NOAA Global Systems Laboratory

Verification and Evaluation

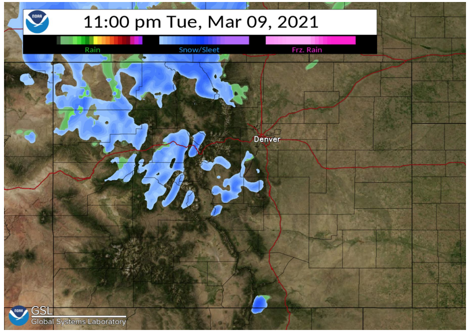
AVID/Verification and Assessment Branch
EDS/Forecast Impact and Quality Assessment Services Branch

Global Systems Laboratory

Why Verification?

Users

Model Development



Managers

Verification

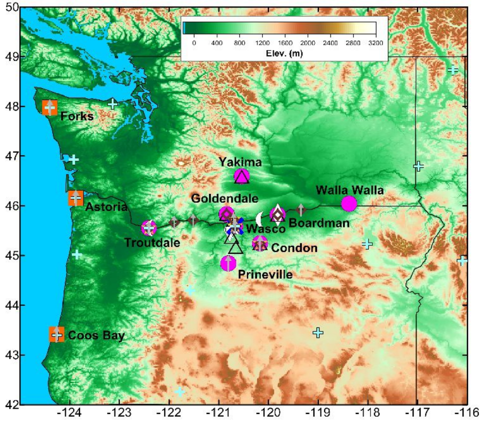
NOAA Global Systems Laboratory

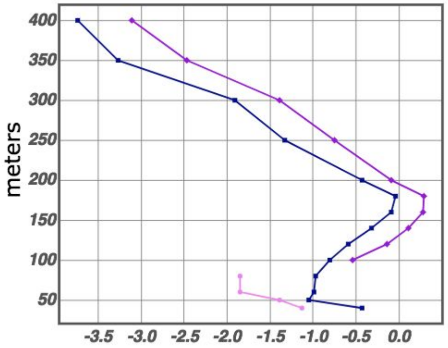
Model Development

Global Systems Laboratory

Verification System - Process Oriented Verification

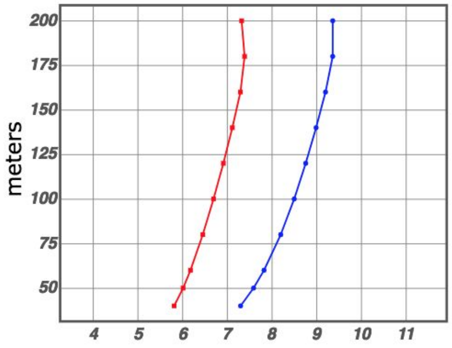
- Aggregated statistics over regions required to keep database size reasonable with rapid response times
- However, these objective metrics are hard to deconvolve to determine what atmospheric (physical) process(es) are the source of model forecast errors
- Process-Oriented Verification is a possible solution - WFIP-2 prototype





Vertical Profile of model bias when compared to Sodar (pink), Lidar (blue), and Profiler (purple) networks

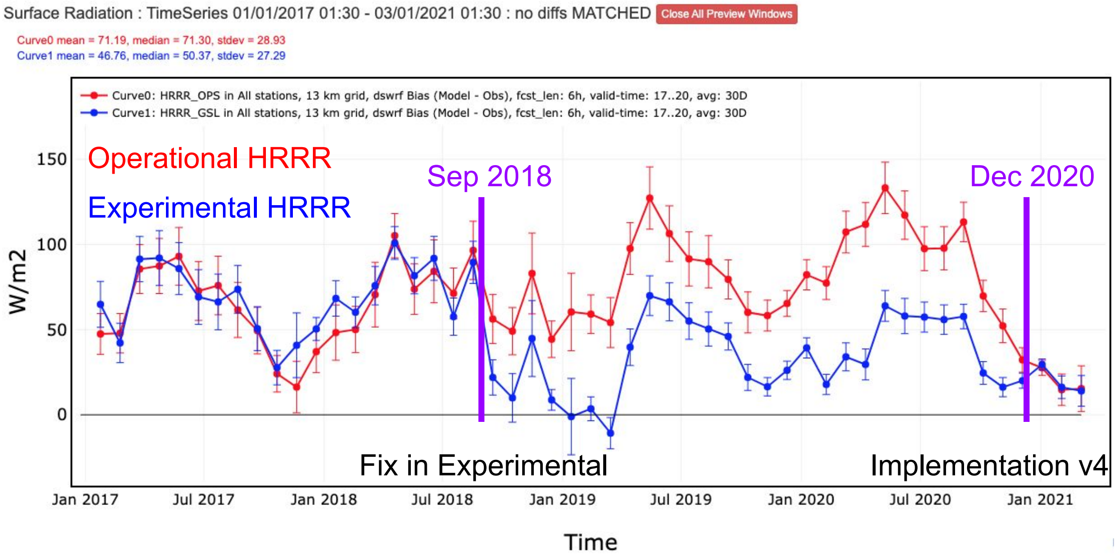
List of Discriminators
Surface Bulk Richardson Number
Surface Heat Flux
Latent Heat Flux
Stability Parameter
Friction Velocity
Max Cloud Fraction
10-m Wind Speed
10-m Wind Direction
80-m Wind Speed
80-m Wind Direction
PBL Height
Wind Shear (200-10m)
Wind Shear (50-10m)
Wind Shear (200-50m)
Vertical Temperature Difference (200m-surface)
Downward Shortwave Radiation



Vertical Profile of model wind speed filtered by downward solar radiation (red) and unfiltered (blue)

Verification System - R2O example

- Issue with how clouds were treated at the sub-grid scale, below the scale at which the High-Resolution Rapid Refresh (HRRR) can explicitly resolve them
- Modelers utilized internal solar radiation verification at all stages of development
  - Diagnosis
  - Testing
  - Long-term viability
  - Operational implementation



NOAA Global Systems Laboratory

Users

Global Systems Laboratory

Aviation Weather Research Program (AWRP)

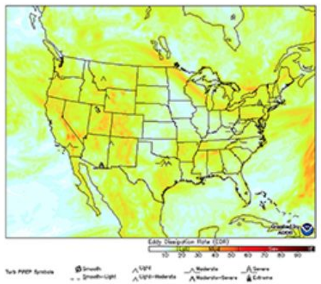
FAA's AWRP: "The goal of the research is to transition new or improved weather capabilities…[to] enhance aviation safety and efficiency."

GSL provides third-party, independent, evaluations of FAA-funded weather products to inform R2O decisions.

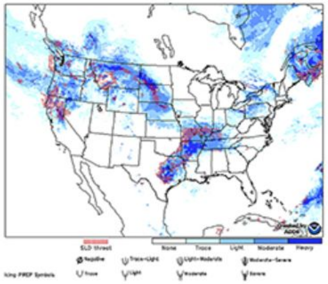
Assessments drove decisions to...

- Select the Localized Aviation MOS Product (LAMP) to provide ceiling and visibility data for the Helicopter Medical Emergency Services (HEMS) Tool.
- Delay implementation of the Offshore Precipitation Capability (OPC) to improve product performance.

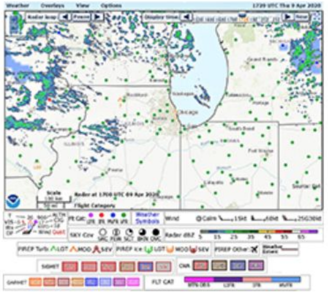
Graphical Turbulence Guidance



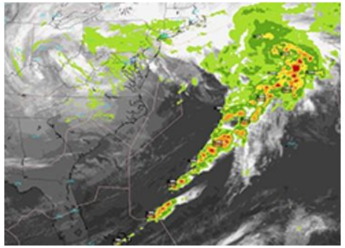
Forecast Icing Product



Helicopter Emergency Medical Services (HEMS) Tool



Offshore Precipitation Capability

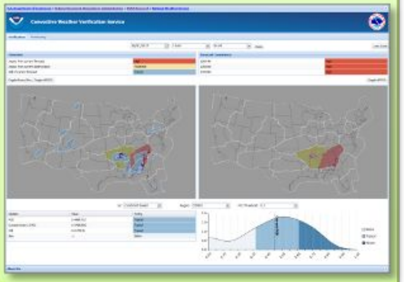


Verification Tools--NWS

CWVS

Convective Weather Verification Service

Impact-based verification of TCF



TFVT

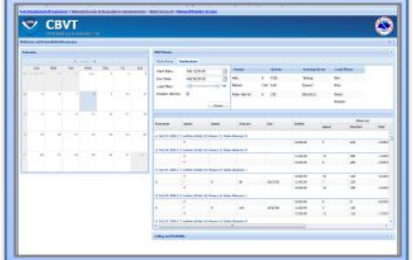
TRACON Gate Forecast Verification Tool

TRACON Gate forecasts for convection at the terminal

CBVT

CWSU Briefing and Verification Tool

Forecaster briefings to TMU for wind shift, C&V events



EVENT

Event-based Verification and Evaluation of NWS Gridded Products Tool

Gridded products for thunderstorm events, terminal and en-route

NOAA Global Systems Laboratory

Management/Economic Benefit

Economic Impacts (EI) - Overview

- Developing NWP models is expensive; is the Nation getting a good return on its investment?
- **EI from some tool can only occur if a decision is changed based upon that tool (e.g., the new forecast changes behavior)**
- EI is only important for certain weather conditions
- EI is usually regionally dependent
- Working with CSU Economics Dept. to evaluate EI of the regular updates to our regional model HRRR
- **Estimating EI of a decision is seldom easy**

Yellow: Major metropolitan areas
Black/blue: Precipitation

Economic Impacts - Wind Energy Example

Only focused on “overprediction” errors → largest financial impact

	New ”better” than old		
	Electricity Generated [MW]	Extra costs [$M]	Potential Savings [$M]
“Actual”	579,260		
Based on HRRR1	3,328,406	75.6	$59.0M
Based on HRRR2	1,170,927	16.6	
“Actual”	873,022		
Based on HRRR2	2,162,395	36.7	$17.4M
Based on HRRR3	1,547,034	19.3	

14 month period

11 month period

Summary

- Developed range of tools for objective NWP model evaluation
  - Goal: to help the model developers improve the models more efficiently
  - Many different statistics over numerous variables and regions
  - Working to identify physical processes that aren't represented correctly
- Impact-based assessment
  - Goal: demonstrate improved functionality and accuracy of new models for specific, high-visibility stakeholders
  - Evaluate the forecast systems from the stakeholder's perspective
  - Primarily objective based, but subjective interpretation also
- Economic assessment
  - Using economic models to translate model improvements into societal savings

Summary

Global Systems Laboratory

- Automated verification tools and metrics to quantify model improvements demonstrate *performance*
- Innovative techniques such as impact-based and process-oriented evaluations demonstrate *quality*
- Independent model assessments used by NWS and FAA as decision points for transitioning models and tools to operations within NWS demonstrate *relevance*
- Significant economic benefits in several sectors demonstrate *quality and relevance* of the HRRR model to society

Summary

Performance

- Automated verification tools and metrics to quantify forecast model and improvements

Quality

- Innovative techniques
  - Impact-based
  - Event-based
  - Process-oriented
- Improved models to NWS operations

Relevance

- Independent model assessments used by NWS and FAA as decision points for transitioning models
- Involved in nearly everything GSL does

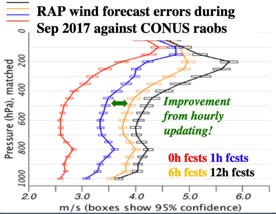
Verification across GSL

Data Assimilation | Prediction | Decision Support | Advanced Technology

Data Assimilation--Obs Impact

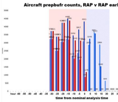
Frequent updates allow:
- Use of the **most recent weather observations** for improved forecasts.
- **Updated information** for decision makers.



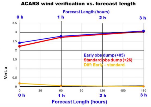



Prediction Across Scales

Transition from 6-hr to 1-hr global data assimilation cycles at 13 km scale
- Collaboration underway between GSL, PSL, EMC and JCSDA
- Produce hourly global "data dumps"
- Test assimilation strategies to
  - mitigate noise from frequent analyses
  - produce lower-latency shorter-term forecasts with smaller obs windows
  - collect more latent observations for longer-term forecasts to "do no harm"

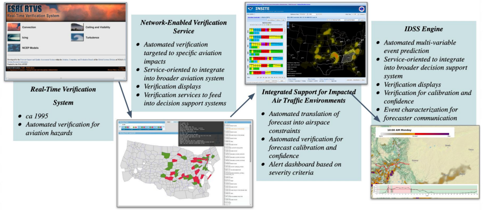


Removal of higher latency aircraft observations (red) in HRRR due to 20 min earlier data cutoff



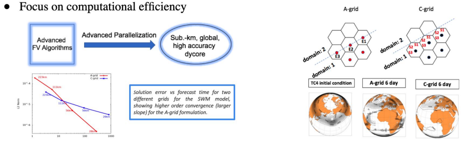
Small increase in HRRR wind forecast error from data loss

Decision Support



Advanced Technology - Exascale Computing

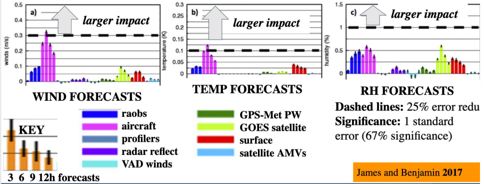
- Comparison of numerical accuracy for shallow water icosahedral A vs. C grid
- Improved algorithm resulted in better stability and higher accuracy
- Focus on computational efficiency



Yu, Y. G., Wang, N., Middlecoff, J., Peixoto, P. S., & Govett, M. W. (2020). Comparing Numerical Accuracy of Icosahedral A-Grid and C-Grid Schemes in Solving the Shallow-Water Model. Monthly Weather Review, 148(10), 4009-4033.

Data Assimilation--Obs Impact

- **RAPv3** data denial experiments carried out for three multi-season 10-day periods. Verification against North American raobs.

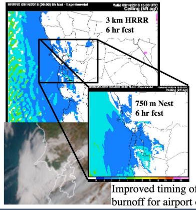


Dashed lines: 25% error redu

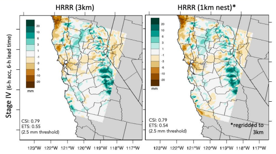
Significance: 1 standard error (67% significance)

James and Benjamin 2017

Prediction Across Scales



Improved timing of SFO fog burnoff for airport operations

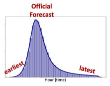


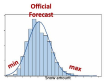
Evaluation of QPF biases with finer orographic detail (English et al. 2021)

Decision Support

Uncertainty Information

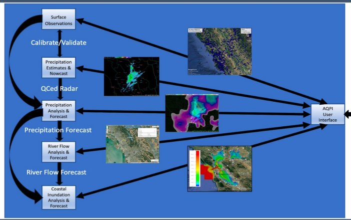
Confidence Information





- Do these distributions represent the range of solutions to the communication of uncertainty?
- Can **verification** information improve them?
- Can **verification** influence forecasters' confidence?

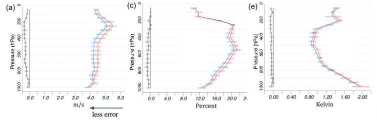
Advanced Technology - AQPI



Data Assimilation

- 3DEnVar is very successful for regional hourly data assimilation in RAP/HRRR and has superior skill over purely a 3D-variational analysis (shown below).
- Ensemble background error covariance data from the global system, GDAS, is effective.

6-hour Forecast RMS Wind, Relative Humidity, Temperature

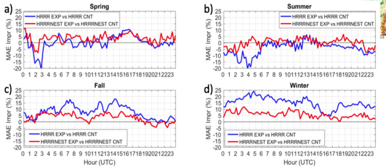


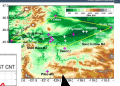
3DVar
3DEnVar
Difference

Hu, M., et al., 2017.

Prediction Across Scales

Seasonal dependence of 80-m wind speed validation in Columbia River Gorge with 750 m HRRR nest accuracy (red)

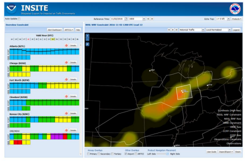




Validation against 19 sodars (white and pink crosses)

Olson et al 2019 BAMS

Decision Support



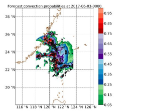
- Automated verification for forecast calibration and confidence
- Verification can influence forecasters' confidence!

One of the first applications to quantify confidence based on past performance of forecast sources.

Advanced Technology - Machine Learning

Improve performance and speed of traditional methods

Convection Initiation Detection and Monitoring
- Using AI to create probabilistic forecast of areas of likely convection initiation with various lead times (+30 to +180 minutes)
- Collaborating with Taiwan Central Weather Bureau



Cyclone Detection
- Using AI to detect and target areas for further data extraction
- Working with ECMWF to test in pre operational capability