Evaluation of Probabilistic Forecasts for Denver International Airport Snow Operations

Dana Mueller, Rebecca Morss, Matthew Wandishin

NOAA/ESRL/GSL/Forecast Impact and Quality Assessment Services Branch 13 October 2020



Agenda

- Background & Objective
- About the Forecasts
- Methods
- Observed Events
- Results for All Airports
- Case Studies
- Summary

Background & Objective

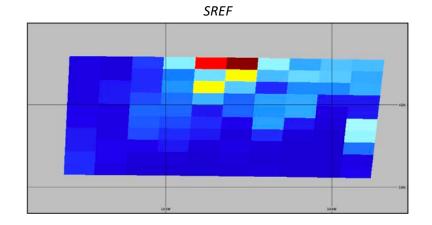
- Collaboration between NCAR, Denver International Airport (DIA or KDEN), and NOAA's Global Systems Division and National Weather Service
 - NCAR observed and conducted interviews to determine DIA airside personnel's decision-making processes ahead of and during winter weather events
 - Event management, alert declarations, staffing, etc.
 - Goal of understanding how DIA currently uses forecast uncertainty information and how it could be used in the future

Objective of this work within the broader scope:

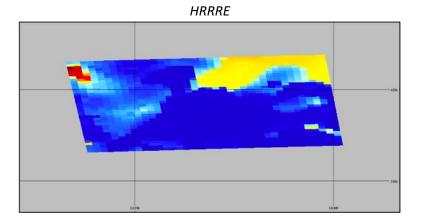
- Evaluate ensemble forecasts with respect to snowfall, wind, temperature, and visibility for airports around Mountain West over the 2018-19 winter season
 - Short-Range Ensemble Forecast (SREF) and High-Resolution Rapid Refresh Ensemble (HRRRE)
 - Other airports besides DIA included to increase sample size
 - Highlight skill of forecasts with an operational focus
 - Explore what type of verification can be done for probabilistic forecasts
- Evaluate DIA Probabilistic Snow Accumulation (PSA) product from Boulder WFO

SREF and HRRRE

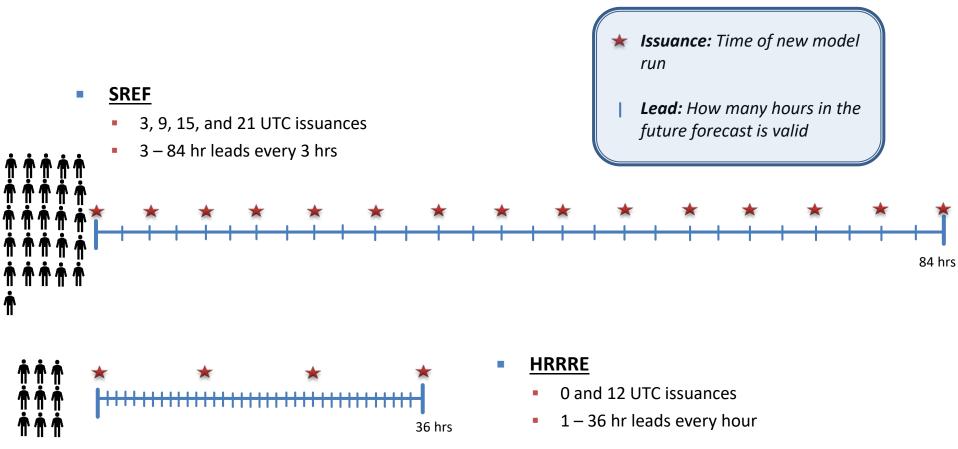
- Short-Range Ensemble Forecast (SREF):
 - Operational ensemble
 - 26 members
 - ~16 km grid cells



- High Resolution Rapid Refresh Ensemble (HRRRE):
 - New, experimental ensemble
 - 9 members
 - ~3 km grid cells



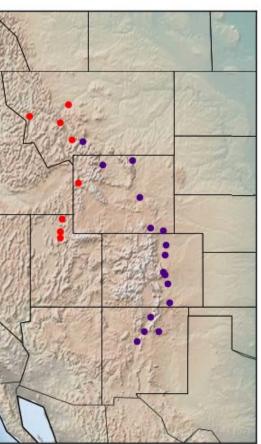
Forecast Timescales



Airports

- SREF evaluated at 24 airports
- Red airports excluded in HRRRE analysis due to domain

Airport	Code	Elevation (m)
Bozeman	KBZN	1361
Helena	KHLN	1182
Great Falls	KGTF	1119
Livingston	KLVM	1418
Missoula	KMSO	975
Jackson	KJAC	1961
Cody	KCOD	1553
Sheridan	KSHR	1202
Casper	KCPR	1621
Cheyenne	KCYS	1868
Laramie	KLAR	2216
Air Force	KAFF	2003
CO Springs	KCOS	1856
Denver	KDEN	1640
Greeley	KGXY	1420
Pueblo	KPUB	1420
Trinidad	KTAD	1756
Albuquerque	KABQ	1618
Las Vegas	KLVS	2091
Santa Fe	KSAF	1930
Taos	KSKX	2161
SLC	KSLC	1286
Logan	KLGU	1355
Ogden/Hill AFB	KHIF	1459



Observed Events Methodology

Event 1: Snow falling; no rain *** METAR = METeorological 38 Aerodrome Reports 36 Event 2: Snow falling Hourly observations of Temp no rain 34 weather variables $T \le 32^{\circ}F$ 32 Snow event definitions: 30 30 Event 3: Snow falling 25 WS no rain $WS \ge 25 \text{ mph}$ 20 15 1.00 0.75 **Event 4: Snow falling** Vis no rain 0.50 Visibility $\leq \frac{1}{2}$ mi 0.25

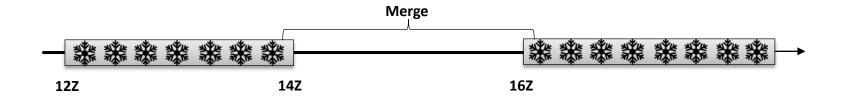
0.00

7

Time \rightarrow

Merging

- Merge events that occur within 12 hours of each other
 - This example considered a single event from operational perspective despite 2 hour lull:



Methodology

Ensembles

Creating Forecast Events

- Made SREF and HRRRE events for each member and each time series (all leads across an issuance)
- Same event definitions (events 1-4; slide 5) as observed events
 - Events merged if close together in same manner as observed events
- Snowfall based on 15:1 snow-liquid ratio
 - Climatological average for Denver (Baxter et al. 2004)

Pairing Forecasts to Observations

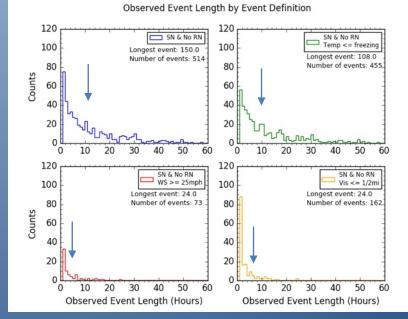
- Considered 24 hours of advanced planning (12 hours for HRRRE due to shorter leads) for airport decision-makers
 - 4 hour latency
 - What is the latest forecast available to users at that time?
- Match closest forecast event to observed event
 - Events not paired if start times are more than 24 hours apart
- Also recorded false alarm events (forecasts predicted snow when none occurred)



Results

Observed METAR Events

- Events often less than 10 hr
 - Some persist much longer due to merging
- Events with low visibility or high wind speeds are shorter



Results - All Airports SREF & HRRRE

Similar to the National Weather Service,

10th, 50th, and 90th percentiles will be used to evaluate the ensembles.

- 90th percentile \rightarrow Low end amount (90% chance of higher value)
- 10th percentile \rightarrow High end amount (10% chance of higher value)

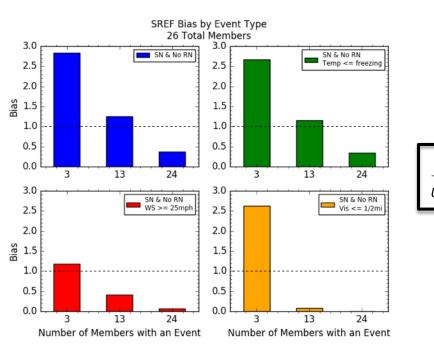
Sections:

- Number of Events
- Amount of Snow
- Timing

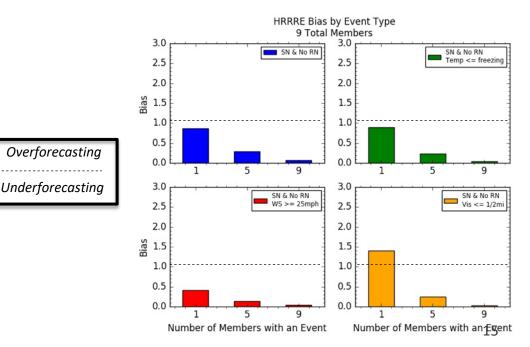
Number of Events

Frequency Bias – Correct # of Events?

- Generally, SREF forecasts slightly too many events while HRRRE forecasts too few
 - Both underforecast wind and visibility events
- Requiring nearly all members to forecast an event will result in missed events



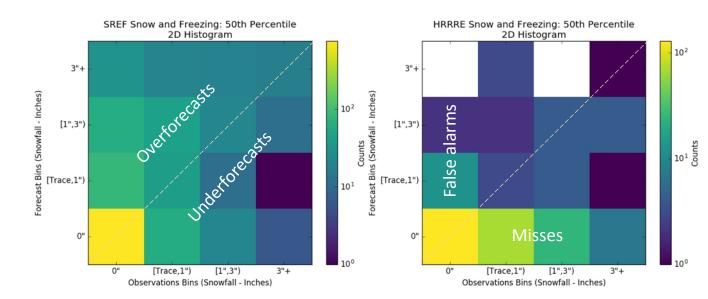
Product	# of Members	% of Total Members	
SREF	3	11.5%	
	13	50%	
	24	92.3%	
HRRRE	1	11.1%	
	5	55.6%	
	9	100%	



Snowfall Amount

Snowfall Amount by Bins

- SREF generally over forecasts
- HRRRE misses many events (bottom row)
- <u>Ideal results:</u> warm, bright colors along diagonal

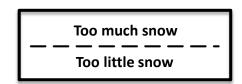


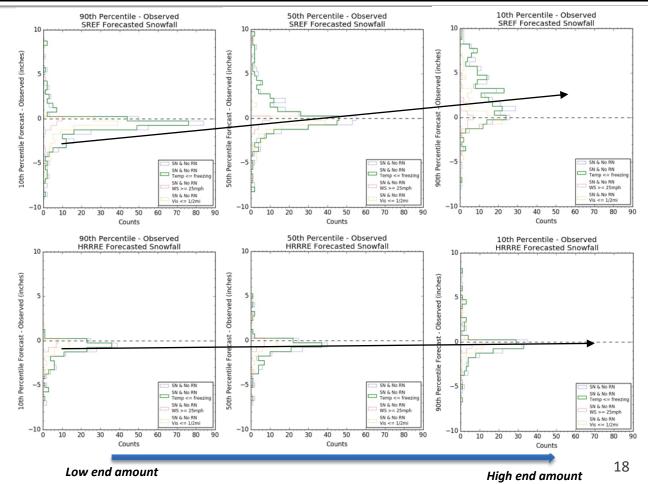
SREF

HRRRE

Snowfall Amount

- SREF has slightly too much, HRRRE too little
- SREF 50th percentile is centered on ~zero error
 - 90th amount is too low and 10th amount is too high
- HRRRE mostly forecasts too little snow and lacks resolution
 - 10th, 50th, and 90th amounts are all about the same
- Determining snow amount:
 - Lower bound: HRRRE 90th percentile
 - <u>Upper bound:</u> SREF 10th percentile
- Only paired events
 - False alarm events had little to no accumulation



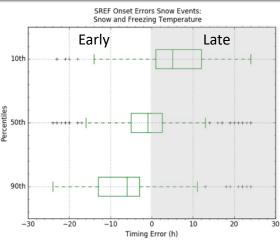




Onset Errors

SREF:

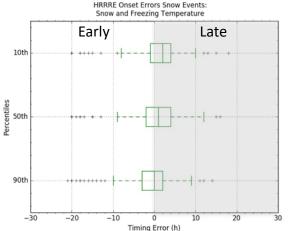
- Small errors for 50th percentile
- Conservative approach is to use 90th percentile start time
 - Events unlikely to start before expected



Small onset errors

HRRRE:

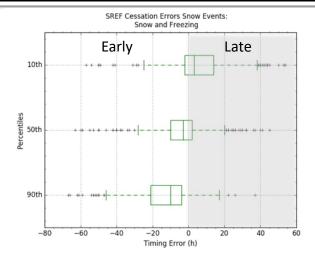
- Like amounts, HRRRE has little spread for onset time
 - Can be caught unaware
- Events mostly start slightly too late



Cessation Errors

SREF:

- Small errors for 50th percentile
- Results more similar across 90/50/10th percentiles compared to onset
- Conservative approach is to use 10th percentile end time
 - Likely won't end event too early



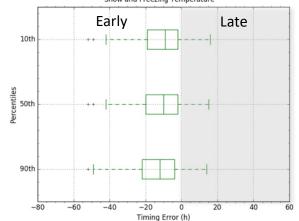
*Note, results are biased toward early ending due to event being forced to end at last lead time

HRRRE Cessation Errors Snow Events: Snow and Freezing Temperature

Larger cessation errors

HRRRE:

- Again, little difference between 90/50/10th percentiles
- Events end too early



Summary

of Events

- SREF has too many events
- HRRRE has too few events
- Both underforecast wind and visibility events
- Neither is well-calibrated for high-end snow events

Snow Amount

- SREF has close to (or slightly over) the observed snow amount, while HRRRE has too little
- HRRRE has poor resolution and misses frequently
- HRRRE has fewer false alarms than SREF, but snowfall is minimal in false alarm events

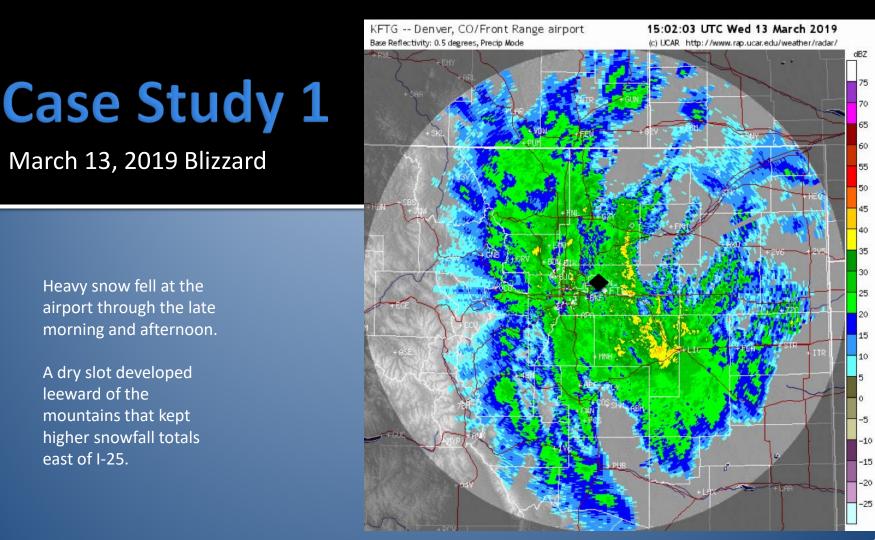
Timing

- Larger cessation errors than onset errors
- SREF 50th percentile timing is accurate
 - Can use other percentiles to avoid surprises
- HRRRE events too short: start too late and end too early
 - Has poor resolution

Heavy snow fell at the airport through the late morning and afternoon.

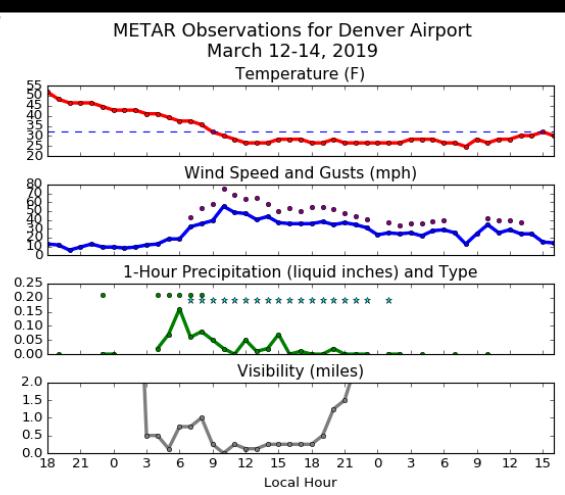
March 13, 2019 Blizzard

A dry slot developed leeward of the mountains that kept higher snowfall totals east of I-25.



Airport Observations

- Rain → snow at 9 am
- High winds
 - Gusts near 80 mph
- Low visibility (< ¼ mile)
 - 9 am to 6 pm

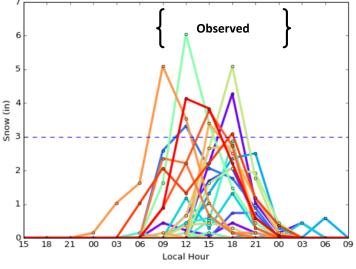


Ensemble Forecasts

SREF forecast from 3/12 at 9 AM:

- Some members indicate heavy snow
- Start times: 6 am to 3 pm
- End times: 3 pm to 9 pm

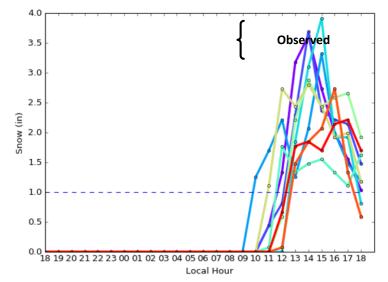




HRRRE forecast from 3/12 at 6 AM:

- All members indicate heavy snow
- Start times: 9 am to noon
- End times: ~6 pm

HRRRE 1-Hour Snow Forecast for Denver Airport March 12-13, 2019



10/14/2020

Forecast Impact and Quality Assessment Section

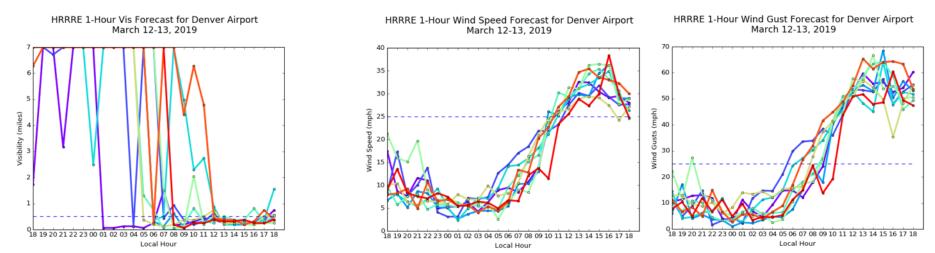
HRRRE Visibility and Wind

<u>Visibility</u>

- All members indicated visibility would drop ≤ ½ mile
- Variation in timing

Wind Speed and Gusts

- Members in good agreement
- Wind speeds: 30 to 40 mph
- Gusts: 50 to 70 mph
- Timing: 11 am to 6 pm



Forecast Impact and Quality Assessment Section

PSA from 9:16 AM on March 12, 2019

Developmental Probabilistic Snow Accumulation Forecast National Weather Service Denver/Boulder, Colorado 0916 MDT Tuesday, March 12, 2019

This developmental product contains the probabilities (%) that the given snow amounts (inches) will occur at Denver International Airport during the respective 6- or 12-hour periods. The probabilities indicate the amount of snow that will fall from the clouds. They do not consider the amount of melting that may or may not occur on various ground surfaces. [T (trace) means more than zero.]

Short Term Discussion (12/12Z-13/12Z): Expect scattered rain showers to develop later this afternoon and continue tonight but no snow is expected through tonight.

Day/Hour GMT	0"	T or more	1" or more	3" or more	6" or more	10" or more
12/12-12/18 12/18-13/00 12/12-13/00	100 100	0 0	0 0	0 0	0	0
13/00-13/06 13/06-13/12 13/00-13/12	100 100	0 0	0 0	0 0	0	0

Long Term Discussion (13/12Z-14/12Z): Rain showers will develop wendesday morning and will become heavier near midday. A push of colder air will then arrive during the late morning hours and will change the rain over to snow between 14-17Z Wednesday. The initial snow will melt on the wet and warm surfaces. Northerly winds will increase drastically, with sustained speeds around 25 to 35 kts, and gusts up near 50 kts during the late morning and afternoon. Snow will become heavy by noon and continue through the afternoon, where snowfall rates will be between 1 to 2 inches per hour, and visibilities will be down near 1/4 to 1/2 mile in blowing snow. A high temperature of 38 degrees will occur in the morning, then fall throughout the day. Snow will taper off Wednesday night and will end between 02-05Z. Expect total snowfall amounts to be between 6-10 inches with large drifts.

Day/Hour GMT	0"	T or more	1" or more	3" or more	6" or more	10" or more
13/12-13/18 13/18-14/00	30 0	70 100	40 100	10 100		
13/12-14/00	1				70	30
14/00-14/06 14/06-14/12	30 100	70 0	40 0	10 0		
14/00-14/12	i				0	0

- 100% chance of 6"+
- Predicted heavy snow ($\geq 1''/hr$)

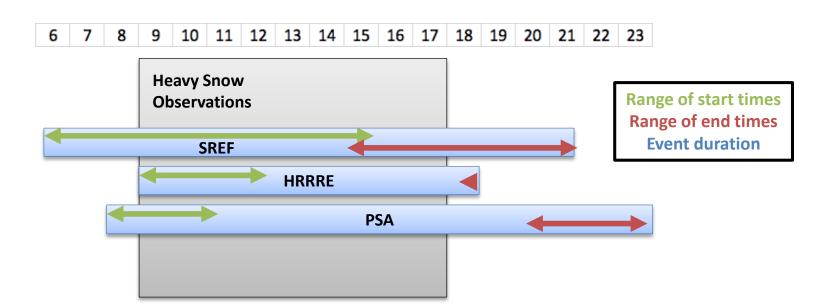
From Long Term Discussion:

- Gusts to 50 kts (~58 mph)
- Visibility down to ¼ mile
- Start: 8 am 11 am
- End: 8 pm 11 pm



Summary of Forecasts from March 13, 2019

All of the models handled this situation well and there was strong agreement on the timing of the heavy snow event.



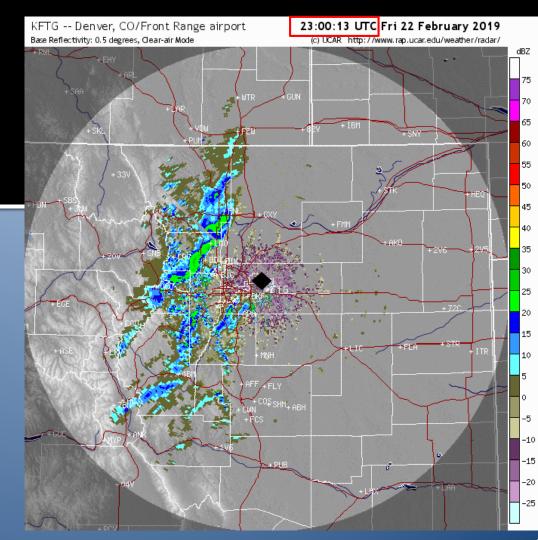
Case Study 2

February 22-23, 2019

The evening of February 22, 2019, a band of heavy snow moved over the airport that was not well forecast.

The heaviest snow occurred between 7 and 10 pm with rates exceeding 1"/hr.

In total, 8.3" of snow was recorded at the airport.

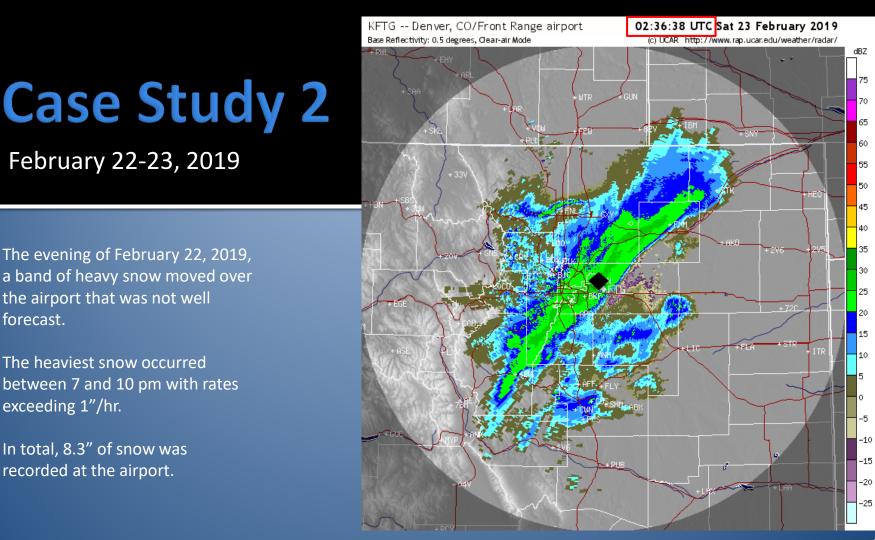


The evening of February 22, 2019, a band of heavy snow moved over the airport that was not well forecast.

February 22-23, 2019

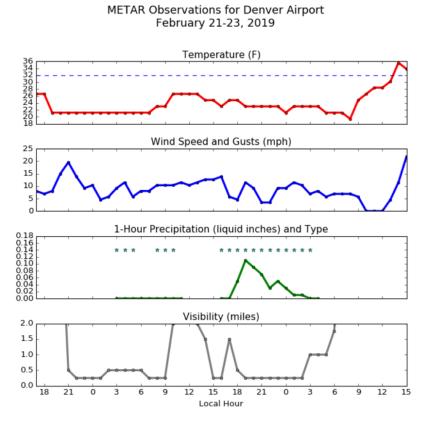
The heaviest snow occurred between 7 and 10 pm with rates exceeding 1"/hr.

In total, 8.3" of snow was recorded at the airport.



Airport Observations

- Heavy snow from 7 pm to midnight
 - Light snow to 4 am
- Light winds
- Visibility: ¼ mile

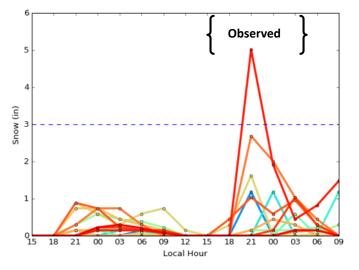


Ensemble Forecasts

SREF forecast from 8 AM on Day Prior:

- Only two members got near heavy snow threshold (≥ 1"/hr)
- Many members didn't have snow at all

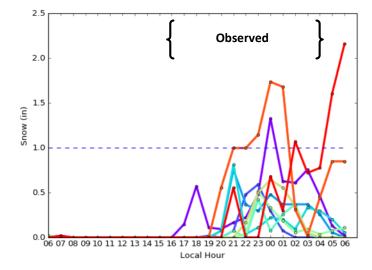
SREF 3-Hour Snow Forecast for Denver Airport February 21-23, 2019



HRRRE forecast from 5 PM on Day Prior:

- Forecast from previous evening shows more snow than SREF
- All 9 members had some snow
- 3 out of 9 members had heavy snow





PSA from 9:33 AM on February 21, 2019

Developmental Probabilistic Snow Accumulation Forecast National Weather Service Denver/Boulder, Colorado 0933 MST Thursday, February 21, 2019

This developmental product contains the probabilities (%) that the given snow amounts (inches) will occur at Denver International Airport during the respective 6 - or 12-hour periods. The probabilities indicate the amount of snow that will fall from the clouds. They do not consider the amount of melting that may or may not occur on various ground surfaces. [T (trace) means more than zero.]

Short Term Discussion (21/12Z-22/12Z): It will be dry through the early evening before a push of cold and moist air arrives around midnght. At that time, stratus will increase with a chance of light snow showers and/or freezing drizzle/fog through Friday morning. Snowfall amounts will not be much of an issue through early Friday morning as amounts are expected to be less than an inch. However, the freezing drizzle or fog will make pavement and elevated surfaces slick and icy.

Day/Hour GMT	0"	T or more	1" or more	3" or more	6" or more	10" or more
21/12-21/18 21/18-22/00	100 100	0 0	0 0	0 0		
21/12-22/00 22/00-22/06	70	30	0	0	0	0
22/06-22/12 22/00-22/12	50	50	0	0	0	0

Long Term Discussion (22/12Z-23/12Z): Areas of freezing drizzle and possibly snow are expected to continue Friday morning, with some warming and drying after 18Z. Snowfall amounts would be less than a quarter inch, but surfaces may be slick due to freezing drizzle. As the storm system progresses east across the Rockies, a cold front will push south after 00Z with the best chance for snow between 03-00Z Friday evening where around a half to one inch of snow is anticipated. There is a low chance of getting up to 3 inches during that time frame if a band sets up across the airport. Confidence in forecast is lower than usual, drier air could end the freezing drizzle sooner than expected Friday morning.

Day/Hour GMT	0"	T or more	1" or more	3" or more	6" or more	10" or more
22/12-22/18 22/18-23/00 22/12-23/00	50 80	50 20	0 0	0 0	0	0
23/00-23/06 23/06-23/12 10/14/2(23/00-23/12	30 60	70 40	30 0	10 0	0	0

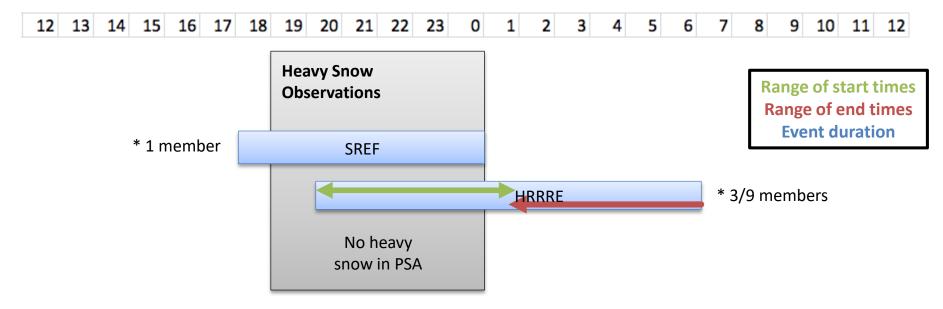
 Most likely snowfall amount is trace to 1" between 5 pm and 11 pm

Long Term Discussion:

- Best chance of snow between 8 pm and 2 am
 - ½ to 1" of snow
 - Up to 3" possible if banding
- Lower confidence forecast

Summary of Forecasts from February 21, 2019

The SREF and HRRRE had hints that this event would occur, but the timing was slightly off.



Forecast Impact and Quality Assessment Section

Day of the Event Forecasts

10/14/2020

Forecast Impact and Quality Assessment Section

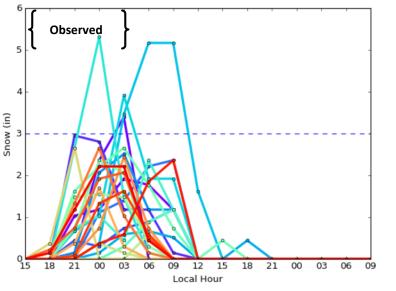
35

Ensemble Forecasts

SREF forecast from 8 AM on Day Of:

- Now more members depict heavy snow (or nearly)
- Start: 5 pm to 8 pm
- End: 5 am to 8 am the next day

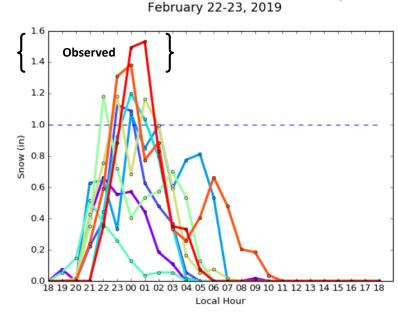




HRRRE forecast from 5 AM on Day Of:

HRRRE 1-Hour Snow Forecast for Denver Airport

- Now more members depict heavy snow
- Start: ~9 pm to 10 pm
- End: 10 pm to 1 am



10/14/2020

Forecast Impact and Quality Assessment Section

PSA from 9:25 AM on February 22, 2019

Developmental Probabilistic Snow Accumulation Forecast National Weather Service Denver/Boulder, Colorado 0925 MST Friday, February 22, 2019

This developmental product contains the probabilities (%) that the given snow amounts (inches) will occur at Denver International Airport during the respective 6- or 12-hour periods. The probabilities indicate the amount of snow that will fall from the clouds. They do not consider the amount of melting that may or may not occur on various ground surfaces. [T (trace) means more than zero]

Short Term Discussion (22/12Z-23/12Z): Stratus will continue with very light snow showers or freezing drizzle through this morning. Conditions should gradually improve after 18Z this morning. Snow is expected to develop late this afternoon and evening with the most likely onset time of around 01Z or 02Z. Localized bands of moderate to heavy snow may occur through about 10Z Saturday morning. 3 to 6 inches of storm total snow is expected with the highest snowfall rates expected to be between 1/2 to 1 inch per hour. Gusty north to northwest winds will develop later tonight after 06Z with gusts of 15-20 kts.

Day/Hour GMT	0"	T or more	1" or more	3" or more	6" or more	10" or more
22/12-22/18 22/18-23/00 22/12-23/00	60 60	40 40	0 0	0 0	0	0
23/00-23/06 23/06-23/12	0 10	100 90	80 70	70 30	Ŭ	· ·
23/00-23/12					10	0

Long Term Discussion (23/12Z-24/12Z): Snow should be winding down after 12Z/5AM but there are indications that it could hold on through 8 or 9 AM on Saturday. Additional amounts after 5 AM should be under an inch but pavement temps should continue to be below freezing until about 18Z/11AM when clouds thin and some sun starts to come out. Winds out of the north should remain 10 mph or so, limiting the blowing snow potential once it stops falling from the sky. Conditions should rapidly improve for snow/ice melting on the concrete surfaces after 11AM Saturday.

Day/Hour GMT	0"	T or more	1" or more	3" or more	6" or more	10" or more
23/12-23/18 23/18-24/00	50 100	50 0	10 0	0 0		
23/12-24/00 24/00-24/06 10/14/2 24/06-24/12 24/00-24/12	100 100	0 0	0 0	0 0	0	0

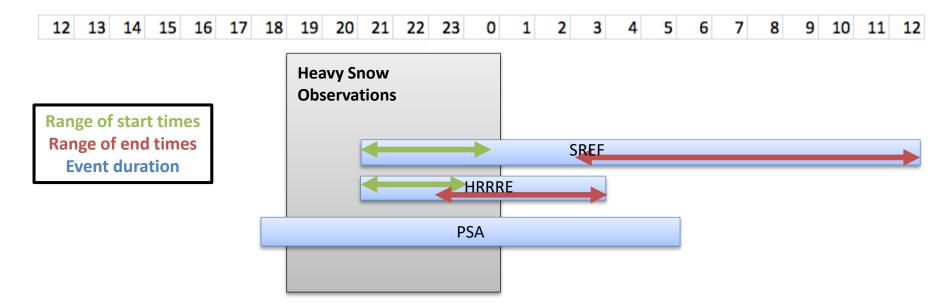
 Updated to higher likelihood of more snow

Short Term Discussion:

- Start time of 6 or 7 pm
- Bands of heavier snow possible through 3 am
- Storm total = 3" to 6"
- Snow rates of ½ to 1"/hr

Summary of Forecasts from February 22, 2019

Forecasts the day of the event improved, but still did not have a good handle on the timing or total accumulation of the event.



Forecast Impact and Quality Assessment Section

Assessment Conclusions

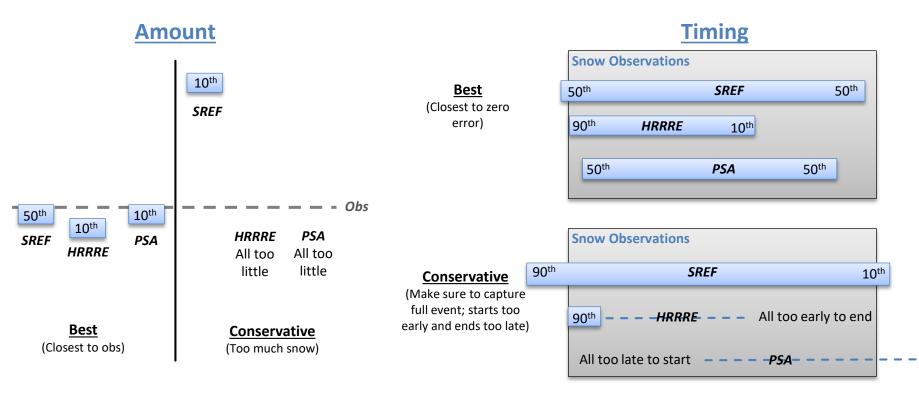
While models have improved in recent years, forecasting specific snow amounts is still challenging. These ensembles do a decent job of predicting when snow will start, but are less skilled at predicting when snow will end.

Details...

- SREF generally predicts too many events while HRRRE forecasts too few
 - HRRRE produces less false alarms, but misses often
- Snow amounts:
 - SREF has slightly too much snow
 - HRRRE has too little
 - If HRRRE doesn't have anything, doesn't mean it won't snow
- Events are mostly too short in duration
 - Larger cessation errors than onset errors
- Less skill when including high winds and low visibility

Summary Schematics

- Choosing the optimal percentile for each product produces fairly accurate amount and timing forecasts
 - Exception: HRRRE always ends too early
- The conservative approach can have larger errors, but less likely to be caught off guard



90th

Communicating Uncertainty in Weather, Climate, and Hydrological Predictions: Recent Progress and a Path Forward

> Rebecca E. Morss National Center for Atmospheric Research Boulder, Colorado

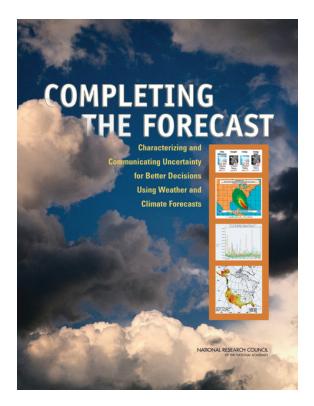
Thanks to: colleagues, funding from NSF and NOAA



http://www.mmm.ucar.edu/wrads



Communicating Uncertainty



2006 National Academies study on Characterizing and Communicating Uncertainty

"Hydrometeorological prediction is **inherently uncertain** and information about such uncertainty **should be helpful to users** in their decisions.

Although some products contain uncertainty information (e.g., ...), relatively little headway has been made in **supplying actionable uncertainty information** to most of the user community." (p. 10)

Recent progress (since ~2006)

- Continued advances in:
 - Generation of forecast uncertainty information
 - Community discussion
- New advances in:
 - Research on forecast communication, interpretation, and use (including uncertainties)
- Some progress in:
 - Incorporating perspectives from other fields and risk domains
- Overall: Significant progress ... but still significant gaps between the production of forecast information and its communication and use



BY PAUL A. HIRSCHBERG, ELLIOT ABRAMS, ANDREA BLEISTEIN, WILLIAM BUA, LUCA DELLE MONACHE, THOMAS W. DULONG, JOHN E. GANNOR, BOB GLAHN, THOMAS M. HAMILL, JAMES A. HANSEN, DOUGLAS C. HILDERBRAND, ROSS N. HOFFMAN, BETTY HEARN MORROW, BRINDA PHILIPS, JOHN SOKICH, AND NEIL STUART

The AMS Board on Enterprise Communication set goals and prepared a road map of tasks for enterprise sectors—led by the National Weather Service—to work on together to make uncertainty information integral to hydrometeorological forecasts.

AMS Ad Hoc Committee on Uncertainty in Forecasts (BAMS, 2011)

Why the Gaps?

- Myth: If only forecast users were willing / able to understand and use uncertainty information ...
 - "If you build it, they will [should] come"
 - Focus on education
- *Reality*: Most people understand that forecasts and decisions are uncertain, and they have (often deep) knowledge about those intersecting uncertainties
 - "All the numbers are inaccurate." floodplain manager (Morss et al., BAMS, 2005)
 - Effective communication is a partnership

Lessons from risk communication research and practice

(adapted from Fischhoff 1995)

Developmental stages in risk communication:

- 1. "All we have to do is get the numbers right"
- 2. "All we have to do is tell them the numbers"
- 3. "All we have to do is explain what we mean by the numbers"
- 4-5. All we have to do is show them that it has value
- 6. "All we have to do is treat them nice" (when communicating)

"People want to be treated respectfully ... "

7. "All we have to do is make them partners"

"Other things being equal, risk data should be collected, vetted, and presented in ways that suit the audience that they are meant to convince."

What is currently limiting progress?

- Myth: People only want / are willing to use deterministic (single-valued) forecasts ...
 - ... Because they want forecasters to make the decision for them
- Reality: "People are immersed in a vast sea of continuously evolving risk information..." (Morss et al., BAMS, 2017)
 - Other uncertainties can swamp [weather] forecast uncertainties
 - Given the many other complexities involved in decisions, complex meteorological information may overcomplicate

Example: Airport snow and ice control

- Airport decision makers decide on staffing for snow and ice control crews 8-24 hours before snow begins
- Decision is forecast-based, so they monitor forecasts from multiple sources starting several days in advance
- A good venue for providing probabilistic forecasts?

Perhaps, but:

- Most forecasts are for snowfall, but impacts are from accumulation
- The amount of information involved in decisions creates challenges for using complex forecasts



www.m-bco.com

What is currently limiting progress?

- Myth: People make one-time, yes/no decisions based on static, definable weather-related thresholds
 - "Supply actionable information" by providing probabilities of exceeding those thresholds
- *Reality*: Decisions are complex, multi-dimensional, interactive, and evolving
 - Decisions (and thus "thresholds") vary with circumstances
 - What appears to be a one-time, threshold-based decision may be much more complex

Example: Airport snow and ice control

 Airport decision makers use forecasts to declare a "snow alert" level, which determines snow/ice control staffing

Snow Alert Level A Declaration Criteria

- Total forecast accumulation is greater than 1 inch up to 3 inches of snow and/or ice on runway surfaces.
- The AOM considers it is in the best interest of the airport.

Snow Alert Level B Declaration Criteria

- Total forecast accumulation is greater than <u>3 inches up to 10 inches</u> of snow and/or ice on runway surfaces.
- The AOM considers it is in the best interest of the airport.

Upgrade/Downgrade of Snow Alert Levels

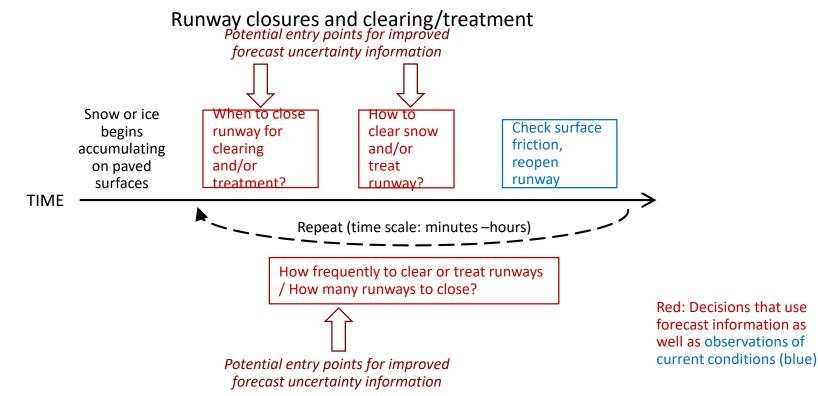
Levels can be increased or decreased based on duration of storm, temperatures, snow type, winds or in the best interest of the Airport.

What is currently limiting progress?

- Approach: Generate and communicate forecast uncertainty information by starting with meteorological variables that (may) project onto impacts or decisions
 - z = f(x,y), where x = meteorology, z = decision criteria
 - forecasting meteorology → forecasting impacts or risks
- *Reality*: Projection of meteorological variables onto decisions can be very small or complex
 - Decision makers start with their decisions, not meteorology

Example: Airport snow and ice control

Airport decision timeline during snow/ice event:



A proposed path forward

- Start with societal perspectives, problems, and decisions, and explore whether, what, and how meteorology-related information and uncertainties project onto those
 - Meet people where they are
- How to bridge the Forecast-to-Decision divide in a generalizable, rather than case by case, way? (adapted from Neilley and Williams)
 - A grand challenge question that requires synthesizing ideas and approaches across domains of expertise — i.e., a fundamentally *interdisciplinary* problem (Morss, Lazrus, and Demuth, in press)

