From: Richard Grumm - NOAA Federal <<u>richard.grumm@noaa.gov</u>> Subject: Re: here we go again - the "Monday" storm - harmlessly out to sea?????

Date: 29 January, 2015 at 20:06:50 GMT To: <<u>MAP@LISTSERV.ALBANY.EDU</u>>

Reply-To: Richard Grumm - NOAA Federal <<u>richard.grumm@noaa.gov</u>>

Jeff

It would appear that the NCEP models and SREF runs around 06-09-12 UTC Sunday caused people to jump. I suspect, do not have, the EC aided with the *confirmation bias*.

The most amazing thing here here is how the SREF and GFS/GEFS backed off after 1200 UTC. The GFS jumped earlier and backed off faster. Why?

Images show sequences of GFS, NAM, and SREF around the time people thought it was time to fish or cut bait and then covering the time when thing began to diminish in potential.

Once the jump was made, there was apparently no undo button.

So, why does one act when the see the potential but they do not act when the see the diminishing potential? *Loss aversion?*

Rich

On Thu, Jan 29, 2015 at 1:42 PM, Jeff

Halblaub <jeffrey.halblaub@noaa.gov> wrote:

Hi Brian. I fully agree with the findings from that survey. The other thing I will add is this: my confidence in using ensembles is lessened by the fact that I'm not sure the actual event (in the real atmosphere) is within the ensemble probability distribution. If it is, it could be a low probability in the ensemble. This was especially true early in the life of the GFS ensembles (2000s), they were way too under-dispersed. While they have improved considerably, I still have doubts because of low ensemble membership. It makes me wonder why we don't have time-lagged GFS ensembles, where the full membership of the last 4 cycles are grouped together in a forward moving window.

I know there has been substantial improvement to physics diversity in the SREF. Does anyone know if this is the case in the GFS ensembles? If not,

that is an area for improvement.

The other thing that concerns me, as a forecaster, is (in theory) the majority of ensemble members should be grouped around the actual event (i.e., what will actually occur in the atmosphere is the highest probability in the ensemble). However, my perception is the ensembles aren't that good yet. So, how am I to determine if the real event is an "outlier" (a lower probability within the ensemble)? The ensemble mean/best clustering should be telling me the most probable event is X. If that is incorrect, then I am being lead astray by the ensemble mean/best clustering. Given the complexity of the atmosphere (and NWP issues/parameterizations now and in the future), will ensembles ever consistently provide the most probable outcome? I'm sure they will at times, but probably not every time. How do we as forecasters deal with that?

Jeff

On Thu, Jan 29, 2015 at 11:01 AM, Brian Colle <<u>brian.colle@gmail.com</u>> wrote:

Justin and Map:

Your point about more access to probabilistic information (which is mostly on web pages, etc...) and training matches up with a survey we did a few years ago as part of our CSTAR, which included over 100 NWS forecasters, 37 NWS admin/managers, and 18 researchers/model developers. Lots of reports were being written at the time (as Tom and Cliff mentioned), but we wanted to get some numbers from the field offices and management about some of the challenges in using ensembles in operations, which are highlighted below.. I can share some more details with those who are interested...

- Nearly all forecasters use ensembles; however, only ~45% use them "sometimes."
- The highest ranked issues for forecasters in order are (1) the lack of tool/graphics (to interpret ensembles or product dissemination), (2) ensemble data access (more members, not just mean/spread), (3) ensemble forecaster training, (4) not enough probabilistic verification to have confidence in ensembles, and (5) limited time in operations to view/interpret ensembles. The other 3 issues to chose from did not make the top 5: most clients do not want probabilistic info, ensemble resolution is too coarse to be

useful for most sensible weather, ensemble deficiencies are too large to be useful.

- For forecasters who only "sometimes use ensembles," it appears that the time to view/interpret ensembles is a greater issue (ranked #3). Some of this may be from the lack of good tools and data access as suggested by open ended questions.
- In contrast to the forecasters, managers/admin view the issue of time to view/interpret ensembles as the smallest problem on average (ranked #8 out of 8).
- Most researchers and model developers use ensembles extensively when they look at the models. The lack of tool/graphics is the largest issue on average. All others issues (verification, ensemble construction, training, etc...) were nearly tied and less of an issue than the tools on average.

There may be plans out there from the past, but to make change requires organized efforts to immerse forecasters with these new products, training, and iteration with modelers/post-processing. One NWS program I am been surprised that has not taken off (or promoted more within NCEP) is the Winter Weather Expt (WWE) at WPC. For 4 weeks mid-Jan-mid-Feb (they are in the middle of the expt now but the week off), they invite a handful of forecasters and sometimes an academic (4-5 total per week) to evaluate new winter weather tools (snow ratios, ensemble probabilities, etc...), and a lot of ensemble information is discussed for these winter storms. They also have a daily weather briefing early in the afternoon (all SOOs were notified). On the first day there were over 30 on the webinar, but it slowly reduced to 10-15 (including the Fri before the big storm). There are challenges... many of the new products, which includes the new parallel SREF, probabilistic snow products, etc... can not be shared outside the experiment walls, so it is tough for the community to iterate and help. Also, even though probabilistic snow amounts are looked at, still a deterministic snow map is made manually for a region, with the confidence level attached. If the NWS and community are serious about change, these sort of winter storm test-bed opportunities should be expanded (BIG TIME) to include a broader effort involving more

forecasters/developers/researchers, etc..., even if it is remote telecom each day. I am guessing that about 1/2 on this listserve were not aware of the WWE, in which one goal is to use ensembles more in operations..

Brian

On Thu, Jan 29, 2015 at 10:28 AM, Justin Arnott <justin.arnott@noaa.gov>

wrote: Rich/others,

The automation discussion is an interesting one.

Forecasters today are in a tough position. After this event, we're writing that they degraded the forecast and that we need to move towards a more probabilistic approach...perhaps automating all of it.

But let's look at what they're up against. We have seen awesome displays of probabilistic/reforecast information on this email string, but these all come from various webpages. So, a forecaster today, without a big push in training in how to use probabilistic information sees a handful of deterministic models in AWIPS 1/2, half of the GEFS and some SREF information. They go to their favorite ensemble webpage and see a bit more probabilistic information. Many of the resources that have been posted on this email list are completely foreign to them. The forecaster at their neighboring offices may/may not look at those same webpages so collaboration is a challenge.

At the end of the day, they have a SnowAmt grid to populate that sends a <u>single value</u> to the public. How can they win? So they lean towards the Euro and get burned. Who can blame them...esp after Sandy?

In my opinion...the forecasters didn't degrade the forecast....the system did. Give forecasters a system (which we have seen encouraging emails about string) with full access to probabilistic information <u>and</u> a way to incorporate it into their forecasts. Train them in these "higher-level" data and how to get out of the "model of the day concept" which is so prone to personal biases. With this system in place, forecasters may surprise us with what they can do.

The final step is to show forecasters that people want (and can use) this information. I commend OKX for their facebook post (attached) that highlighted the min/max/most likely snowfalls. If you read the comments, though, many people took the graphic to mean "we have no clue". Perhaps we're not talking about higher-end users here, but I'm sure there are plenty of our core partners that wouldn't know what to do with a 2-24" range either. Education needs to happen inside and out before we can leverage these data effectively to move towards a Weather Ready Nation.

My opinion only,

Justin

Justin Arnott - Science and Operations Officer National Weather Service - Gaylord, Michigan <u>989-731-3384 x766</u> justin.arnott@noaa.gov

On Thu, Jan 29, 2015 at 8:39 AM, Richard Grumm - NOAA Federal <<u>richard.grumm@noaa.gov</u>> wrote: Neil

So, all the examples and presentations you've done about using EFS data , you biased your forecasts **on a single deterministic model** and not on ensemble forecast system (EFS)? How human and how refreshingly honest.

I would bet an AI program using a single EFS, deterministic models, or all available EFS data would never do that.

The books Automate his: How Alogrithims Took over Markets, our Jobs, and the World; The Second Machine age, and the more recent Isaacson book: The Innovators: How a group of Hackers Geniuses and Geeks Created the Digital Revolution should be on your reading list.

I think a good blended forecast might have and *in fact did* show the QPF often did not support the robust snow forecasts, at least on the edges to any comfortable degree. It also showed very sharp edges and *life on the edges is certainly risky*. All forecast systems showed lower probabilities of higher QPF amounts and considerable spread on the edges, particularly the western and northern edges. And across Long Island!

I also think this past storm had some predictability horizon issues so the window to actually predict this is was relatively short, on the order of a few days. Trevor A. showed this with the images he has sent along with GEFS QPF and QPF M-Climate data. Tim H showed this quite nicely too.

At the risk of tar and feathers, one could easily see a worded forecast from the SREF plumes and how it would have spoken of the large spread in the NYC area and the QPF for snow amounts probably had a great range. The Mel Brooks BOT would have noted the high uncertainty...or was it anxiety? Using Tim H's example a BOT would have said something like...as for heavy snow over 1 or 2 feet in NYC...*I feel strongly either way*. (We need jocular BOTS).

I think our inability over the past 10-15 years to properly leverage these data is screaming for more A U T O M A T I O N. We could call the first such system SPOCK and it could be free of human bias and emotion. We just need to pay a tad extra for the green wires. LOL

Food for thought and of course not to be taken quite so seriously.

Rich

To unsubscribe from the MAP list, click the following link and send the email generated:

MAP-SIGNOFF-REQUEST@LISTSERV.ALBANY.EDU

To unsubscribe from the MAP list, click the following link and send the email generated: MAP-SIGNOFF-REQUEST@LISTSERV.ALBANY.EDU

Jeff Halblaub NOAA / National Weather Service 6365 North Osborne Drive West Hastings, NE 68901-9163 <u>402.462.4287</u> weather.gov/hastings

To unsubscribe from the MAP list, click the following link and send the email generated:

MAP-SIGNOFF-REQUEST@LISTSERV.ALBANY.EDU

Richard Grumm Scientific Operations Officer National Weather Service Office State College PA 16803 http://nws.met.psu.edu/