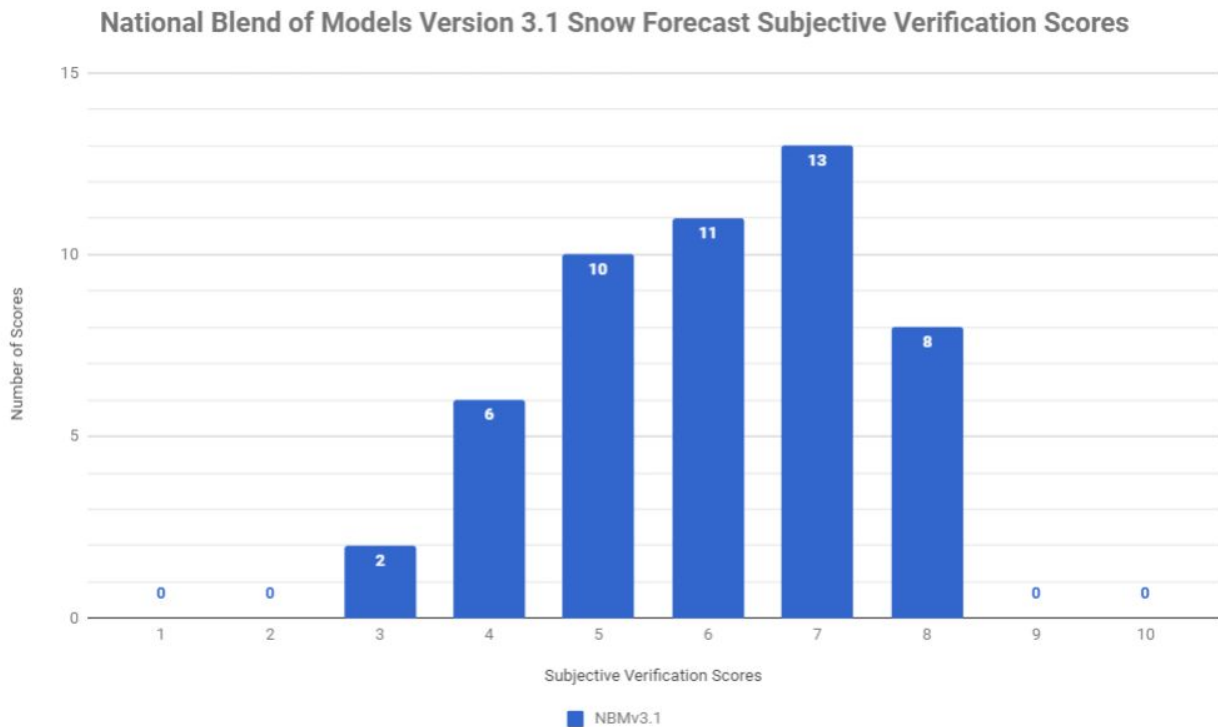


Figure 28. Participant scoring results for the NBM snowfall accumulation forecasts, between 1 (poor) and 10 (excellent).

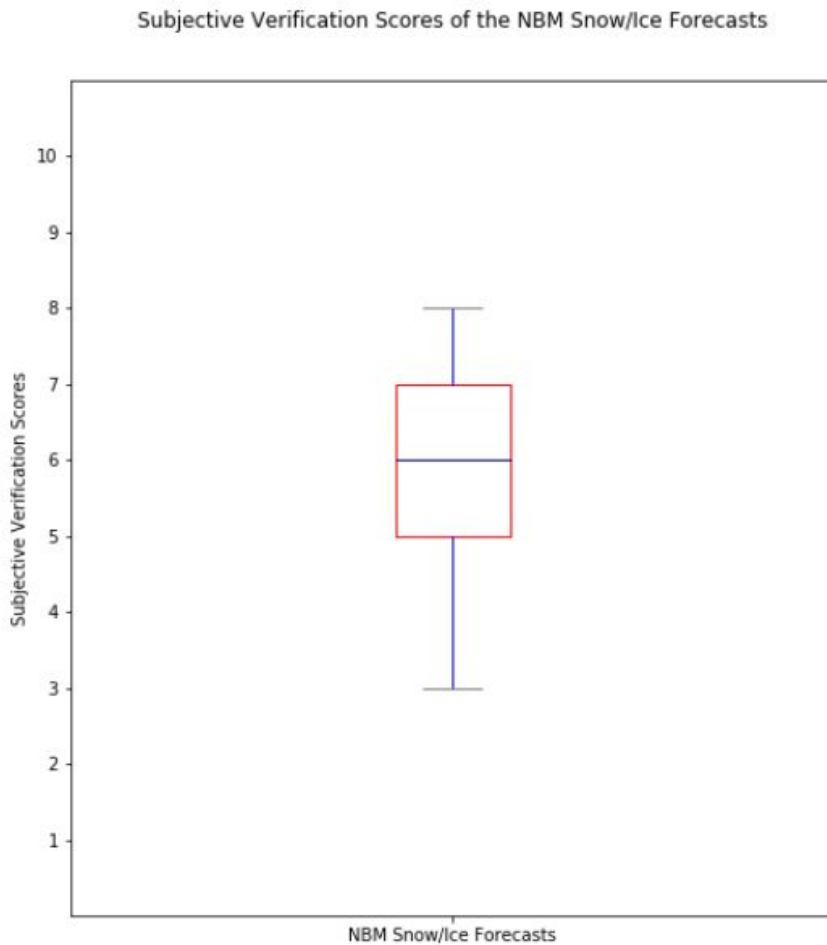


Figure 29. The box plot showing the distribution of scoring results for the NBM precipitation type forecasts, between 1 (poor) and 10 (excellent).

Feedback

The participants often commented that the NBM snowfall was underdone and missed many lighter or more marginal snowfall events. Additionally, the NBM snowfall missed large accumulating “warm snow” events, some producing 6 inches or more around the Great Lakes and in the northeast, due to the current temperature profile constraints to which is referred in the previous section. The March 7-8th, 2018 example in Figure 30 shows an event in the northeast for which the NBM was predicting precipitation but was too warm and failed to produce snow from Philadelphia, southern New Jersey, and along the seaboard through Boston. Of particular interest is the Boston area where the temperature during most of the event was 34 or above with accumulating snow.

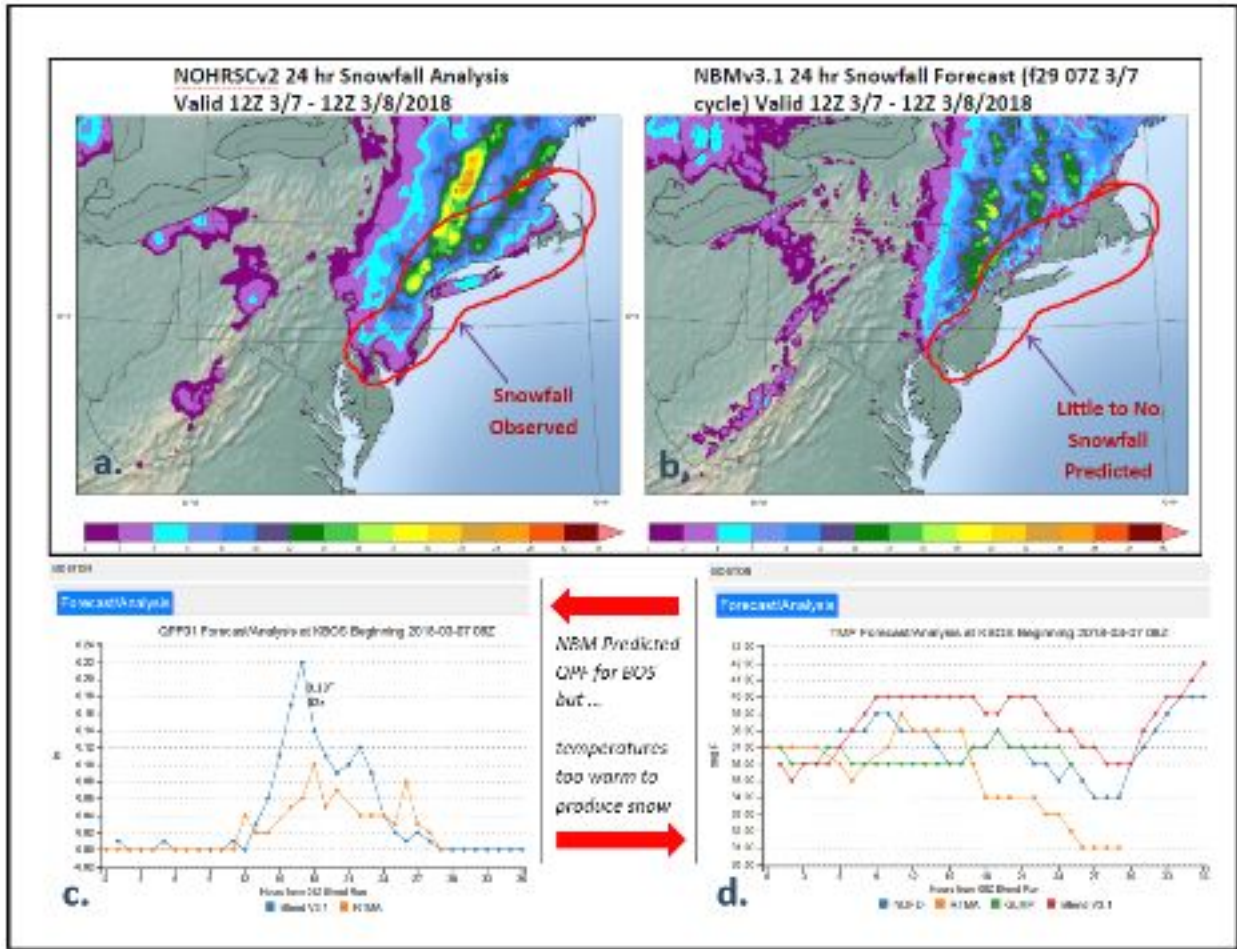


Figure 30. a. NOHRSCv2 24-hr snowfall analysis valid 12Z March 7th - 12Z March 8, 2018; b. NBMv3.1 24-hour snowfall forecast valid 12Z March 7th - 12Z March 8, 2018; c. 1-hour precipitation analysis (RTMA) in orange and NBMv3.1 QPF in blue for Boston, MA valid 08Z on March 7th, 2018; d. 1-hour temperature analysis (RTMA, orange) and predicted temperature from the NDFD (blue), GLMP (green), and NBM (red) valid 08Z on March 7th, 2018 for Boston, MA.

Recommendations

It is recommended that the NBM adjust calculations to raise the magnitude of snowfall events, particular those of lower density. Errors mentioned previously regarding temperature constraints which cause the NBM to not produce accumulating snow within a warmer temperature profile should be improved in future builds.

Environmentally-Assessed Ensemble Probabilities

Fuzzy Clustering

The application of fuzzy clustering was tested to assess predictability of precipitation type from the perspective of storm track. The fuzzy clustering methodology developed at Stony Brook University (Zheng et al. 2017) partitions a multi-model ensemble consisting of 20 individual members from the Global Ensemble Forecast System (GEFS) run at the National Center for Environmental Prediction (NCEP), 20 members from the Canadian Meteorological Center Ensemble (CMCE), and 50 members from the European Center for Medium-Range Weather Forecasts' Ensemble System (ECENS) into one of five clusters derived from principal component (PC) analysis of mean sea level pressure (MSLP).

Daily Derivation of Clusters

At 0000 UTC for each day of the forecast cycle, an Empirical Orthogonal Function (EOF) analysis was applied to the ensemble sensitivity of the multi-model dataset to determine the dominant spatial patterns of model uncertainty in the ensemble forecast of MSLP over three prescribed regions (East Coast, Central U.S., and Eastern Pacific). The EOFs are computed across the model ensemble member dimension, and the resulting modes show the dominant patterns of the difference between individual ensemble members and the 90-member ensemble mean. The first and second PCs for the 90 ensemble members are used as input into a fuzzy clustering routine, which is used to group ensemble members with similar forecast scenarios. The two leading PCs are the projections of the dominant EOFs onto the difference between each of the ensemble members and the 90-member mean. Since the leading PCs and the associated EOF patterns contain the main uncertainty information across the entire ensemble, they are used as the basis for performing the cluster analysis.

An example of the EOF analysis applied to a Day 2 forecast for the March 8, 2018 Nor'easter is shown in Figure 31 below. The forecast cycle is from 0000 UTC March 6, for 0000 UTC March 8. Plotted in Figure 31a is the 90 member mean of MSLP in the solid black contours, and the spread of MSLP in shaded blue. Figure 31b depicts EOF1 of the MSLP which explains approximately 36% of the ensemble variance. The overlap of the gradient of the negative sign of EOF1 with that of the ensemble spread in Figure 31a suggests that EOF1 represents the placement of the cyclone center. So ensemble members that fall into a cluster aligned with the positive sign of EOF1 will feature a cyclone closer to or along the New Jersey coast. Conversely, ensemble members that align with a cluster in the negative sign of EOF1 will likely feature a cyclone center further away from the coast. EOF2 in Figure 31c, which explains approximately 19% of the ensemble variance, also suggests differences in placement of the cyclone. However the axis of the EOF sign is more orthogonal to the ensemble spread axis, and captures less of the MSLP differences over the Great Lakes.

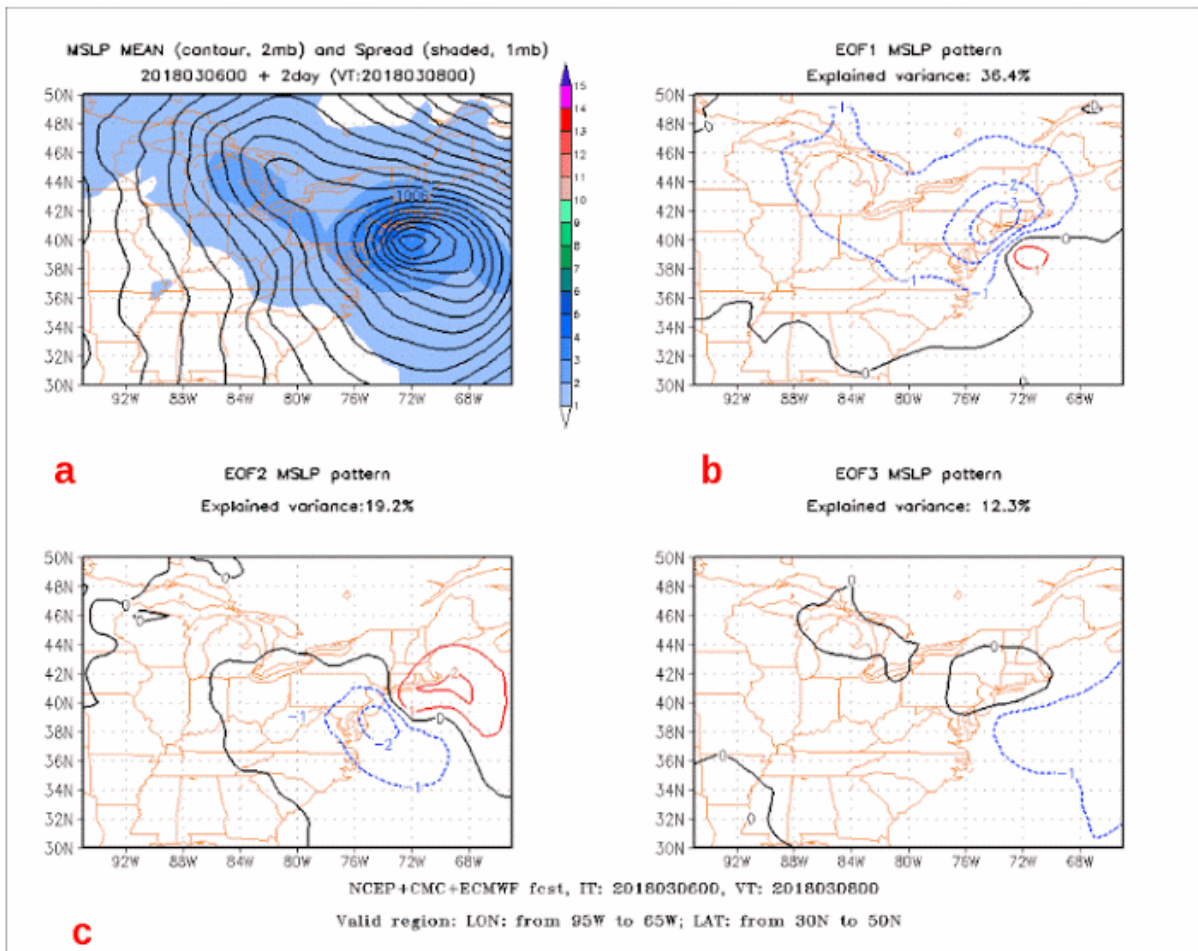


Figure 31. EOF analysis applied to a Day 2 forecast for the March 8, 2018 Nor'easter where (a) is the 90 member mean of MSLP (black contours) and spread (shaded blue), (b) depicts EOF1 of the MSLP, and (c) depicts EOF2 of the MSLP.

The partitioning of the clusters with respect to the EOF and sign they align with are plotted in Figure 32. Cluster 1 is always closest to the 90 member mean (black dot on origin) and has a diverse mix of ECENS, GEFS, and CMCE membership. Cluster 2 is aligned with positive EOF2 (as shown in Figure 31c) and consists of mainly GEFS and CMCE members. Cluster 3 also contains mostly GEFS members with a few CMCE, and is aligned with positive EOF1 (shown in Figure 31b). Cluster 4 is mostly ECENS and a few CMCE members, and aligned with negative EOF1, while Cluster 5 is all ECENS members and negative EOF2.

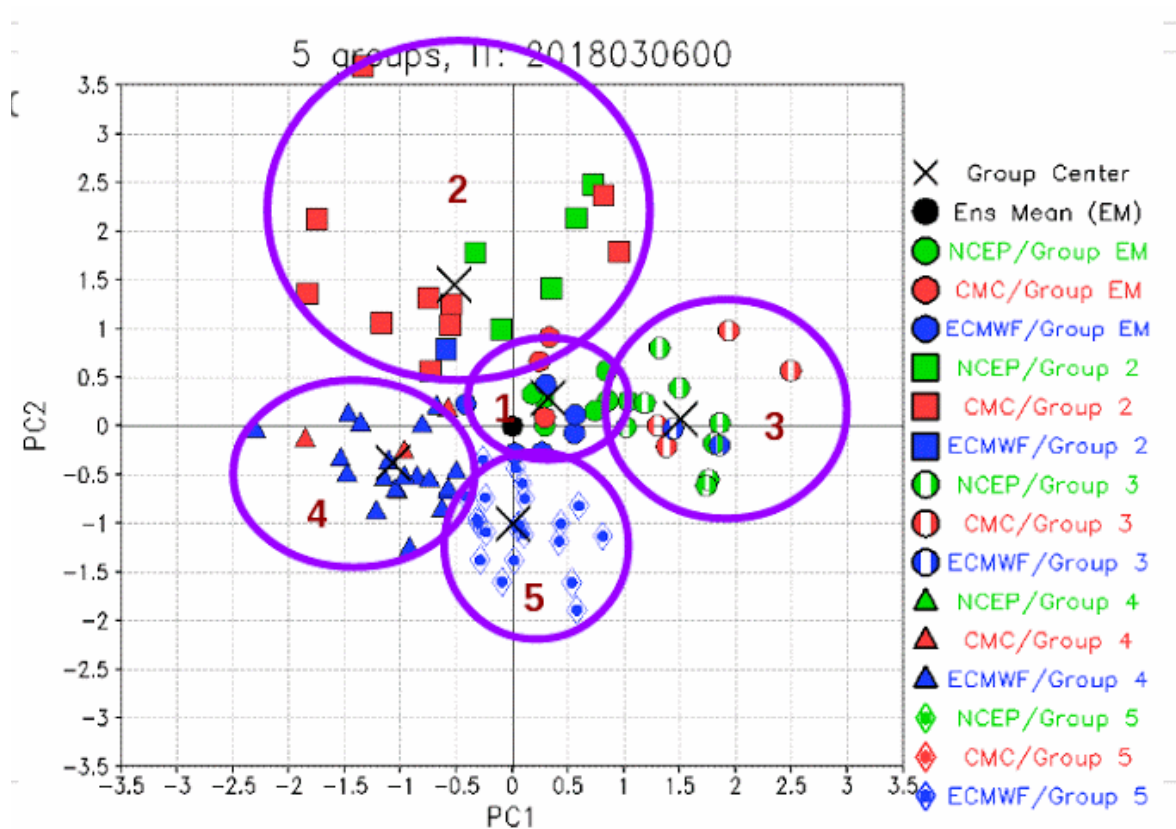


Figure 32. The partitioning of the clusters (delineated by red numerals and contained in purple) with respect to the EOF and sign with which they align: EC (Blue), GEFS (Green), CMCE (Red).

Application to Forecasting Precipitation Type

The cluster analysis can be more intuitively assessed by depicting 850 hPa temperatures. During each forecasting session, an image similar to Figure 33 was shown to illustrate the cluster partitions based on the location of 850 hPa 0°C temperatures. In each cluster, the 90 member mean 850 hPa 0°C line was depicted, then overlaid with the 850 hPa 0°C isotherms from members assigned to the cluster. Focusing in on the forecast area of concern, forecasters can identify which clusters are running warmer or colder than the 90 member ensemble mean.

For example, we can isolate the Philadelphia to New York City corridor. The experimental tools for precipitation type, as well as the operational guidance, suggested rain, or mixed precipitation type over this area during March 8. Upon looking at the 850 hPa temperature analysis for Clusters 4 and 5, it is apparent that the cluster member 850 hPa 0°C isotherms are

colder than the 90 member mean off the New Jersey coast, thereby implying a potentially colder sounding from Philadelphia to New York. Referring back to Figures 31a and 31b, and the negative sign of EOF1 and EOF2 respectively, one can see the signal for higher MSLP inland thereby implying a faster cyclone and positioned further offshore.

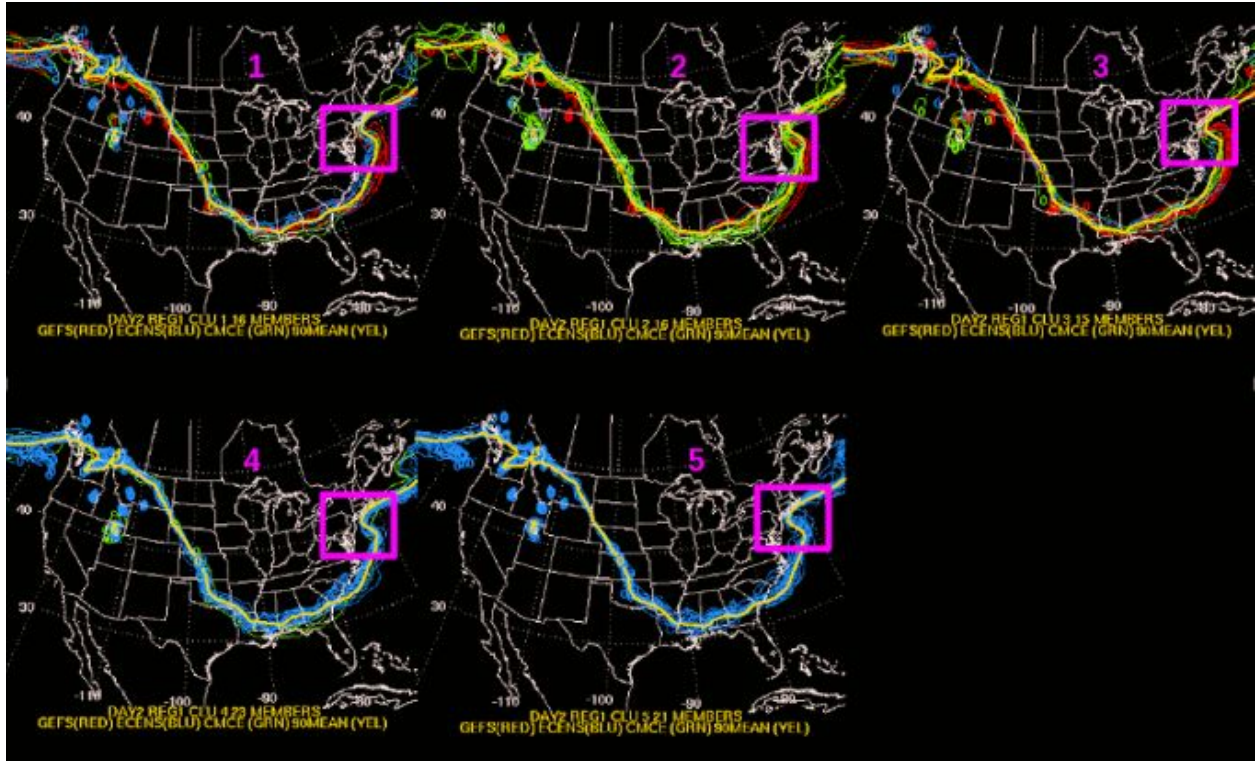


Figure 33. An example of cluster partitions based on the location of 850 hPa 0°C temperatures.

The WPC-PTA, computes an estimated, instantaneous precipitation type for each of the aforementioned 90 ensemble members prior to the clustering. After the ensemble members and their assigned precipitation type were partitioned into their respective cluster, an ensemble probability of snow, sleet, and freezing rain was computed. Plots of the probabilities were made available to forecasters to ascertain spatial differences in precipitation type between the clusters.

Figures 34a and 34b depict the ensemble probability of snow for the warmer Clusters 2 and 3. Note 40-60 percent probabilities of snow over the Philadelphia area with these two clusters. However, the colder Clusters 4 and 5 are plotted in Figures 34c and 34d where the ensemble probability of snow increases to the 70-90 percent range.

During this event, there was a quick change from rain to heavy snow over this densely populated area, and the cluster analysis was a guide for forecasters to trend towards a colder solution with the snow line farther east and through the two large cities.

Feedback

There was consensus among experiment participants that the use of clustering techniques shows promise in improving the forecast process by providing information on ensemble sensitivity, and WFO forecasters expressed interest in and having access to this data displayed in a web site. There was a unified desire to have the tool more efficiently expedite the ingestion and interpretation of the ensemble data through more intuitive displays and interactive capabilities (such as zooming in and out and interrogating the data). The 850 hPa 0°C lines were often the preferred field for predicting precipitation type. Also requested was extensive training and documentation on how to better apply clusters to the forecast process.

A limitation of the fuzzy clustering is that it is surface cyclone dependent, with the best use for significant cyclogenesis and ensuing storm track scenarios. It was difficult to apply the tool in more benign MSLP situations. The tool also does not provide information on which synoptic features are leading to the uncertainty, focusing more on the presence of the groupings.

The cluster showed good utility over the central and eastern U.S. but not so over the high terrain of the western U.S. The high variance in elevation made PC identification from MSLP EOFs not very useful in this region.

Recommendations

WPC-HMT recommends the development of training material to enable increased use among forecasters as well as exploring ways to more intuitively visualize and interrogate the data in the display. It is also recommended to experiment with clusters derived from other large scale mass fields other than MSLP to address the terrain issue, and allow application in forecast scenarios other than large cyclone events.

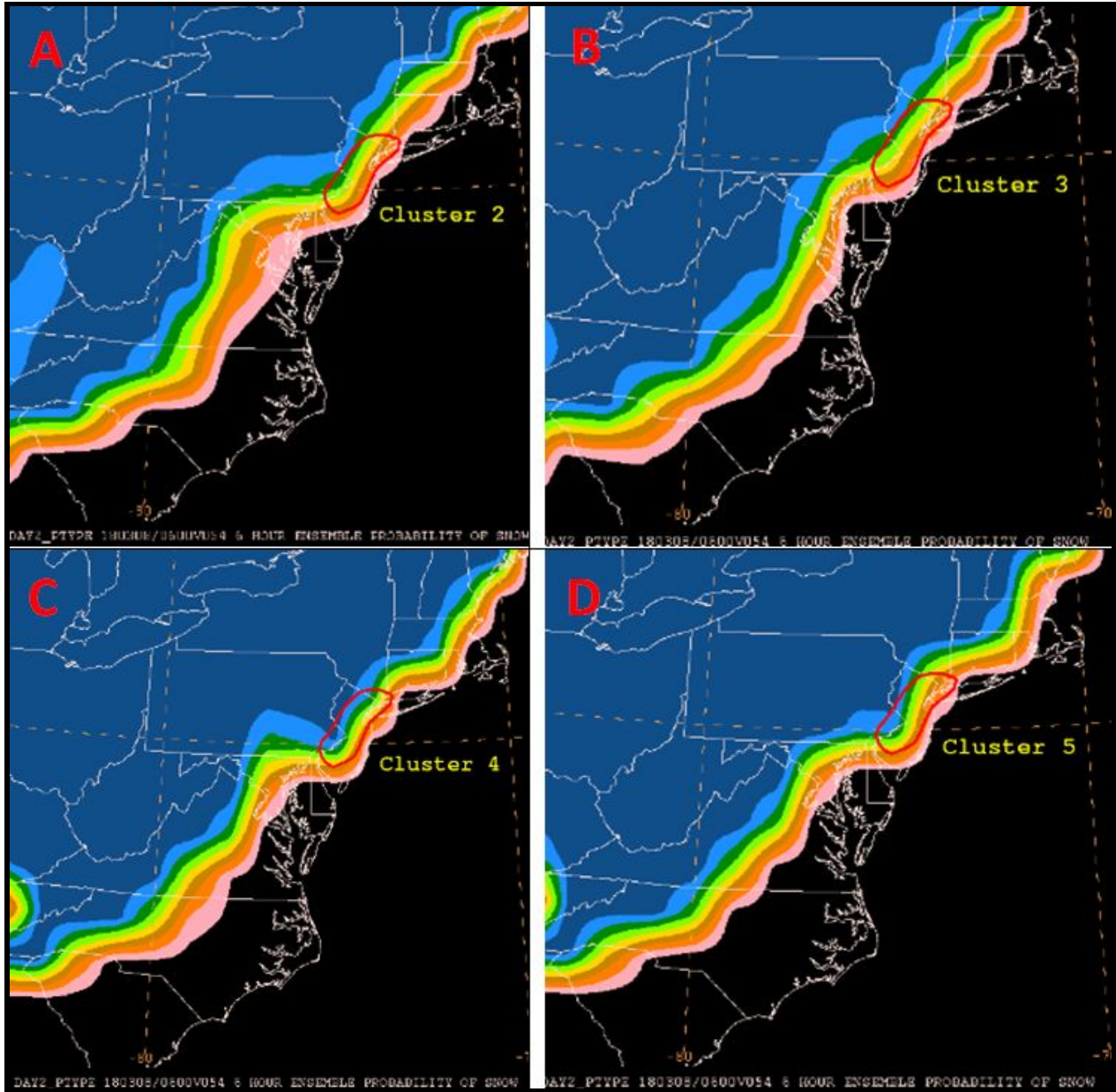


Figure 34. (A) Ensemble probability of snow for the warmer Cluster 2; (B) Ensemble probability of snow for the warmer Cluster 3; (C) Colder Clusters 4 showing the ensemble probability of snow increasing to the 70-90 percent range; (D) Colder Clusters 5 showing the ensemble probability of snow increasing to the 70-90 percent range.

HREFv2 Point/EAS Precipitation Type Probability Methods

Two different probability methods for forecasting precipitation type were evaluated and scored from the HREFv2 ensemble system throughout the experiment. The first method was the

traditional point probability method where an ensemble probability is determined at each grid point. The second method was based upon ensemble agreement scale (EAS) similarity criteria outlined in Dey et al. (2016) that uses a varied neighborhood radius size between 10 km and 100 km according to member to member similarity criteria. This method was referred to as the “EAS probability scheme” in the experiment. More details on the two probability methods can be found in Appendix () under the HREFv2 description.

For each verification session, the point and EAS probabilities from the HREFv2 for either snow or freezing rain precipitation type, depending on the case and which was determined to be a more useful case study, were shown to the participants. As verification, the RAP/MRMS-GC QPE and precipitation type algorithm were used. Participants would subjectively rank each scheme on a scale of 1 (very poor) to 10 (very good) based on how well the probabilistic values represented what happened and comment on the utility of each probabilistic scheme.

Both the EAS and point probability method had very similar overall average subjective evaluation scores. The EAS method had an overall average score of 6.58 and a standard deviation of 1.28. The point probability method had an overall average score of 6.49 and also had a standard deviation of 1.28. Figure 35 is the box plot for the subjective evaluation scores collected over the entire experiment and Figure 36 shows the distribution of individual scores for each probability method.

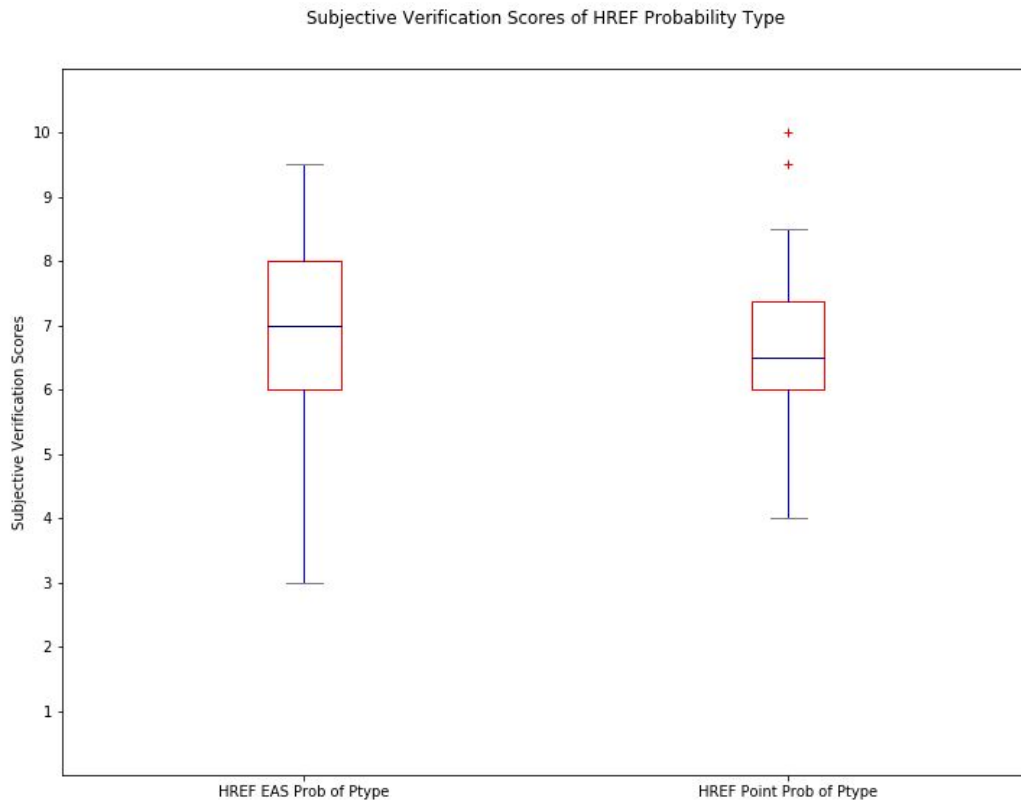


Figure 35. Box plot of the subjective scores for the HREFv2 EAS and point probability methods for precipitation type.

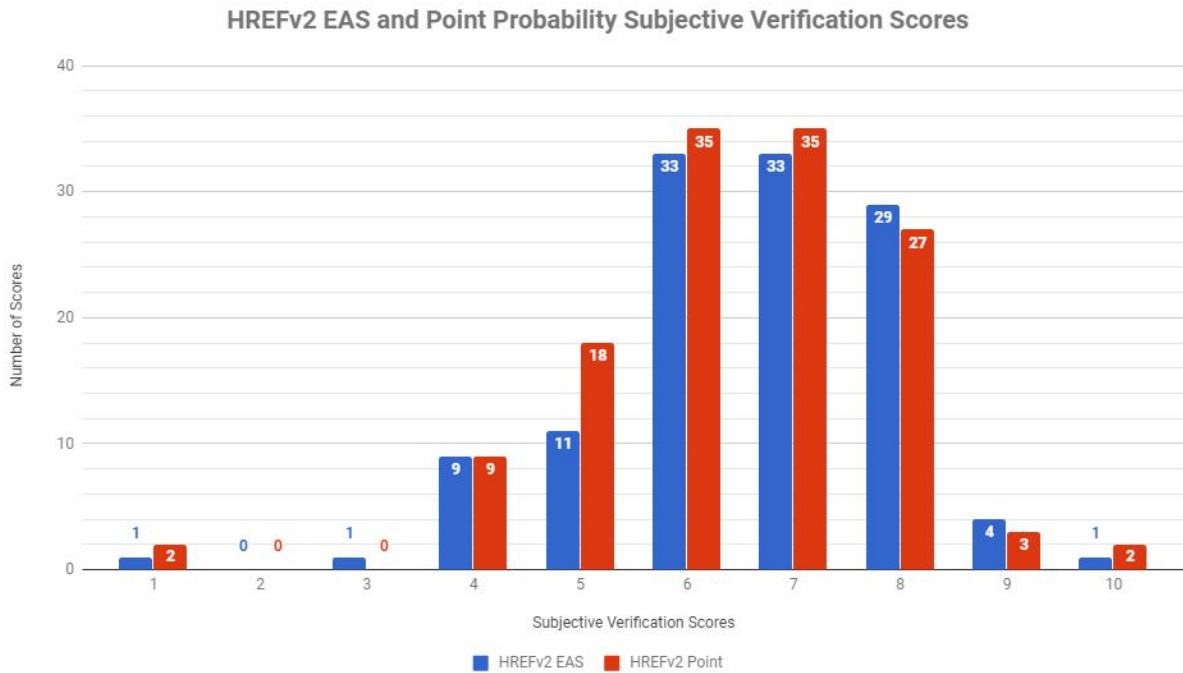


Figure 36. Individual subjective evaluation score distribution over the course of the entire experiment for the HREFv2 EAS and point probability methods for precipitation type.

Findings

The precipitation type point and EAS probabilities from the HREFv2 generated a lot of discussion each verification session. Many forecasters and participants preferred the detail that the point probabilities provided. The detail was often preferred in areas of complicated terrain such as the mountains and valleys in New England, over the Appalachian Mountains, or the many mountain ranges out in the Western U.S. In these cases, the EAS probabilities were too smooth and had probabilistic values in unrealistic areas. Figure 37 shows one of these examples where the EAS probabilities in (B) have low end snow precipitation type probabilities extending to the coast of southern California where the point probabilities in (A) have more realistic zero values for snow and the higher probabilities confined to the mountains in that region.

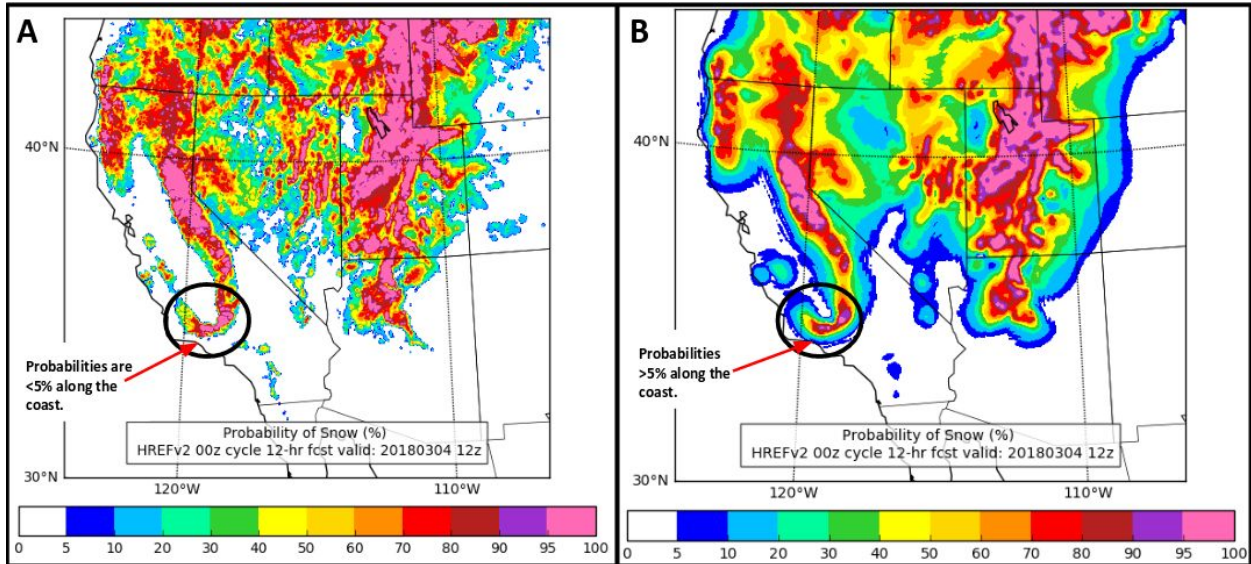


Figure 37. (A) 12 hour point probabilities for snow from the 00Z HREFv2 valid at 12Z 3/4/2018 and (B) 12 hour EAS point probabilities for snow from the 00Z HREFv2 valid at 12Z 3/4/2018. Images courtesy of Ben Blake/EMC.

Many of the forecasters and participants preferred the EAS method as a way to better communicate the precipitation type probabilities in a decision support services (DSS) role as opposed to the point probabilities due to the decreased levels of “noise” in the data. Within the EAS calculation, there is a similarity criteria parameter (alpha) that can be tuned to alter the sharpness of the probabilities. Alpha is related to the amount of bias tolerated in the forecast, where an alpha = 0 signifies no bias is tolerated at the grid-scale and an alpha = 1 signifies any bias is tolerated. Consequently, for a smaller alpha, it is more difficult to satisfy the similarity criteria at smaller radii, and vice versa. During the WWE, alpha was set to 0.5 for the EAS method and this was based on feedback received during the 2017 Flash Flood and Intense Rainfall Experiment (HMT-WPC, 2017). Figure 38 shows an example of different alphas and how the EAS probabilities change. Figure 38B is an example with alpha set to 0.5, which is the value that was used during the experiment. Figure 38C increase the alpha to 0.75. In Figure 38D, the alpha is set to 0.90. Apparent in the different images, increasing the alpha closer to 1 makes the EAS probabilities look more similar to the point probabilities (Figure 38A) and the benefits of the EAS method smoothing that have been shown in verification from the Environmental Modeling Center (EMC) are lost.

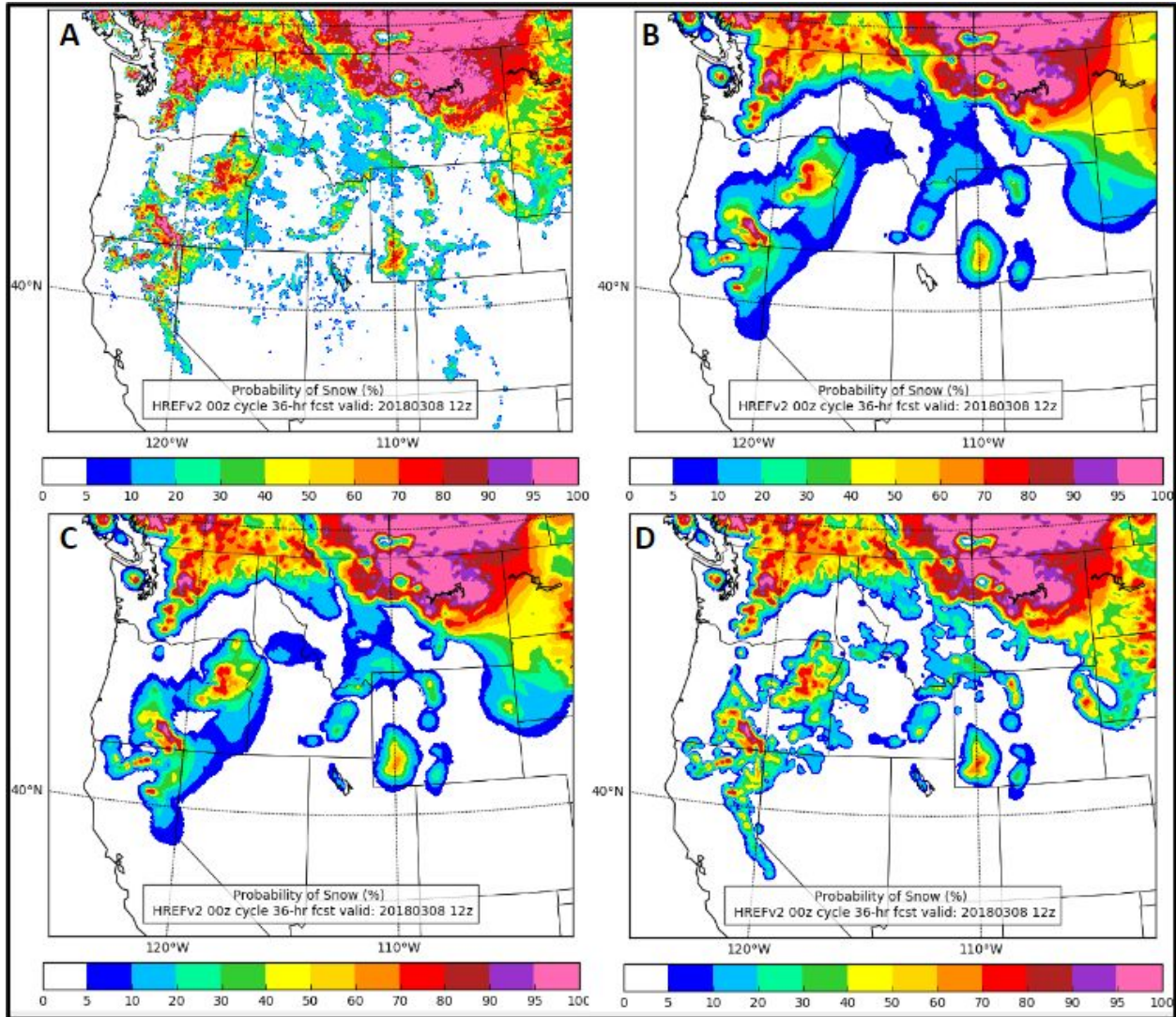


Figure 38. Probability of snow valid at 12Z 3/8/2018 displayed using (A) point probabilities, (B) EAS method with $\alpha = 0.50$, (C) EAS method with $\alpha = 0.75$, and (D) EAS method with $\alpha = 0.90$. Images courtesy of Ben Blake/EMC.

Recommendation

Because of the overall differences in forecaster preference regarding the two methods, it is recommended that both point and EAS probabilities be made available to forecasters for application in multiple scenarios. It is also recommended that the alpha value remain at 0.5 unless further verification from EMC demonstrates otherwise.

WWE Precipitation Type Forecasts

Each forecast day began with a standard weather briefing that included examining current satellite, radar, and model mass fields in order to assess the environment. Participants were then led through the available experimental precipitation type guidance, and WPC's 5km QPF in 6-hour intervals to assess areas to insert a forecast precipitation type. WPC's Watch Collaborator (a tool available to field forecasters displaying the probabilities of WPC PWWF exceeding WFO warning criteria for snow and ice) was also consulted in the forecast process.



Figure 39. *Participants discuss the experimental guidance while making precipitation type forecasts.*

Four forecasts (three for Day 1), each in 6-hour increments, were created over either the Day 1, Day 2 or Day 3 period (example with verification shown in Figure 40). The forecast period was chosen by the participants and WPC-HMT based on which day presented the highest likelihood to use the experimental precipitation type guidance. Day 1 forecasts were valid from 18Z-12Z, rather than the full 12Z-12Z for Days 2 and 3. The HMT facilitators, with the input and assistance of the participants, drew contoured areas for snow, sleet, and freezing rain, at times using WPC QPF as areal guidance for each 6-hour period over the entire CONUS. During the first half of the experiment, the WPC SLR was then applied for forecasts areas of snow, a 2:1 ratio for areas of sleet, and the FRAM algorithm for freezing rain to create a snow and ice accumulation forecast. As the experiment matured, the snow and ice accumulation forecast was dropped to focus more on the forecasting and verification of precipitation type.

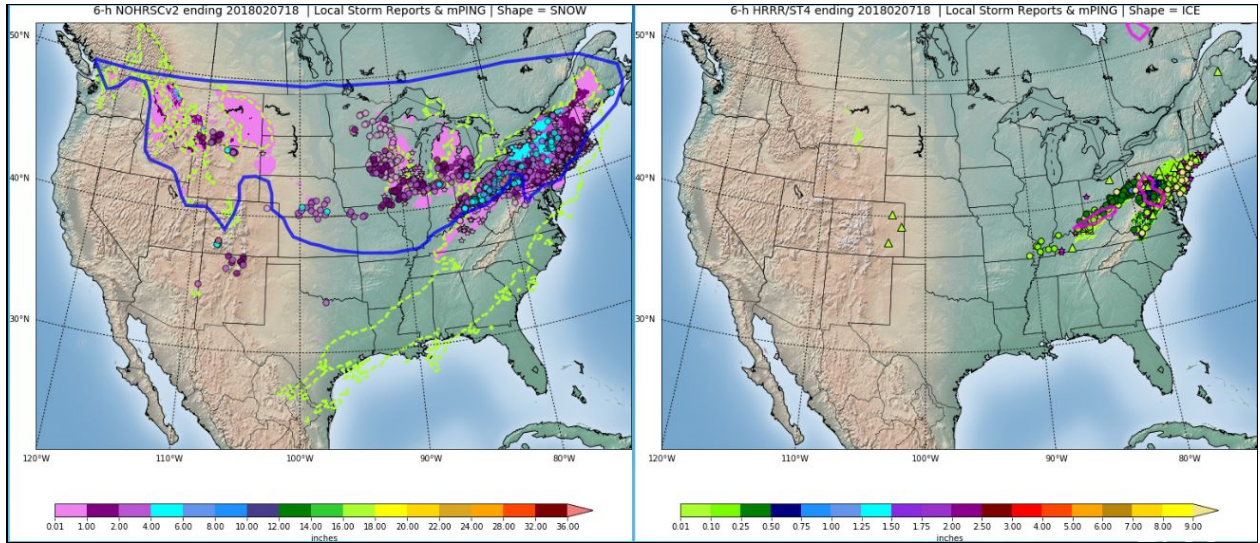


Figure 40. An example of the contoured snow (left, blue contour) and ice (right, pink contours) experimental precipitation type forecasts created by the participants valid 12Z 02/07/2018 to 18Z 02/08/2018, with verification underlaid. Please refer to the verification section in this document for more details.

Figure 41 shows that on a scale of 1 to 10, the participants scored the 6-hour experimental precipitation forecasts 6.6 (12-18Z) , 6.4 (18-00Z), 6.9 (00-06Z), and 7.4 (06-12Z). As shown in the box plot in Figure 42, the lowest score given was a 3 and the highest score a 10. Most scores fell between 5 and 8, with the concentration of highest scores given in the last period (06-12Z).

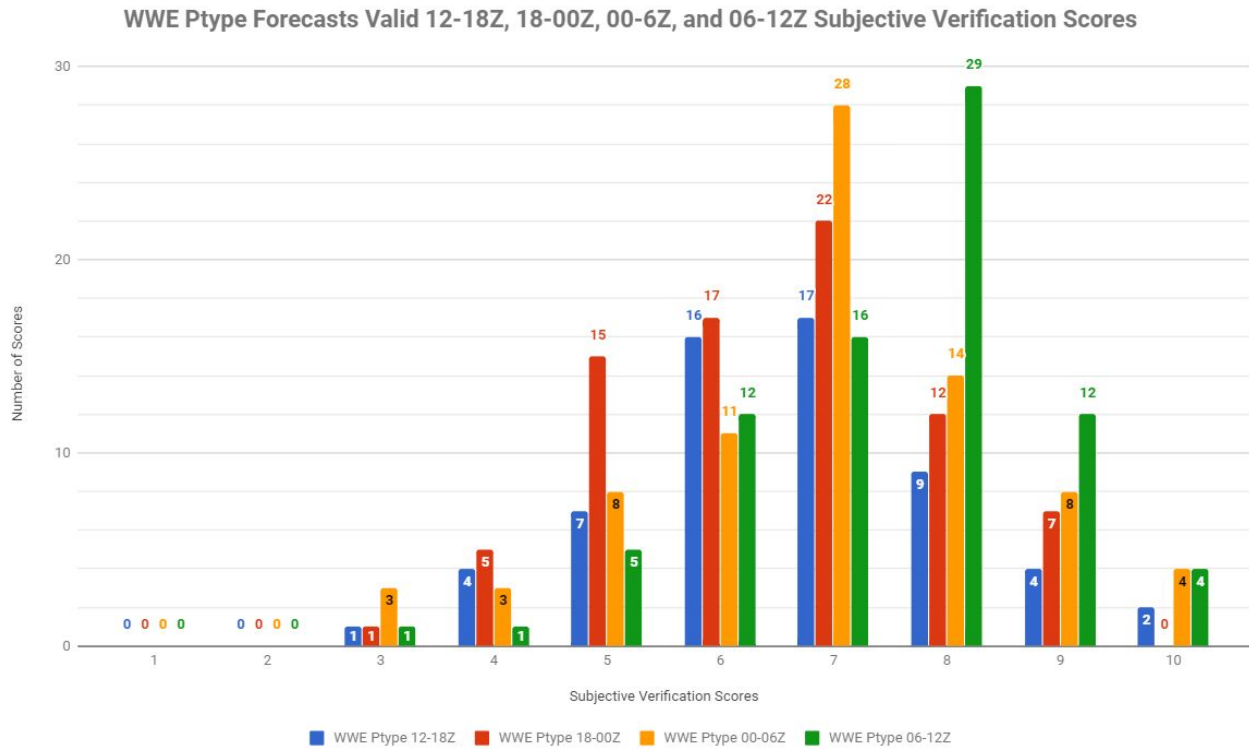


Figure 41. A graphical representation of the participant ratings of each 6-hour precipitation type forecast that comprised the 24-hour experimental precipitation type forecast. Participants were asked to rate each 6-hour forecast on a scale from 1 (poor) to 10 (very good).

Subjective Verification Scores of the WWE Precipitation Type Forecast

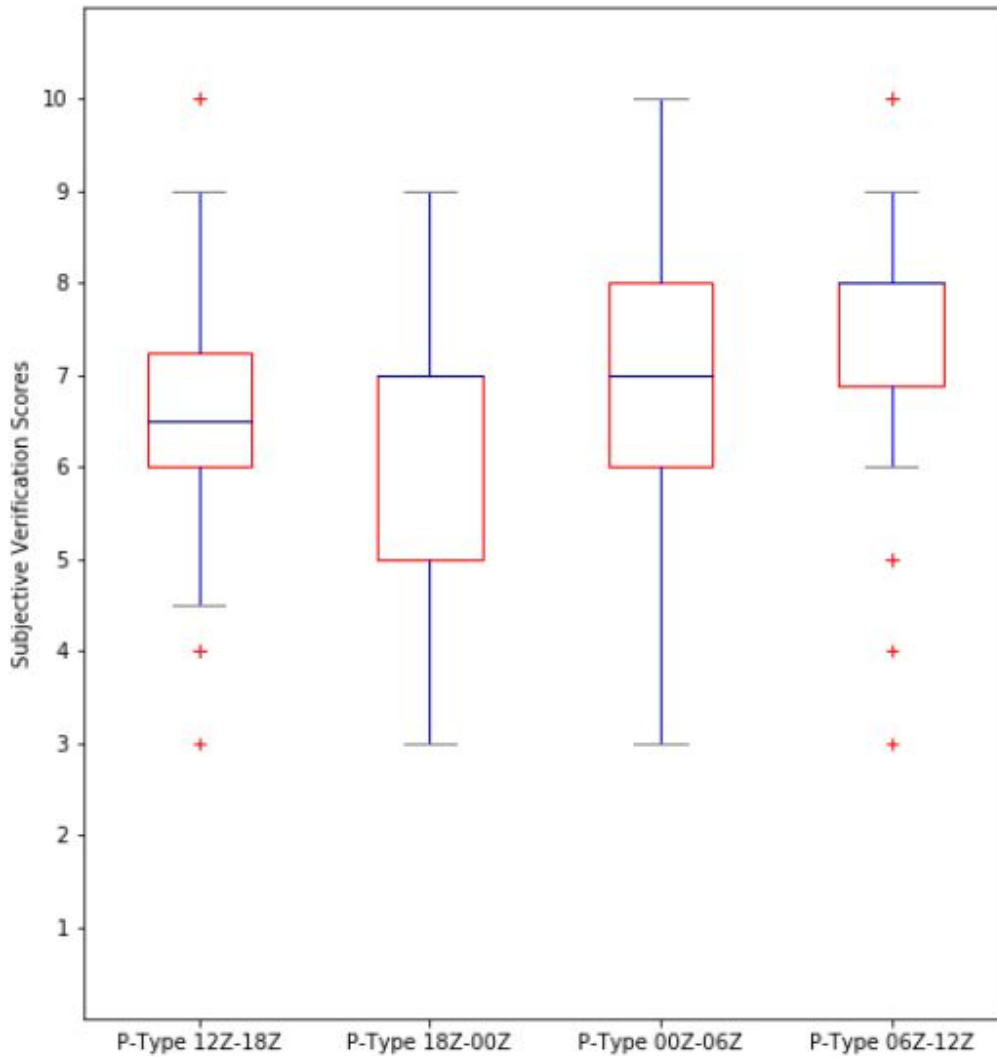


Figure 42. Box plots showing the distribution of the participant ratings of each 6-hour precipitation type forecast that comprised the 24-hour experimental precipitation type forecast. Participants were asked to rate each 6-hour forecast on a scale from 1 (poor) to 10 (very good).

Feedback

Participants often felt the snow contours were too broad-brushed, and the ice contours were too small. There was significantly higher confidence in the snow forecasts as opposed to those for freezing rain and/or sleet. This was reflected in the scores and comments. Errors in and along transition zones and the gradient along the areal extent of the winter precipitation were

theorized to be a result of model errors forecasting parameters such as 2-meter temperature, wet bulb temperature, and snow levels responsible for resolving complex terrain.

The Pacific Northwest coastline was particularly difficult when predicting snow versus rain. This region was often predicted to receive rain when snow was observed. Regions of the northeast along the New England coastline (Figure 43) and the region wedged between the Appalachian Mountains and the coast where cold air can become trapped presented challenges not only for the presence of sleet and freezing rain, but transitioning between multiple precipitation types over the 6-hour period as seen in the hourly verification tools and station observations. Areas around the Great Lakes were a challenge for freezing rain and sleet prediction as well.

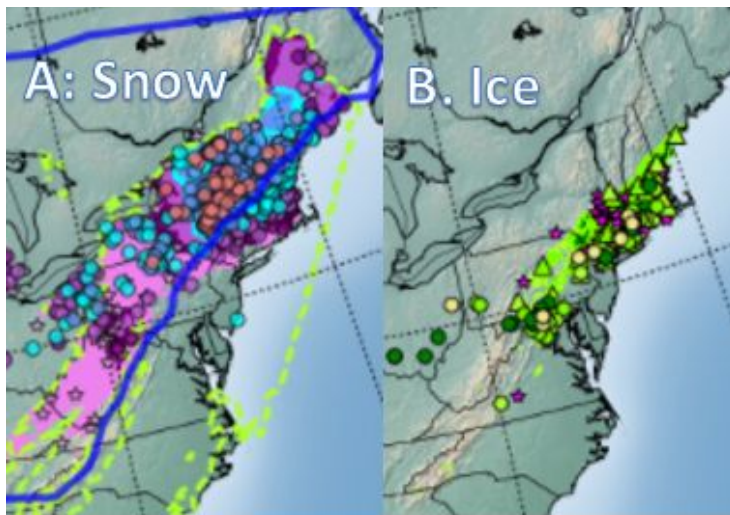


Figure 43. Example of a 6-hour experimental forecast, valid from 18Z 02/07/2018 to 00Z 02/08/2018, exhibiting difficulty capturing transition zone along the New England coast. Many ice reports are found when no contour for ice was drawn (B). The snow contour did not stretch far enough east to capture all of the observed snowfall (A). Ice guidance did not indicate the risk of freezing rain or sleet. In this case, the ability to overlap the contours would have been useful.

It should be noted that limitations in technical capability did not allow for the overlap of precipitation type forecast contours for this year's experiment, which prevented participants from being able to delineate areas in which more than one type of precipitation was predicted within the forecast period. The experiment was designed to predict snow or ice or freezing rain, and not a mix of more than one precipitation type.

Round-Table Discussions

During the two in-house weeks of the Winter Weather Experiment, each day featured a focused topic of discussion among the participants related to improving the prediction of precipitation

type. Topics spanned from NWS WFO tools, techniques, and allotted forecast time to the different predictive methods in the NWP models. The discussions were complemented by presentations from subject matter experts throughout the week.

Most forecasters feel that the predictive starting grids from Forecast Builder in AWIPS (more on Forecast Builder: <https://www.youtube.com/watch?v=cMY21ORJDrg&feature=youtu.be>) were adequate in assisting the forecasters with the first 6 to 12 hours of the precipitation forecast despite necessary adjustments to these grids. Due to the short time forecasters have to produce a forecast and the emphasis on decision support, editing these grids is easier than assessing the high-resolution models individually. WFO adjustments to the probabilistic ice grid, applying the Bourgoiu Energy Areas method (Bourgoiu, 2000), and applying snow level grids enhance the existing Top-Down method used in Forecast Builder for the determination of precipitation type.

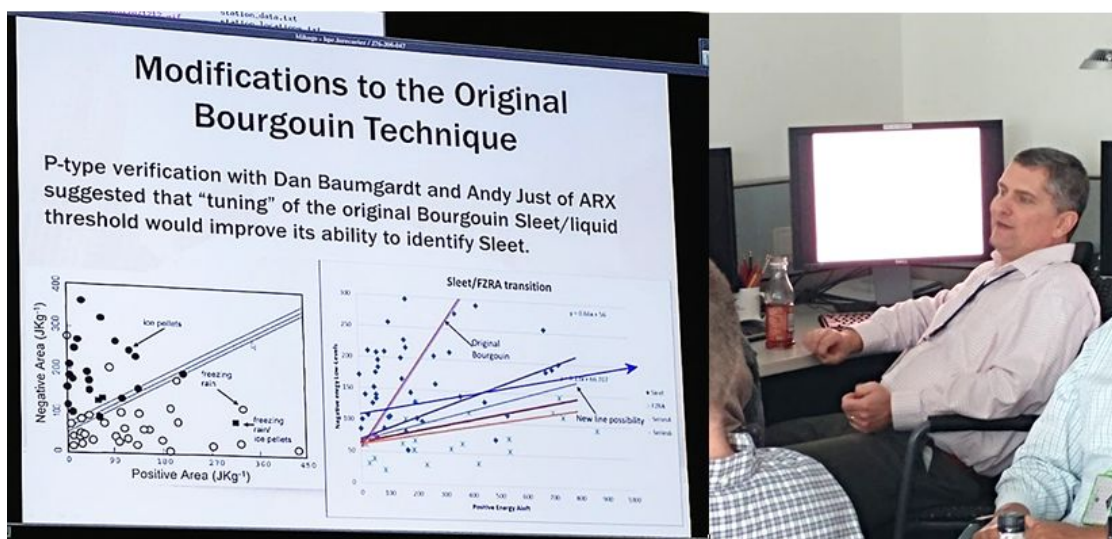


Figure 44. Eric Lenning (NWSFO Chicago) discusses the modified Bourgoiu Technique for determining precipitation type with the WWE participants. This technique calculates total layer energies to better account for melting and refreezing of hydrometeors.

The majority of participants see the direct explicit precipitation determination from NWP models as the easiest to interpret and to edit. There is also agreement that microphysics schemes are now sophisticated enough to do an adequate job of predicting precipitation type.

At the end of the week, the participants were asked to anonymously submit survey responses to expand on their assessment and vision for advancing the science and forecasting of precipitation type. When asked of their opinion on the biggest precipitation type determination challenges and where the industry should place its largest effort, the consensus was to devote resources to the improvement of the NWP models. Any improvements to the models that contribute to blends and ensembles will improve the prediction overall. More

testing is recommended to resolve localized details such as terrain and transition zones. High resolution models continue to struggle with shallow arctic air near the surface which impacts precipitation type (freezing drizzle, light snow). Additional NWP improvements are needed to account for wet bulb and diabatic cooling that results from melting.

A common call from WWE participants for advancing the prediction and forecasting of precipitation type was increased availability of probabilistic ensembles that provide uncertainty information, and improved observation and verification resources for evidence-based decision making both by the forecaster and for advancing NWP methodologies.

Finally, there was extensive discussion on winter weather verification during the residence sessions. Verification of snowfall and snow precipitation type has always presented a challenge. The NOHRSCv2 snowfall analysis showed great improvement over the 24-hour interval. Moreover, the first-time effort to visually verify sleet and freezing rain, both deterministic and probabilistic, presented even more of a challenge to participants. As stated in the verification section earlier in this report, WPC provided multiple data sets to verify forecasts. However, each presented their individual shortcomings in the verification process. There was a general consensus that a concentrated effort to build a precipitation type verification methodology is needed.

Acknowledgements

The WPC-HMT would like to sincerely thank WPC's Kathy Gilbert, Jim Nelson and Mark Klein for their commitment and diligence to enabling this newly-formatted experiment to be successfully executed. We would also like to recognize Geoff Manikin for fostering deep collaboration with EMC, not only for data and subject expertise but with precise scheduling of EMC participants.

This experiment would also not have been possible without the commitment of our dedicated NWS SOO team: David Beachler (WFO Marquette, MI), Steven Zubrick (WFO Sterling, VA), Gene Petescu (WFO Anchorage, AK), Kirby Cook (WFO Seattle, WA), Christopher Buonanno (WFO Little Rock, AK), Eric Lenning (WFO Chicago, IL), and Michael Dutter (WFO Wakefield, VA).

References

- Bourgouin, P., 2000: A Method to Determining Precipitation Types. *Wea. Forecasting*, **15**, 583–592.
- Hamrick, D., 2011: Mid-Atlantic and Northeast U.S. Winter Storm January 26-27, 2011. Available at: http://www.wpc.ncep.noaa.gov/winter_storm_summaries/event_reviews/2011/Mid-Atlantic_Northeast_WinterStorm_Jan2011.pdf.
- HMT-WPC, 2017: The 2017 Flash Flood and Intense Rainfall Experiment Final Report. Available at: http://www.wpc.ncep.noaa.gov/hmt/2017_FFaIR_final_report.pdf.
- , 2015: The 2015 HMT-WPC Winter Weather Experiment Final Report. Available at: https://hmt.noaa.gov/experiments/pdf/WWE2015_final_report.pdf.
- Just, A., 2017: Personal collection and analysis of precipitation type sounding data. *NWS Central Region Headquarters*.
- Manikin, G. S., 2005: An overview of precipitation type forecasting using NAM and SREF data. *21st Conference on Weather Analysis and Forecasting/17th Conference on Numerical Weather Prediction, Washington, D.C.*, 8A.6.
- Novak, D. et al, 2017: Northeast Winter Storm 13-15 March, 2017. Available at: http://www.wpc.ncep.noaa.gov/storm_summaries/event_reviews/2017/WPC_EventReview_March2017.pdf.
- Roebber, P. J., S. L. Bruening, D. M. Schultz, and J. V. Cortinas, 2003: Improving snowfall forecasts by diagnosing snow density. *Wea. Forecasting*, **18**, 264-287.
- , Melissa R. Butt, Sarah J. Reinke, and Thomas J. Grafenauer, 2007: Real-Time Forecasting of Snowfall Using a Neural Network. *Wea. Forecasting*, **22**, 676–684.
- Zheng, M., E.K.M. Chang, B.A. Colle, Y. Luo, Y. Zhu, 2017: Applying fuzzy clustering to a multi-model ensemble for U.S. east coast winter storms: scenario identification and forecast verification. *Wea Forecasting*, **32**, 881-903, <http://doi.org/10.1175/WAF-D-16-0112.1>.

Appendix A: Featured Guidance and Tools for Experimental Forecasts

In addition to the full multi-center suite of deterministic and ensemble guidance available to WPC forecasters, participants were asked to consider several different precipitation type forecasting data sets and techniques while preparing their Day 2 and Day 3 precipitation type forecasts. Participants also had access to the experimental short range HRRR (HRRRX) and the HRRR Ensemble (HRRRE) system from the Earth System Research Laboratory (ESRL) as well as WPC's operational PWWF ensemble that is used to derive WPC's probabilistic winter precipitation forecasts. Table 3 summarizes the model guidance that was the focus of the short range portion of the experiment, and more information about each dataset is provided below.

Table 3. Featured Day 1-2 guidance for the 2017-2018 HMT-WPC Winter Weather Experiment. Experimental guidance is shaded.

Provider	Model	Resolution	Forecast Hours	Notes
EMC	NAM/Nest	12/3 km	60	12 km parent NAM and its 3 km CONUS and 3 km nest <i>Deterministic Ptype Guidance using WPC algorithm</i>
WPC	PWWF ensemble (68 members)	20 km	72	Operational PWWF ensemble includes the WPC deterministic forecast, SREF members, GFS mean, ECMWF mean, and deterministic NAM, GFS, CMC, and ECMWF; SLR is an average of multiple techniques <i>Weighted Probabilistic Ptype Guidance</i>
ESRL/GSD	HRRRX	3 km	18 (all hours) 36 (00/06/12/18Z) 48	Deterministic Experimental HRRR (HRRRv3 and v4 flux over WWE) <i>Deterministic Ptype Guidance and Snowfall</i>
ESRL/GSD	HRRR Ensemble (HRRRE)	3 km	18 hours at 12Z, 15Z, 18Z, 21Z 36 hours at 00Z 48 hours at 09Z	9 HRRR members, Sub-CONUS domain, stochastic <i>Probabilistic Ptype Guidance</i>
EMC	HREFv2	3 km	36	Hydrostatic Multiscale Ensemble with 8 members including NAM Nest, NMB, and ARW which produces probabilistic winter output <i>Probabilistic Ptype Guidance Day 1</i>
EMC	FV3-GFS	13 km	10 days	3D hydrostatic dynamical core; vertically Lagrangian; GFS analyses initialization/physics

				<i>Deterministic Ptype Guidance, Model implicit SLR</i>
MDL	NBMv3	2.5 km	Hourly out 36 hrs 3-hrly to Day 8 6-hrly Days 8-10	Runs every hour with 15 different deterministic and ensemble systems Weighted Conditional <i>Probabilistic Ptype Guidance</i>

Operational Guidance

3 km NAM CONUS NEST and NAM12

The 3 km NAM CONUS NEST and NAM 12KM feature Ferrier-Aligo microphysics. The fraction of frozen precipitation and rime factor parameters of the microphysics scheme have been used at WPC to modify various SLR forecasts, and make appropriate snowfall forecast adjustments in precipitation type transitions area for the past 5 years.

In WWE 2018, elements of the Ferrier microphysics were used to test a precipitation type algorithm.

Snow	Fraction of Frozen Precipitation > 90%
Sleet	Fraction of Frozen Precipitation 50-89% Rime Factor > 20
Freezing Rain	Fraction of Frozen Precipitation 49% or < 2-meter Temperatures < 32F

WPC PWWF Ensemble

The WPC PWWF Ensemble is generated internally by WPC and is used extensively in the WPC Winter Weather Desk forecast process. The ensemble membership is as follows:

- 6 SREF ARW (09z)
- 9 SREF NMMB (09z)
- 10 ECMWF Ensemble (00z)
- 10 GFS Ensemble Members (06z)
- 4 ARW Hires Window (00z/12z) Day 1 or GFS (06z)/SREF NMMB Day 2,3
- 2 NMB Hires Window (00z/12z) Day 1 or GFS (00z)/SREF NMMB Day 2,3
- 1 GFS 12z

- 1 ECMWF Deterministic (00z)
- 1 NAM Nest (Days 1-2) or NAM 12 km (Day 3) (12z)
- 1 GEFS Ensemble Mean (06z)
- 45 Total Members

The WPC 24-hour deterministic snow or freezing rain accumulation forecast is also an ensemble member (making it a total of 46) serving as the mode or “most likely” solution, while the multi-model ensemble provides the variance of the distribution. The Winter Weather Desk (WWD) at WPC uses an ‘average’ SLR for generation of its operational products. This scheme equally weights different SLR schemes to calculate the final SLR used operationally:

- 1) Roebber SLR modified with rime factor and fraction of frozen precipitation data from 12 km NAM (post-processed and mapped to 40 km)
- 2) Roebber Method applied to the 70 km GFS (post-processed, mapped to 40 km)
- 3) Climatological mean SLR (106 km) for the corresponding season (Baxter, 2005)*
- 4) Climatological mean SLR modified with rime factor and fraction of frozen precipitation derived from the NAM
- 5) Fixed 11:1 snow ratio

*WPC replaces any grid points that do not have a climatological SLR value (e.g. locations in the deep south) with a default 10:1.

Experimental Guidance

EMC/ESRL Experimental Deterministic High-Resolution Rapid Refresh (HRRRX) and Ensemble HRRRE

The Experimental HRRR (HRRRX) in the WWE Experiment is a WRF-ARW-based 3 km model initialized with the latest 3-D radar reflectivity using a digital filter initialization (radar-DFI) technique (via the parent 13 km RAP) and is updated hourly. HRRRX contains numerous model changes including an update to WRF-ARW version 3.9.

The HRRRE is an experimental convective-allowing ensemble analysis and forecasting system run at NOAA/ESRL/GSD. It is being developed and tested for three main reasons: (1) improving 0-12 h high-resolution forecasts through ensemble-based, multi-scale data assimilation, (2) testing ensemble-design concepts for 0-48 h forecasts produced with a single model, and (3) providing a foundation for experimental, on-demand, very-high-resolution applications such as Warn-on-Forecast.

Model

- WRF-ARW version 3.8+, combining elements of versions 3.8 and 3.9 plus other GSD-specific features
- Configuration identical to experimental HRRR (<https://rapidrefresh.noaa.gov/hrrr/>), except that domain covers central and eastern US only (55% of HRRR domain), and a standard vertical coordinate is used instead of a hybrid coordinate

Data-Assimilation Ensemble

- 36 members
 - 3-km horizontal grid spacing
 - Initial ensemble-mean atmospheric state from RAP analysis
 - Atmospheric spatial perturbations from members 1-36 of GDAS ensemble
 - Initial ensemble mean of land-surface state from HRRR
 - Random soil-moisture perturbations added to each member at initial time
 - Random perturbations to MU, U, V, T, and QVAPOR added to boundary conditions of each member
- Initialization at 0300 and 1500 UTC, followed by hourly cycling for 9 h to 1200 and 0000 UTC, respectively

Hourly Data Assimilation

- Observations
 - NCEP bufr conventional observations, as in HRRR (http://www.emc.ncep.noaa.gov/mmb/data_processing/data_processing/)
 - MRMS gridded radar reflectivity observations, thinned in horizontal and vertical directions
- Gridpoint Statistical Interpolation (GSI) for observation preprocessing and calculation of ensemble priors
- DART ensemble adjustment Kalman filter (EAKF) for assimilation
 - Analysis variables: U, V, T, QVAPOR, PH, MU, QCLOUD, QRAIN, QICE, QSNOW, QGRAUP
 - Gaspari-Cohn compact pseudo-gaussian for localization
 - Horizontal localization radius (full radius, where weight reaches zero) 300 km and 18 km for conventional and radar observations, respectively
 - Vertical localization radius (full radius, where weight reaches zero) 8 km and 6 km for conventional and radar observations, respectively
 - GSI adjustments applied to each member individually after EAKF
 - Soil adjustment, as in HRRR
 - Cloud clearing based on satellite observations, as in HRRR
- Relaxation to prior spread (inflation factor 1.2) after assimilation each hour

Ensemble Forecasts

- 9-member, 48-h forecast initialized from first 9 members of data-assimilation ensemble at 1200 and 0000 UTC
- 3-km horizontal grid spacing
- Random perturbations to MU, U, V, T, and QVAPOR added to boundary conditions of each member
- Stochastic physics

- Post processing: An ensemble post-processing system is applied to the nine HRRRE forecast members to produce all-season weather hazard probabilities including heavy rainfall as is done with the time-lagged HRRR. For the 2017-18 Winter Weather Experiment, HRRR-E probabilities are the fraction of members that exceed a given threshold, or predict a given precipitation type, at a point. The final probability field ($100 \cdot (n/\text{total})$) is smoothed using a Gaussian filter of width 25 km.

HRRRX/HRRR-E Precipitation Type Algorithm

HRRRX and HRRR-E both use a microphysics-based algorithm to predict instantaneous precipitation type (snow, sleet, rain and/or freezing rain). Output is in the form of categorical (0 or 1) grids for each precipitation type, stored in GRIB-2 fields CSNOW, CICEP, CRAIN, and CFRZR. Multiple precipitation types (up to 3) may be forecast simultaneously, with the only exception being that the algorithm will predict either rain OR freezing rain. The algorithm uses 2-m temperature, 3-D hydrometeor mixing ratios and fall rates to provide a first guess of precipitation type reaching the ground. A flow-chart of the algorithm is shown below in Figure 45, with further details available in the following Weather and Forecasting article:

Benjamin, S., J. M. Brown, and T. G. Smirnova, 2016b: Explicit precipitation-type diagnosis from a model using a mixed-phase bulk cloud-precipitation microphysics parameterizations. *Wea. Forecasting*, 31, 609–619.

<http://journals.ametsoc.org/doi/abs/10.1175/WAF-D-15-0136.1>

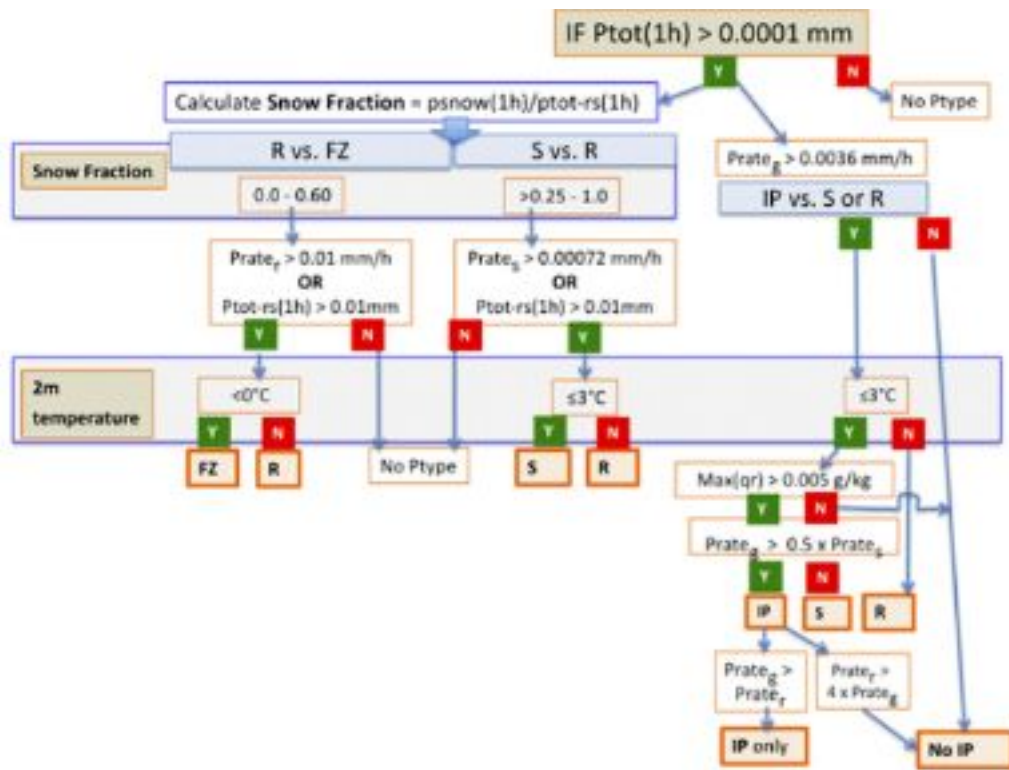


Figure 45. Flowchart showing the HRRR/HRRR-E microphysics-based algorithm to predict instantaneous precipitation type.

HRRRX/HRRR-E Variable-Density Snowfall Algorithm

HRRR and HRRR-E both use a variable-density snowfall accumulation algorithm. The output of this algorithm, contained in the “ASNOW” GRIB-2 field, is the total depth of new snowfall and graupel/sleet accumulation. Note: in the Thompson microphysics scheme and this algorithm, falling sleet is considered graupel.

At every model time step, the algorithm first determines snow-water-equivalent and graupel-water-equivalent precipitation rates using the mixing ratios and fall rates of these hydrometeors at the lowest model level. Next, some melted water equivalent is subtracted from these totals if the land surface temperature is above 0C. Snow and graupel density functions are then applied to determine the rate of increase of total snow+graupel depth. The product of this accumulation rate (m/s) and the model time step (20 s) yields a snowfall+graupel depth accumulation that is added to a running total depth. At every hour, HRRR and HRRR-E output this running total as ASNOW.

Snow-to-liquid and graupel-to-liquid ratios are linear functions of the temperature at the lowest model level (typically ~8 m), given by:

$snow\ density\ (kg/m^3) = \min(250, 1000 / \max(4.179, (13. * \tanh((274.15 - (x + 273.15)) / 3))))$
 $graupel\ density\ (kg/m^3) = \min(50., 1000 / \max(2, (3.5 * \tanh((274.15 - (x + 273.15)) / 3))))$
 The minimum and maximum densities possible from the snow equation are 76 and 250 kg/m³, respectively, equal to snow-to-liquid ratios of 13:1 and 4:1.

GFDL/EMC FV3 GFS

The GFDL Finite Volume Cubed-Sphere Dynamical Core (FV3) is a scalable and flexible dynamical core capable of both hydrostatic and non-hydrostatic atmospheric simulations. The full 3D hydrostatic dynamical core, the FV core, was constructed based on the Lin-Rood (1996) transport algorithm and the Lin-Rood shallow-water algorithm (1997). The pressure gradient force is evaluated by the Lin (1997) finite-volume integration method, derived from Green’s integral theorem based directly on first principles, and demonstrated errors an order of magnitude smaller than other well-known pressure-gradient schemes. Finally, the vertical discretization is the “vertically Lagrangian” scheme described by Lin (2004).

The most unique aspect of the FV3 is its Lagrangian vertical coordinate, which is computationally efficient as well as more accurate given the same vertical resolution. Recently, a more computationally efficient non-hydrostatic solver was implemented using a traditional semi-implicit approach for treating the vertically propagating sound waves. This faster solver is the default. The Riemann solver option is more efficient for resolution finer than 1km, and also more accurate, because sound waves are treated nearly exactly. FV3 core has been chosen for the Next Generation Global Prediction System project (NGGPS), designed to upgrade the current operational Global Forecast System to run as a unified, fully-coupled system in NOAA’s Environmental Modeling System infrastructure. For more information, please see <https://www.gfdl.noaa.gov/fv3/>.

National Blend of Modes Version 3 (NBMv3):

The NBMv3 runs every hour with 15 different deterministic and ensemble systems. For the CONUS, typically 4 to 6 new model runs are available each hour, with up to 7 or 8 (~50% new) on four cycles. The NBMv3 uses the TOD (Time of Day) concept, rather than the “model” cycle. Therefore, a 12z run of NBMv3 does not contain a single 12z model run. It is a data cutoff time (newest models are from 10-11z in this example; several 00z to 06z models are included). NBMv3 is run each hour on top of hour, and available 50-60 minutes later. Models included are found in Table 4.

Table 4. Data Dependencies for the NBMv3.

<u>Global Models</u>	<u>Mesoscale Models</u>
-----------------------------	--------------------------------

GFS - 0.25 degree	HRRR - 3 km
GEFS mean + members - 0.5 degree	NAM Low-Res - 12 km
CMC deterministic (PoP12/QPF only) - 1 degree	NAM High-Res - 3 km (pending)
CMC ensemble mean + members - 1 degree	HIRESW (NMMB and ARW cores) - 3 km
FNMOG - 1 degree	RAP - 12 km
GMOS - 2.5 km	SREF - 40 km
EKDMOS - 2.5 km	GLMP - 2.5 km
CCPA (used in NBM Precip SQM) - ~ 13 km	URMA (CONUS, AK, HI, PR) - 2.5 km
NBM 2.5 Precip Stochastic Quantile Mapping	

For precipitation type the NBM uses the Top-Down method (Figure 45). The NBM uses a number of forecast grids in creating probability of precipitation type, and snow and ice accumulation grids. These grids include:

- Temperature
- Dewpoint
- Wind
- QPF
- SnowRatio
- SnowLevel
- MaxTwAloft (maximum wet-bulb temperature in the troposphere)
- ProbIcePresent (Probability of Ice Present in the Clouds)
- ProbRefreezeSleet (Probability precipitation can re-freeze back into sleet)
- RoadTemp (temperature of the roads)

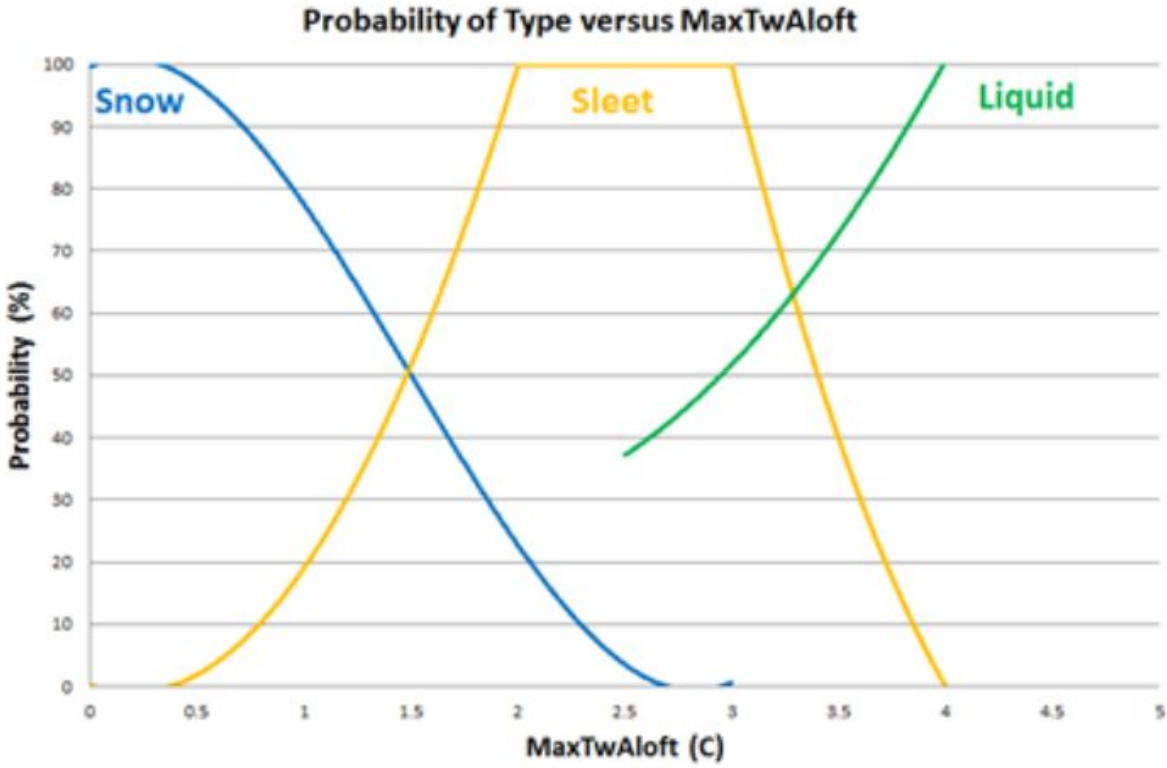


Figure 45. NBM Calculation of Probability of Precipitation Types from MaxTwAloft Based off the top-down approach

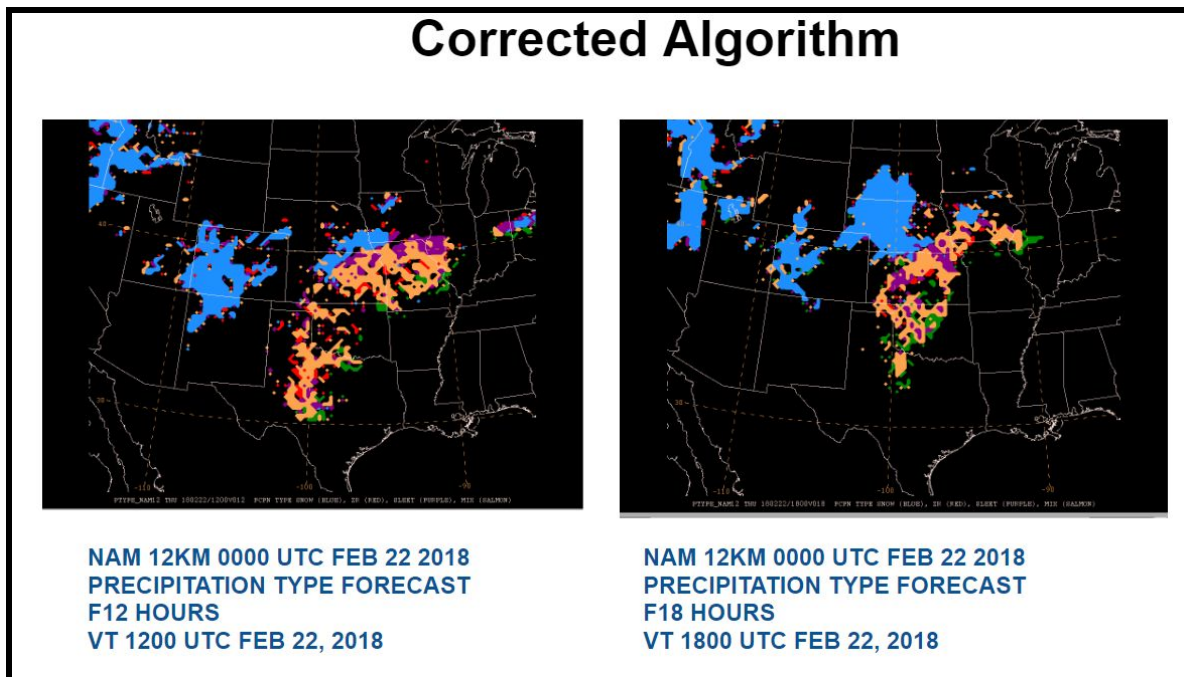
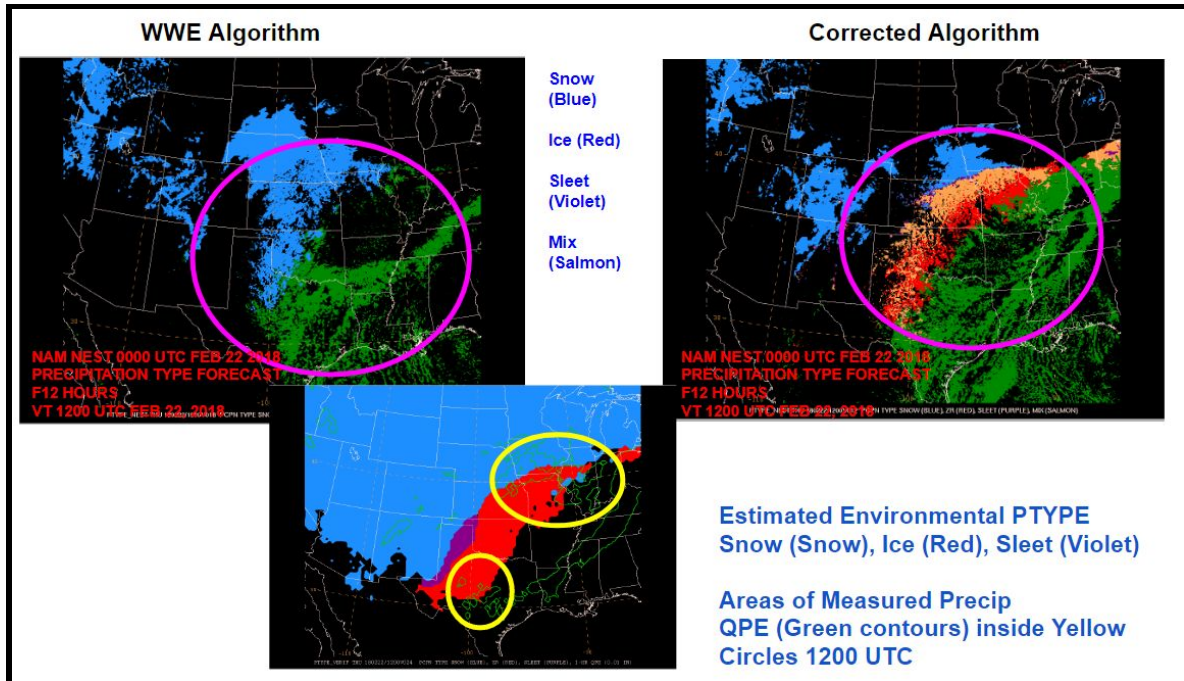
APPENDIX B

WPC MODE Settings for Objective Verification

- 24 and 48 HR model snowfall accumulation forecasts verified against 24 HR NOHRSCv2 snowfall accumulations
 - 12Z to 12Z forecast cycle used
 - Both snowfall accumulation forecasts and NOHRSCv2 snowfall accumulations re-gridded to a common 5km lat/lon grid
 - Thresholds investigated varied.
- MODE
 - Grid stats harvested from MODE CTS
 - Circular convolution radius of 3 grid squares used
 - Double thresholding technique applied

APPENDIX C: Corrected WPC Precipitation Type Algorithm for the NAM Nest

A retrospective example of the corrected WPC precipitation type algorithm applied to the 00Z run of the NAM Nest on Feb 22, 2018. This event was forecast in the last residence week of WWE, which distinctly uncovered the errors in the WPC methodology over Missouri, Kansas, and Oklahoma.



UPDATED PROCESSING:

If Fraction of Frozen Precipitation > 90, P-type is Snow

If Fraction of Frozen Precipitation <= 89, and >= 70, P-type is Sleet

If Fraction of Frozen Precipitation <= 69, and >= 21, P-type is mixed

If Fraction of Frozen Precipitation <= 20,

P-type is Rain if 2-meter temperature > 32F

P-type is Freezing Rain if 2-meter temperature <= 32F

A final check step was done to highlight sleet:

If Rime Factor is > 10, and Fraction of Frozen Precipitation is >= 99 and >= 50 then P-type is Sleet thereby overwriting any previously assigned P-type