NDFD Statistics Viewer

John Wagner & Dana Strom

Planned Updates to Viewer on AWS

- Current NDFD Sample Matching

NDFD	Blend/Blendx		GMOS		WPC (day 4-7)	
IssuanceTime	Model Cycle	Available Time	Model Cycle	Available Time	Model Cycle	Available Time
00z	19z	20z	12z	18z	12z	16z
06z	01z	02z	12z	18z	00z	04z
12z	07z	08z	00z	06z	00z	04z
18z	13z	14z	00z	06z	12z	16z

Planned Updates to Viewer on AWS

- New NDFD Sample Matching - CONUS/Puerto Rico

NDFD	Blend/Blendx Early		Blend/Blendx		GMOS		WPC (day 4-7)	
Issuance Time	Model Cycle	Available Time	Model Cycle	Available Time	Model Cycle	Available Time	Model Cycle	Available Time
00z	13z	14z	19z	20z	12z	18z	12z	16z
12z	01z	02z	07z	08z	00z	06z	00z	04z

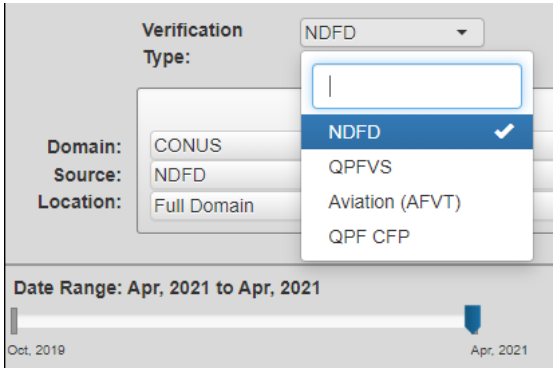
Planned Updates to Viewer on AWS

- New NDFD Sample Matching - Alaska/Hawaii

NDFD	**Blend/Blendx Early**		**Blend/Blendx**		**GMOS**	
Issuance Time	Model Cycle	Available Time	Model Cycle	Available Time	Model Cycle	Available Time
06z	19z	20z	01z	02z	12z	18z
18z	07z	08z	13z	14z	00z	06z

Planned Updates to Viewer on AWS

- Four selectable Verification Types to support unique sample matching
  - **NDFD** - NDFD/NBM/GMOS/WPC verification through 7 days
  - **QPFVS** - QPF Verification Service
  - **Aviation (AFVT)** - near-term hourly NDFD verification through 36 hours
  - **QPF CFP** - Special QPF verification for Collaborated Forecast Process demo



Planned Updates to Viewer on AWS

- Entire period of record (April 2005-present) for NDFD
- Query by Blend version number (in addition to Blend, Blendx)
- Gridded scores for just land areas, just water areas, or both
- High Impact Verification for NDFD
  - Scores specific to valid times when selected WWAs (e.g., winter weather) were in effect
- Additional elements
  - QPF24 and Snow24 in QPF CFP
  - RH/Apparent Temperature in NDFD
  - NDFD, Blend, and GLMP wind speed/gust/direction, temp, dewp, sky cover to 36 hours in Aviation (AFVT)
  - QPF01 - Blend, GFS, and HRRR to 36 hours in Aviation (AFVT)
  - PoP01 - Blend and GLMP to 36 hours in Aviation (AFVT)
  - 12-hr Max Wind Gust in NDFD
- Gridded Verification for the Guam and Oceanic domains

Known Issues

- Station PoP12
  - Probabilistic verification takes significantly longer to run in MET
  - Solution that works for the 134 gridded masks (WFOs/regions/RFCs) does not work for 2600+ stations
- Gridded Monthly PoP12
  - Daily PoP12 brier scores are available
  - Currently missing from Mar-Present due to an offset problem with WPC data
- We rely on unreliable sftp to move data from WCOSS to MDLnet
  - Will no longer be an issue when viewer/db is in the cloud

Data Usage

- Sufficient Data for Contingency Table Scores (Heidke Skill Score, Critical Success Index, Probability of Detection, False Alarm Ratio, etc)
  - Looking at small samples of data for these scores can produce unexpected results. If thresholds are not exceeded, no results will be displayed.
  - We recommend selecting at least 30 days of data when looking at these scores
- Blendx
  - Blend V4.1 started on May 20, 2021
  - Blend V4.0 spans February through September 2020
  - Blend V3.2 spans October 2019 through January 2020
- Quality Control
  - QC for gridded data currently using “common sense” thresholds and known issues that we have encountered.
  - If you encounter data that you believe incorrectly got through our QC process, please let us know via VLab/slack.

More information about MET

- DTC Homepage
- MET V9.1 Users Guide
- Existing builds on NOAA Machines

NDFD Statistics Viewer

- NDFD Statistics Viewer

https://veritas.mdl.nws.noaa.gov/ndfd-stats/comparative/verification.php

- VLab Community

https://vlab.ncep.noaa.gov/group/mdl-forecast-evaluation

- Slack

https://noaa-mdl.slack.com/archives/C01AC41QVFC

- QPFVS

https://veritas.mdl.nws.noaa.gov/qpfvs/verification.php

Backup Slides

NDFD Statistics Using MOS-2000

- Gridded and station verification for months only
- Relied on monthly station obs file from the Statistical Modeling Division
- Run after the 15th of the following month
- 63-step process to run the verification
- Plots and tables available to users limited to monthly averages
- Station data limited to WFO averages

NDFD Statistics Using Model Evaluation Tools (MET)

- Hourly verification
  - 7 day delay for QPF and PoP - to get final URMA grid with RFC updates
  - 1 day delay for other elements
- Quality control station data using the Community Atmospheric Modeling Post-processing System (CAMPS)
- Data is ingested nightly into the database
- Currently automated using crons. Upgrade to rocoto as job scheduler coming soon.
- New viewer allows users to verify custom ranges for grids
- Monthly station verification can be plotted for an individual station
- Monthly station and gridded verification still runs around the 15th of the following month, when all necessary obs are available

How MET Differs From MOS-2000

- Grib2/netCDF files instead of TDLPack
- Run for a single model/element/projection
  - Case and Gridpoint matching done externally
  - Data must be preprocessed for thresholds
- MET is not optimized to run on a supercomputer
- Using MET ensures consistency of scores among users

What is Verified Currently

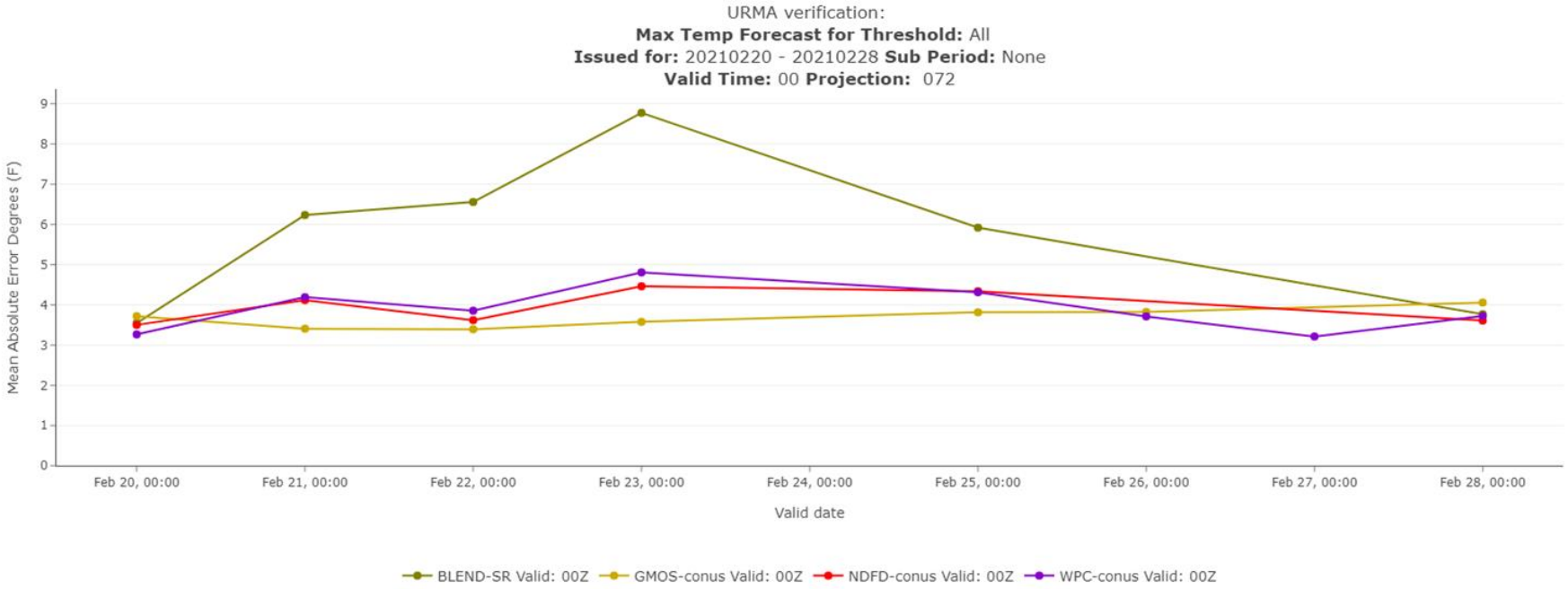
	NDFD				Blend/Blend Para				GMOS				GLMP				WPC			
	CO	AK	HI	PR	CO	AK	HI	PR	CO	AK	HI	PR	CO	AK	HI	PR	CO	AK	HI	PR
Ceiling	X				X								X							
Dewp	X	X	X	X	X	X	X	X	X	X	X						X			
MaxT	X	X	X	X	X	X	X	X	X	X	X						X			
MinT	X	X	X	X	X	X	X	X	X	X	X						X			
PoP12	X	X		X	X	X		X	X	X							X			
QPF06	X	X		X	X	X		X	X	X							X			
Sky Cover	X			X	X			X	X								X			
Snow06	X				X															
Temp	X	X	X	X	X	X	X	X	X	X	X									
Visibility	X				X								X							
Wave Height	X	X	X	X	X	X	X	X												
Wind Direction	X	X	X	X	X	X	X	X	X	X	X						X			
Wind Gust	X	X	X	X	X	X	X	X	X	X	X									
Wind Speed	X	X	X	X	X	X	X	X	X	X	X						X			

Known Issues

- Station PoP12
  - Probabilistic verification takes significantly longer to run in MET
  - Solution that works for the 134 gridded masks (WFOs/regions/RFCs) does not work for 2600+ stations
- Gridded Monthly PoP12
  - Daily PoP12 brier scores are available
  - Currently missing from Mar-Present due to an offset problem with WPC data
- Station MaxT/MinT
  - Working through issue with CAMPS QC process

Known Issues

- Station Verification for Guam
  - Currently working through issues with data for Guam
- Station Sky Cover Obs Require Satellite Data
  - SCE/SCW data was down for Feb-Oct 2020
- Offsets off for WPC cl/dp/mx/mn/po/ws/wd for Feb 26-May 10
- We rely on unreliable sftp to move data from WCOSS to MDLnet
  - This is no longer expected to be an issue after the db/viewer is in the cloud



- Feb 24 - Data missing for all dates means missing data or data failed QC check
- Feb 26-27 - Missing Blend data means that we likely had an sftp issue. Please let us know when you see this.

Score Groups - CNT

- Continuous Scores (CNT)
  - Mean Absolute Error (MAE) - measures accuracy of a forecast
    - Perfect Score: 0
  - Bias - indicates whether a forecast is too high or too low in predicting a certain parameter. Also known as Mean Error (ME).
    - Perfect Score: 0
  - Root Mean Square Error (RMSE) - measures the accuracy of a forecast. Since the mean is computed of the square error, more emphasis is placed on larger errors.
    - Perfect Score: 0

Score Group - CTS/CTC

- Contingency Table Scores (CTS) and Contingency Table Counts (CTC)
  - Critical Success Index (CSI) - ratio of the number of times the event was correctly forecasted to occur to the number of times it was either forecasted or occurred. Also known as the Threat Score.
    - Perfect Score: 1
  - Probability of Detection (POD) - the fraction of events that were correctly forecasted to occur. Also known as the hit rate.
    - Perfect Score: 1
  - False Alarm Ratio (FAR) - the proportion of forecasts of the event occurring for which the event did not occur.
    - Perfect Score: 0
  - Gilbert Skill Score (GSS) - the CSI, corrected for the number of hits that would be expected by chance. Also known as the Equitable Threat Score.
    - Perfect Score: 1

Score Group - CTS/CTC

- Contingency Table Scores (CTS) and Contingency Table Counts (CTC)
  - Heidke Skill Score (HSS) - skill score based on Accuracy, where the Accuracy is corrected by the number of correct forecasts that would be expected by chance.
    - Perfect Score: 1
  - Frequency Bias - the ratio of the total number of forecasts of an event to the total number of observations of the event. A “good” value of Frequency Bias is close to 1; a value greater than 1 indicates the event was forecasted too frequently and a value less than 1 indicates the event was not forecasted frequently enough.
    - Perfect Score: 1
  - Percent Correct - Ratio of forecasted events to the total number of events.
    - Perfect Score: 100

Score Group - PSTD/PCT

- Probabilistic Statistics for Dichotomous Outcome (PSTD) and Probabilistic Contingency Table (PCT)
  - Brier Score (BS) - Mean squared probability error
    - Perfect Score: 0
  - Reliability Plot - Ratio of times an event was forecast to occur versus how often it occurred.
    - Perfect Score: Diagonal line