Strategic Implementation Plan (SIP) for a Community-based Unified Forecast System

Verification and Validation Working Group

Presented by

Tara Jensen, NCAR and DTC

Presented at SIP Coordination Meeting

May 14-16, 2019; College Park, MD

V&V WG *Membership*

- *Tara Jensen (NCAR & DTC)***
- *Geoff Manikin (NOAA/EMC)***
- *Jason Otkin (U of Wisc Madison)***
- *Ivanka Stajner (NOAA/EMC)***
- *Zhuo Wang (U of Ill U-C)***
- *Ben Albright (NOAA/WPC)*
- *Mike Baldwin (Purdue)*
- *Jimmy Correia (NOAA/AFS)*
- *Surya Dutta (JCSDA)*
- *Burkely Gallo (NOAA/SPC)*
- *John Halley Gotway (NCAR & DTC)*
- *Jeff Hamilton (NOAA/GSD)*
- *Matt Jainga (NRL)*
- *Israel Jirak (NOAA/SPC)*
- *Mark Klein (NOAA/WPC)*
- *Arun Kumar (NOAA/CPC)*
- *Jason Levit (NOAA/EMC)*
- *Sarah Lu (SUNY Albany)*
- *Mariusz Pagowski (NOAA/GSD)*
- *Tanya Peevey (NOAA/GSD)*
- *Ricky Rood (U Mich)*
- *Patrick Skinner (NOAA/NSSL)*
- *Nathan Snook (NOAA/NSSL)*
- *Ryan Solomon (NOAA/AWC)*
- *Dana Strom (NOAA/MDL)*
- *Bonny Strong (NOAA/GSD)*
- *Laurie Trenary (George Mason U)*
- *Xuguang Wang (OU)*
- *Betsy Weatherhead (Jupiter Intel)*
- *Dave Zelinsky (NOAA/NHC)*
- *Chidong Zhang (NOAA/PMEL)*

Overhauled: April 2019
31 Members

- *Co-Chair ***

V&V WG *Membership*

- *Tara Jensen (NCAR & DTC)***
- *Geoff Manikin (NOAA/EMC)***
- *Jason Otkin (U of Wisc Madison)***
- *Ivanka Stajner (NOAA/EMC)***
- *Zhuo Wang (U of Ill U-C)***
- *Ben Albright (NOAA/WPC)*
- *Mike Baldwin (Purdue)*
- *Jimmy Correia (NOAA/AFS)*
- *Surya Dutta (JCSDA)*
- *John Halley Gotway (NCAR & DTC)*
- *Jeff Hamilton (NOAA/GSD)*
- *Matt Jainga (NRL)*
- *Israel Jirak (NOAA/SPC)*
- *Mark Klein (NOAA/WPC)*
- *Arun Kumar (NOAA/CPC)*
- *Jason Levit (NOAA/EMC)*
- *Sarah Lu (SUNY Albany)*
- *Mariusz Pagowski (NOAA/GSD)*
- *Tanya Peevey (NOAA/GSD)*
- *Ricky Rood (U Mich)*
- *Patrick Skinner (NOAA/NSSL)*
- *Nathan Snook (NOAA/NSSL)*
- *Ryan Solomon (NOAA/AWC)*
- *Dana Strom (NOAA/MDL)*
- *Bonny Strong (NOAA/GSD)*
- *Laurie Trenary (George Mason U)*
- *Burkely Twiest (NOAA/SPC)*
- *Xuguang Wang (OU)*
- *Betsy Weatherhead (Jupiter Intel)*
- *Dave Zelinsky (NOAA/NHC)*
- *Chidong Zhang (NOAA/PMEL)*

"Community" Membership:
12 Members

- *Co-Chair ***

V&V WG *Membership*

- *Tara Jensen (NCAR & DTC)***
- ***Geoff Manikin (NOAA/EMC)*****
- *Jason Otkin (U of Wisc Madison)***
- ***Ivanka Stajner (NOAA/EMC)*****
- *Zhuo Wang (U of Ill U-C)***
- *Ben Albright (NOAA/WPC)*
- *Mike Baldwin (Purdue)*
- *Jimmy Correia (NOAA/AFS)*
- *Surya Dutta (JCSDA)*
- *John Halley Gotway (NCAR & DTC)*
- *Jeff Hamilton (NOAA/GSD)*
- *Matt Jainga (NRL)*
- *Israel Jirak (NOAA/SPC)*
- *Mark Klein (NOAA/WPC)*
- *Arun Kumar (NOAA/CPC)*
- ***Jason Levit (NOAA/EMC)***
- *Sarah Lu (SUNY Albany)*
- *Mariusz Pagowski (NOAA/GSD)*
- *Tanya Peevey (NOAA/GSD)*
- *Ricky Rood (U Mich)*
- *Patrick Skinner (NOAA/NSSL)*
- *Nathan Snook (NOAA/NSSL)*
- *Ryan Solomon (NOAA/AWC)*
- *Dana Strom (NOAA/MDL)*
- *Bonny Strong (NOAA/GSD)*
- *Laurie Trenary (George Mason U)*
- *Burkely Twiest (NOAA/SPC)*
- *Xuguang Wang (OU)*
- *Betsy Weatherhead (Jupiter Intel)*
- *Dave Zelinsky (NOAA/NHC)*
- *Chidong Zhang (NOAA/PMEL)*

NOAA Total: 19
3 EMC

- *Co-Chair ***

NWS: AWC, CPC, NHC, SPC, WPC, MDL, AFS
OAR: ESRL, NSSL, *PMEL*

V&V WG Accomplishments & Challenges

SIP project milestones completed/progress to date:

Project 13.1 – T&E to demonstrate operational readiness

- Held a Test Plan and Metrics Workshop; Completed FV3-GFS evaluation as part of transition to operations; Workshop findings used in Physics Testing

Project 13.2 – Unified Validation and Verification capability

- METplus integrated into parallel GFS workflow – not on operational system yet but soon – many examples now on GitHub
- Dependencies on UPP and JEDI waiting on Hurr Supp funding so delayed
- Community workflow – poised to start working with CIME project; portions integrated into Rocoto

Project 13.3 – MET component – statistical and diagnostic engine

- Expected Q1FY20 – **initial to expanded capability** for Atmospheric Composition and Air Quality, Marine, Sub-Seasonal, Large-Scale Dynamics, TC genesis, Oceans, Waves, Sea Ice, and Process-based metrics
- Land Surface Model, Hydrology – linking has not started

V&V WG Accomplishments & Challenges

SIP project milestones completed/progress to date:

Project 13.4 – Database and Display Analysis Tools (METviewer/METexpress)

- **In the cloud** with METviewer and METexpress; **In containers** METe on AWS and METv on linux servers at NSSL for HWT, MET on Army Research Lab

Project 13.5 – Protocol for community contributions

- METplus, MET and METviewer codebases in Github repositories; METexpress in VLab; will move to Github when community code
- METplus governance for community contribution in process
- **Community Interest:**
  - **NCAR, GSD, EMC currently contributing to METplus repos**
  - NOAA/MDL, USAF, and NRL funded to transition to METplus
  - The Met Office submitted proposal to adopt METplus and contribute to development (decision due in mid-summer)
  - NCAR exploring METplus as its unifying base

SIP project issues (main challenges):

- METplus is still not installed on operational portion of WCOSS (ETA: 05/24/19)

V&V WG Accomplishments & Challenges

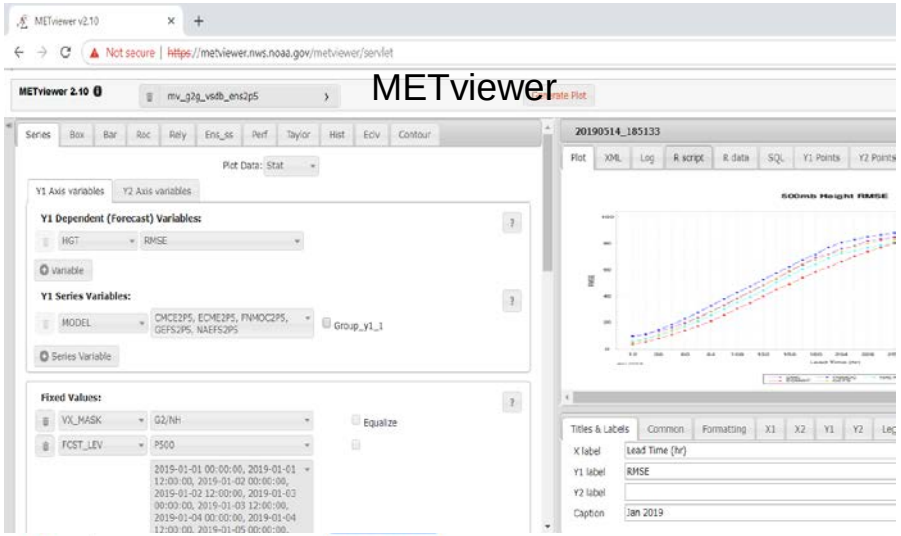
SIP project issues (main challenges):

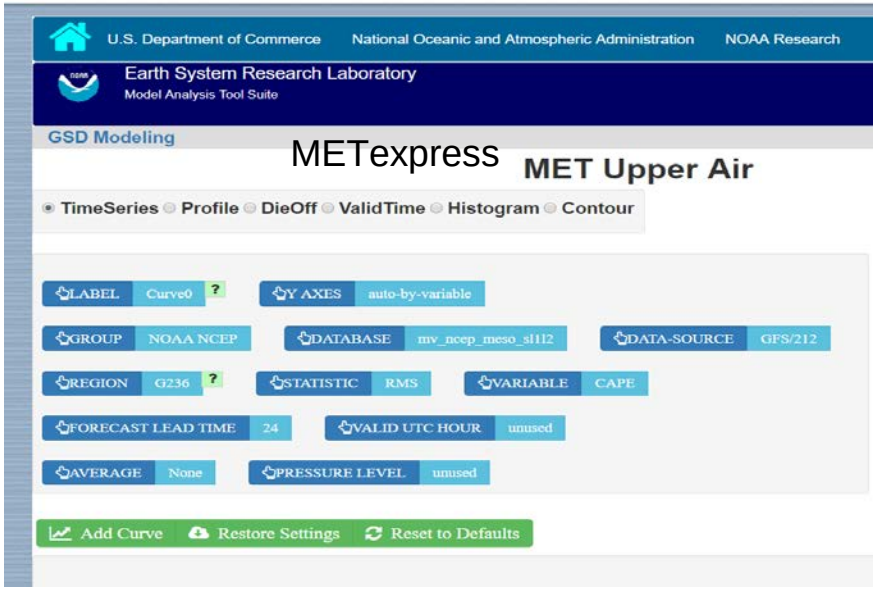
- METplus is still not installed on operational portion of WCOSS (ETA: 05/24/19)
- METplus team (currently: NCAR, GSD, EMC) working as fast as it can but seems to be under-resourced and hence highly pressured and significantly less responsive that it should be
- V&V WG having a lot of great discussions but am still trying to figure out how to move from discussion to action

V&V WG Accomplishments & Challenges

Recent METplus Enhancements

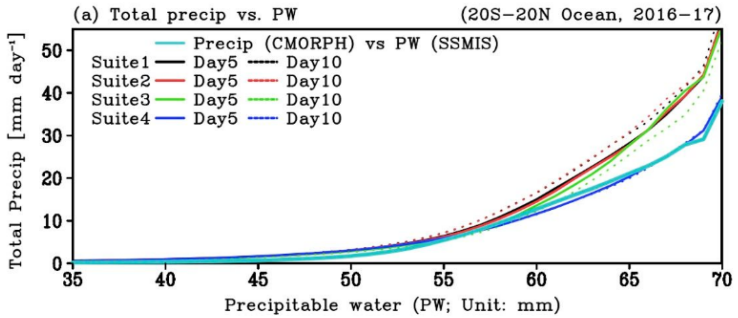
- Fortify cyber-security compliance
- Percentile thresholds
- Land/sea and topography masking
- Additional pre-processing capability in PB2NC and Pcp-Combine tools
- Gaussian interpolation for surrogate severe
- Support for GOES-16/17 fields - including “super-obbing” of high res data and interpolation to desired projection
- METviewer and METexpress sharing databases with support for MySQL, MariaDB, AuroraDB, and Couchbase





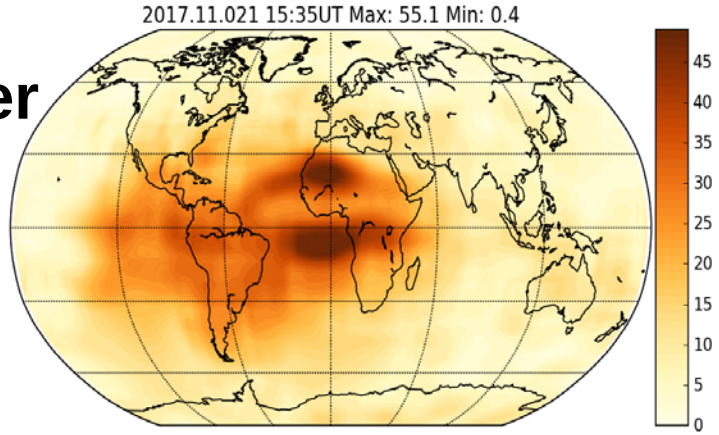
Design Philosophy: One Tool - Many Applications

S2S

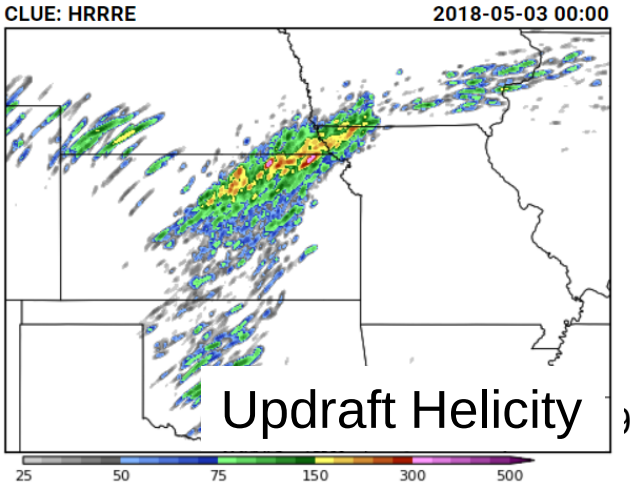


- Inventories the data
- Develops the PDF
- User configurable bins for PDF and percentiles
- Writes out – or holds in memory – bins or percentiles for use by other tools to compute diagnostics or statistics using Grid-Stat, Point-Stat, MODE, MODE Time Domain

Space Weather



CAM



V&V WG
T&E Recommendations

General Recommendations

- Consider ECMWF-like testing cycle: Alpha-phase, Beta-phase, and Release-candidate-phase testing (seems like the HTF we will discuss)
- Metrics and diagnostics need to identify strengths and weaknesses and allow the developers to determine where to look for improvements
- Once weaknesses are identified, select additional metrics to measure what we are trying to “fix” and “maintain”
- Suite of metrics should be complimented by subjective evaluations
- Possible way to define metrics: holistic categories such as large-scale flow, high-impact weather, tropical cyclones, etc ....
- Coupling evaluation needs knowledge of climatologies through reforecasts

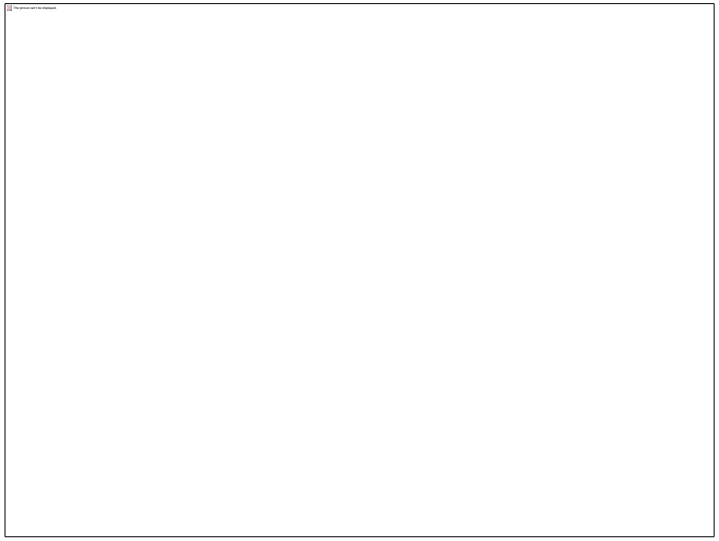
Community Involvement

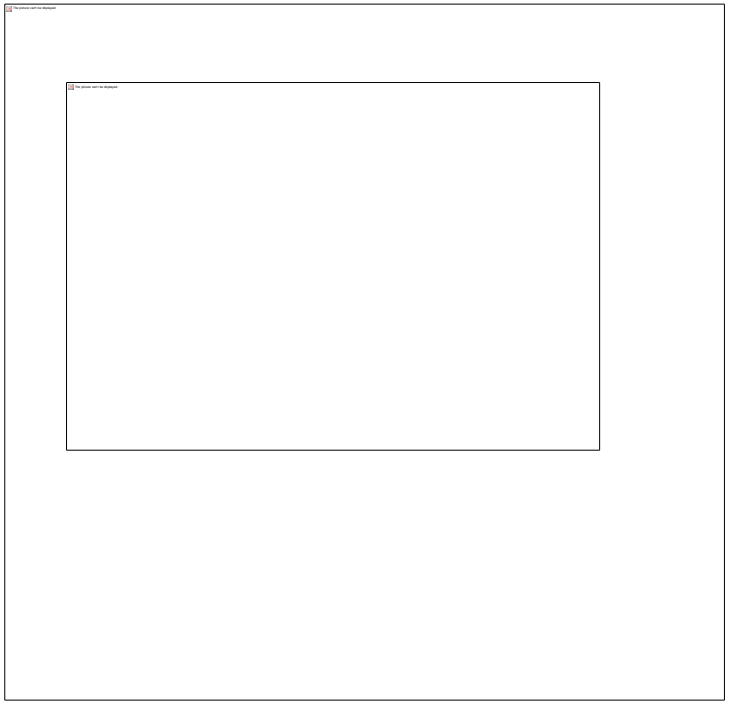
- Work with universities, private sector, other NOAA entities to get more “eyes on” operational vs parallels runs for evaluation
- Publish test plans well in advance so community can identify areas of evaluation not covered by EMC where they can contribute
- Need methods of synthesizing metrics and scorecards
- Need more focus on observation data sources for independent evaluation

V&V WG
Team Coordination and Dependencies

- **Coordination/dependency successes or issues;**
  - Active collaboration: CAM, ACC, System Architecture, Workflow Focus Group
- **What to accelerate**
  - Version 1 test plans and HTF defined to finalize metrics so METplus team can focus on that prioritized development
  - Transition of EMC and GMTB to METplus so METplus configurations can be published for UFS Community use
  - Developing evaluation capability for a fully coupled system
  - Community involvement in evaluation activities to help EMC identify areas of possible improvement
- **Change to working group**
  - Need more “non-NOAA community members” with an emphasis on university and private sector and more diversity in Co-Chairs
  - Group functioning well but need more direction for future

O2R2O2R Reality “TV”





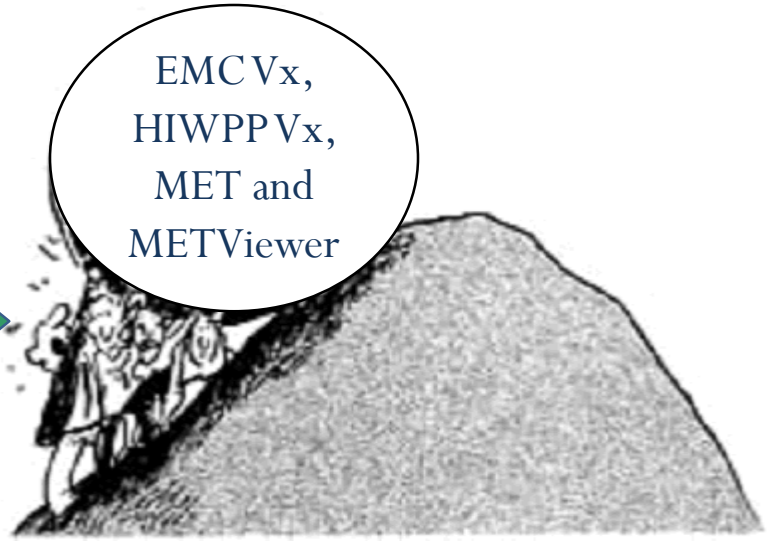
O2R2O2R - Before 2015

- **2005-2006:** DTC partners suggested a project to develop community verification code to replicate EMC mesoscale verification capability and make it platform independent and extensible
- **2006-2007:** Requirements gathering and initial development
- **2008**: MET 1.0 Release
- **2009-2013:** DTC tasks start adopting MET package for T&E activities; Testbed collaborations with HWT and HMT; METviewer developed to replicate EMC mesoscale database interface capability; Limited use at EMC of MET because they had the VSDB package and institutional inertia prevailed
- **2014:** NOAA withdrew support of MET in DTC budget because there was "no business case" to continue support

At 2015 DTC Science Adv. Board

MET was being used by international research community and within DTC but not at most of the organizations that paid for its development

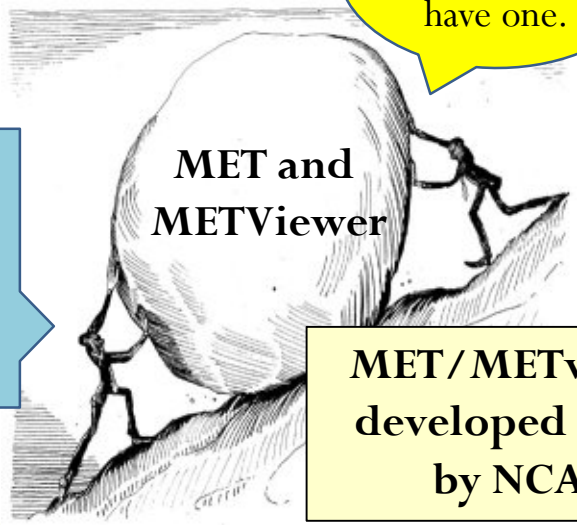
Unification
Maybe if we work together we can save time and money



2015

2012

No Thanks! We already have one.



Consider using this great verification package we've developed

MET/METviewer developed solely by NCAR

Recent developments at EMC, plus NCEP and UMAC push for unification of all systems has led to a discussion surrounding a unified NGGPS system. MET / METViewer is proposed to be the foundation

Challenges & Risks

Unification
Here we go…

MET

GSD and NCAR partnered together to identify needs but most development still at NCAR

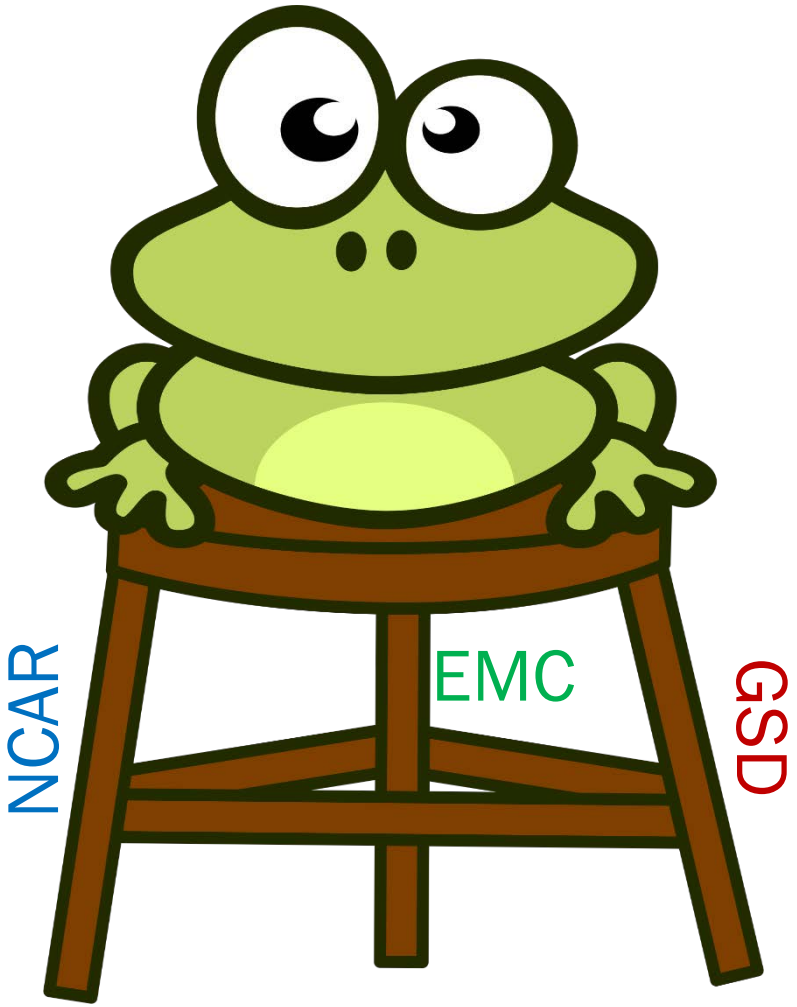
Managed Growth and Proper Funding

- Database – big data problem
- Simplified Interface for some users
- Community contributions & oversight
- Training and user support

Unification Roadmap

- Interviewed 50+ NCEP staff (EMC, WPC, CPC, NCO)
- Included discussions with coupled system "components"
- 99 functional requirements and 19 non-functional broken down by priorities
  - Statistics
  - Plot types
  - Data types
  - Preprocessing
  - Database and display
  - Documentation and help desk

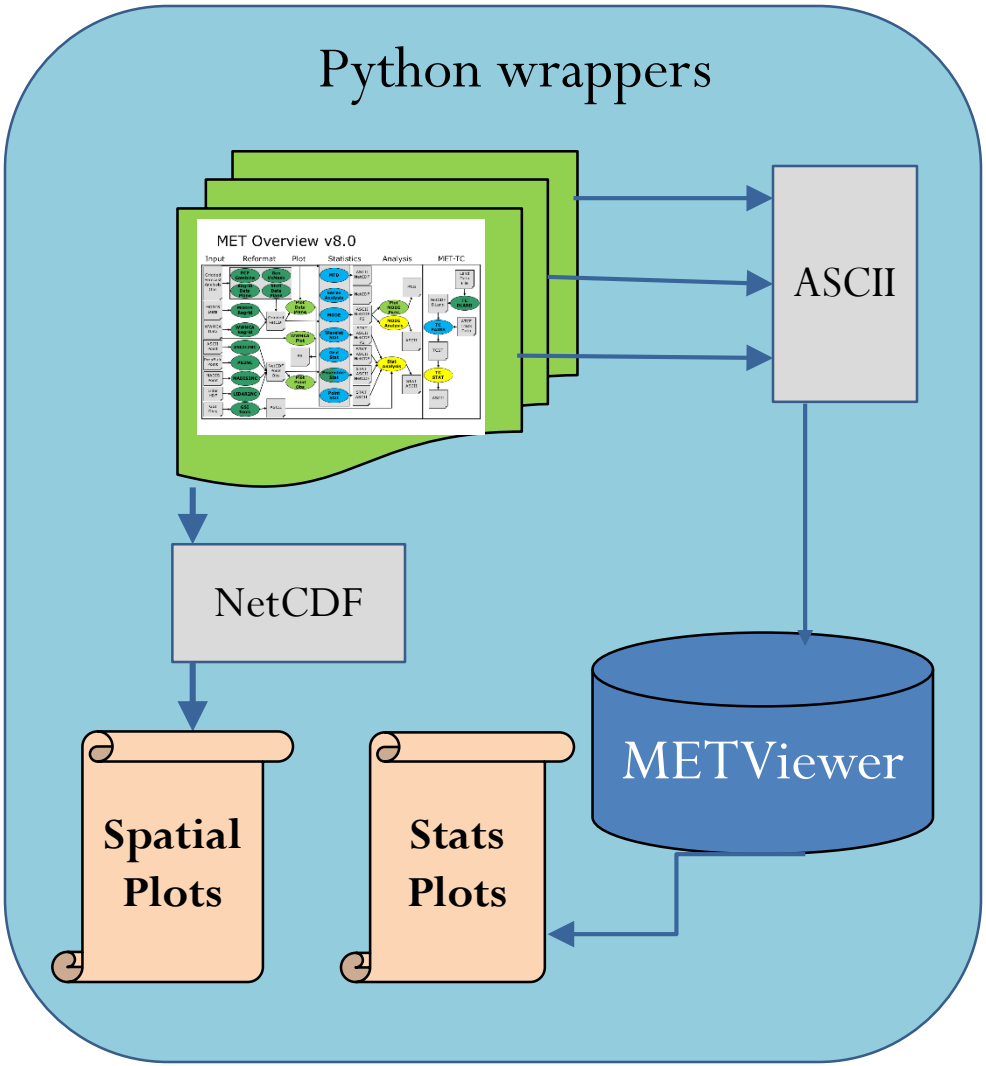
EMC starts collaborating with NCAR and GSD



NGGPS Unification Using METplus

Python wrappers around MET and METViewer:

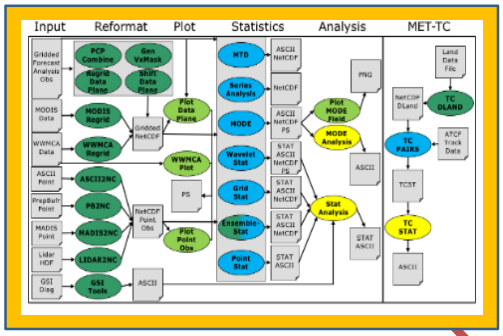
- Simplify set-up and running of MET
- Open up MET's C++ interface to work with other algorithms
- Automated plotting of fields
- Load data into METviewer database and display system and generate plots of scores through the batch engine



Mostly NCAR developers + 1 GSD developer

METplus Umbrella Repository

In Development



MET

METviewer Deep Analysis

METexpress Predefined Queries

METdb (database)

NCAR, GSD and EMC all contributing to METplus and METdb applications

Other UFS Tools

JEDI

UPP

Community Workflow

METplus

Python scripts for low-level workflow

Python scripts for plotting

manage_externals package developed by NCAR CESM team

Connection to Workflow Managers

Other Vx Pkgs

Air Quality

Marine

Conv. Allow. Model

S2S

Seasonal

Hydro

Land

Around 2018 DTC SAB and after

- **Aug 2018:** METplus 2.0 with MET 8.0 and METviewer 2.7 released
- **Oct 2018:** EMC and DTC (NCAR node) completed NCO Request for Change documentation for METplus to be installed in Operations
- **Dec 2018:**
  - NCO accepts METplus RFC
  - EMC VPPPG acquires cloud space on AWS to replace stagnated METviewer installation in the IDP (initially installed in IDP in 2015 but storage was constrained with no hope of expansion)
- **Jan 2019: Government Stutdown** but NCAR and GSD contractors kept working
- **Mar 2019:** METviewer and METexpress available on AWS
- **May 2019:** METplus 2.1 with MET 8.1 and METviewer 2.10 released
- **As of this May 15, 2019**
  - METplus 2.0 and MET 8.0 still not installed in operations but promised around May 24th. There is a chance that the operational install will be METplus 2.1 with MET 8.1
  - METplus team includes 25 core members but only 4 are fully funded by METplus work spread across 20+ different projects with median size of ~$150K

Current R2O Funding Paradigm
Cost to Organizations

- NCAR/RAL: Past year over 220 hours spent on writing or contributing to proposals or letters of support (just my time)
  - Result: Hopefully maintain NCAR METplus team's current level of funding to support our momentum + possibly being able to acquire 1 FTE
  - Reporting and Meetings: 10-20% **each** month are dedicated to reporting; project meetings take up at least 50% of my time and few of the other engineers

- EMC, WPC, SPC, NSSL – maxed out on reviewing proposals, tech transfer documents, and SOWs (personal communication)
- Non-NOAA community – many still don't really understand how to engage and what TRL/RLs are (personal communication)

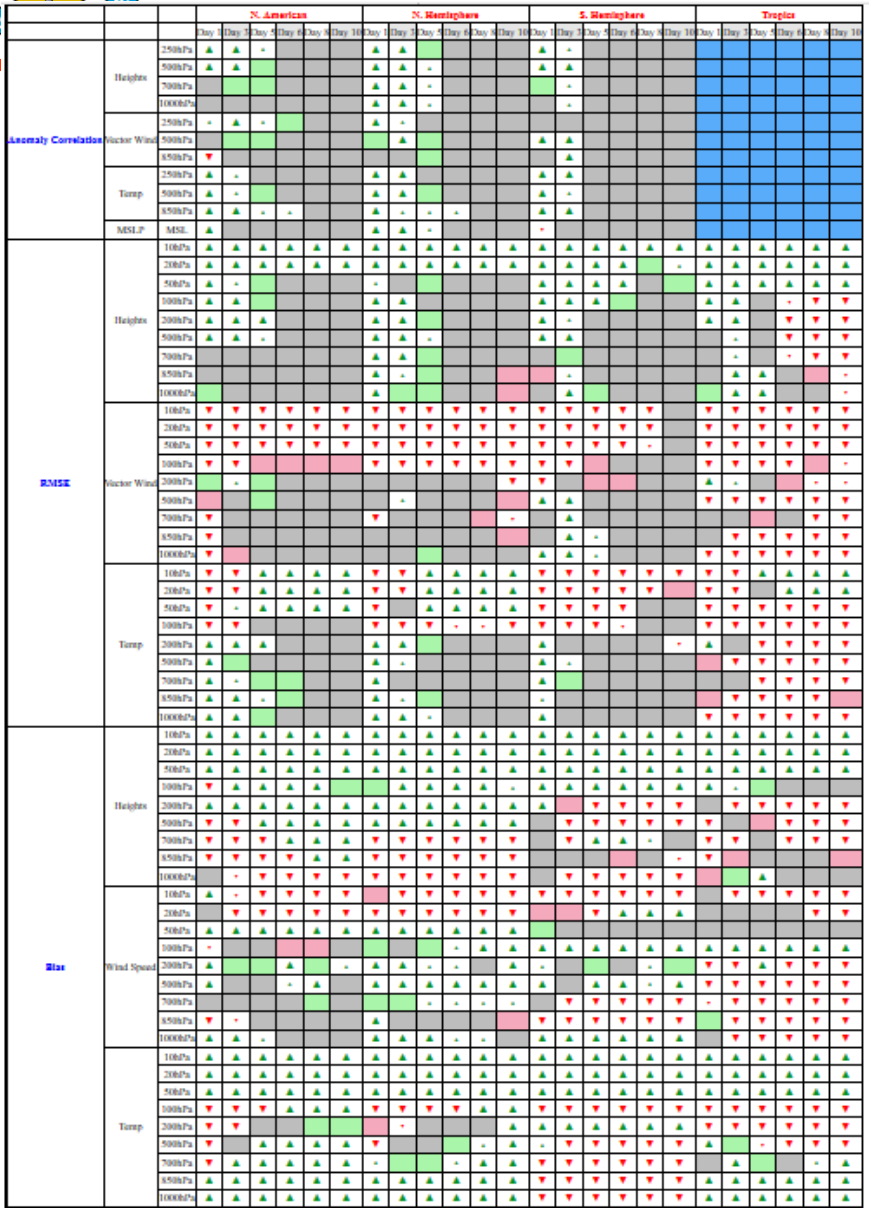
Metrics Refresher

Scorecards and the Metrics we choose

Something to think about:

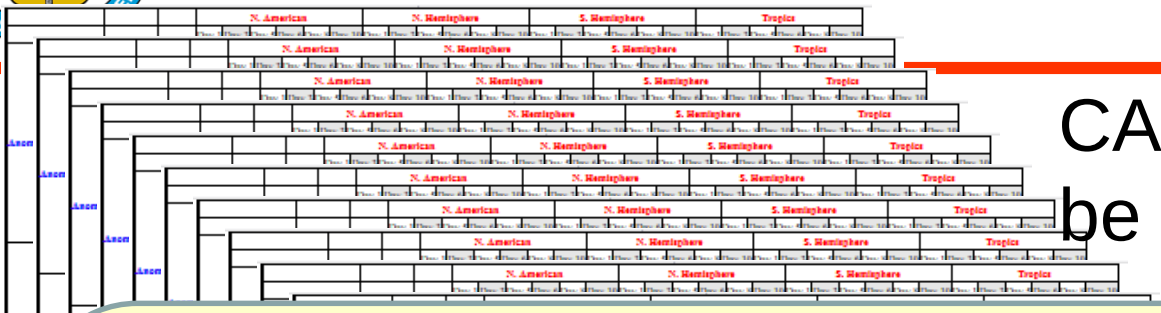
As initially proposed the CAM Metrics with permutations of all of the fields, statistics, thresholds, regions, and neighborhoods would represent just under 2000 rows

Scorecards and the Metrics we choose



Current EMC Scorecard – 65 rows

Scorecards and the Metrics we choose

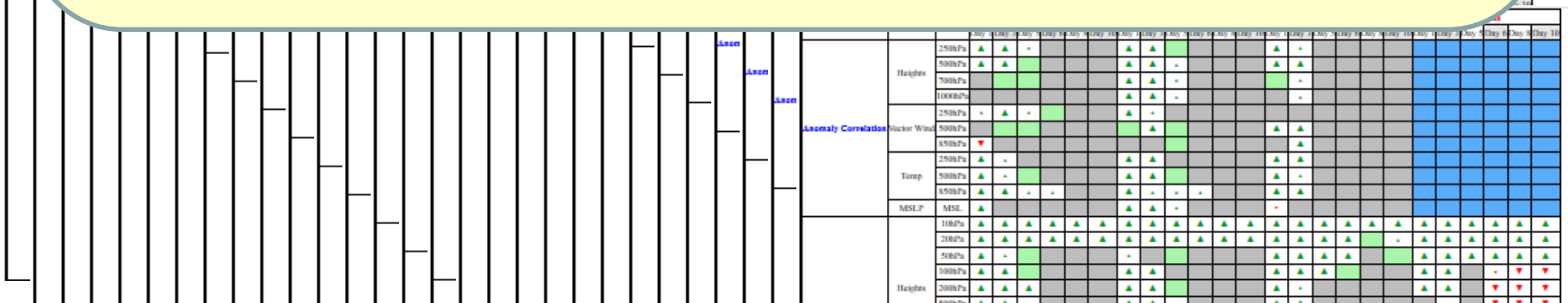


CAM scorecard would be 30 times larger…

We need priorities

1) Important for pre-implementation
2) What should be tested further up the funnel
3) Can sample size vary depending on where tests are in the funnel?

Hence the discussion of Hierarchical Testing



Metrics

Categories: Important to All, Global, Regional/CAM, S2S, Process Oriented, Coupling Validation, **Marine, Tropical, ACC, Land/Hydro,** Upper Air and Space

- Ensembles/Probabilistic measures were embedded in each sheet
- Example Metrics Worksheet - Atmospheric Chemistry and Composition (ACC)

Forecast Field	Specialty	Vertical Attribute	emporal Attribut	Validation Source	Priority	Maturity	Deterministic Methodology	Determinisit
				DRAFT IN PROGRESS				
e.g. Aerosol Optical D	Compositon	Surface	1-hr, 6-hr, 24-hr T	Stage IV 1-hr, 6-hr, 24-hr Precip			Grid-to-Grid	CSI, BIAS, FSS, POD, FAR, AUR, Perform
Aerosol optical depth		total column	1hr, 24hr, monthl	AERONET L1.5, L2.0	1	1	Grid-to-observations	*CSI, BIAS, FSS, POD, FAR, AUR, Perform*
Aerosol optical depth		total column	daily	MODIS, VIIRS	1	1	Grid-to-observations, C	*CSI, BIAS, FSS, POD, FAR, AUR, Perform*
Ozone		surface	1-hr, 8-hr, daily m	EPA AIRNOW, AQS	1	1	Grid-to-observations	*CSI, BIAS, FSS, POD, FAR, AUR, Perform*
PM2.5		surface	1-hr, 24-hr, daily max,ave	EPA AIRNOW, AQS,	1	1	Grid-to-observations	*CSI, BIAS, FSS, POD, FAR, AUR, Perform*
PBL Depth	Environmental/Air Qu	Top of PBL	Instantaneous	WSR-88D, METAR Ceilometer, RAOB, ACARS, BL Profilers	1	2	Grid-to-Grid, Grid-to-Point	RMSE, BIAS, Corr
Downward Shortwave Radiation	Air Quality/Energy	Surface	Instantaneous/Av	ARM, Surfrad (Oak ridge ameriflux)	1	1	Grid-to-Point	RMSE,
Downward UV Radiation	Air Quality/Energy	Surface	Instantaneous/Av	ARM, Surfrad (Oak ridge ameriflux), AIrNow UV r	1	1	Grid-to-Point	RMSE,
Aerosol optical depth		total column	15 min	GOES, Himawari, Meteosat, GEMS	2	2		
Aerosol index			daily	OMPS, OMI, TROPOMI	2	2		
smoke, ash plume height			daily	MISR, CALIPSO, MPLNET	1	1		
			1-hr, 24-hr,					

“Important to Most” Tab

MODIFIED Sheet Gathered During V&V WG after 2018 DTC Community UFS Test Plan and Metri

Forecast Field	Original Application	Vertical Attribute	Temporal Attribute	Validation Source	In METplus
Surface					
Sea Level Pressure	Global	Surface	24-hr instantaneous		Y
Precipitation	Global	Surface	3-hr, 6-hr, 24-hr Totals		Y
Precipitation	Global	Surface	3-hr, 6-hr, 24-hr Totals		Y
Dewpoint	CAM - Environmental	2-m	Instantaneous	METARs, SFC	Y
Temperature	CAM - Environmental	2-m	Instantaneous	METARs, SFC	Y
Wind	CAM - Environmental	10-m	Instantaneous	METARs, SFC	Y
Wind Gust	CAM - Environmental	10-m	Instantaneous	METARs, SFC	Y
SST	Marine	Sea Surface Tempera	*Instantaneous*	RTOFS, OISST	***maybe***
SSS	Marine	Sea Surface Salinity	*Instantaneous*	WOA09	*maybe*
SSU/SSV	Marine	Sea Surface Currents	*Instantaneous*	OSCAR	*doubt it*
Soil Moisture	Land Surface/Hydro	Multi-layer profile	Daily	SCAN, CRN, State Mesor	*maybe*
Snow Covered Area	Land Surface/Hydro	Surface	Daily	MODIS/VIIRS, SNODAS	*maybe*
Radiation	Process Oriented	Shortwave Down		Surfrad	Y

“Important to Most” Tab

Other					
MLD	Marine	Mixed Layer Depth	Instantaneous	Argo MLD data (Holet (2017))	***maybe***
HC300	Marine	Ocean heat Content	Instantaneous	GODAS	*maybe*
Nino3.4	Marine	Nino 3.4 index	Instantaneous	RTOFS, OISST	*maybe*
PBL Depth	CAM;Atmos composi	L0	Instantaneous	Derived from Pbufr, Sondes	MET 8.1 if using PrepBUFR
Cloud Fraction/Coverage	Process Oriented	L0		CERES, Infrared geostationary, MODIS and CloudSat	MET 8.1 if using GOES-16
Cloud Optical Depth	Process Oriented	L0		CERES, Infrared geostationary, MODIS and CloudSat	MET 8.1 if using GOES-16
Cloud base and top	Process Oriented	L0		METARs, SFC	yes
Lapse Rates	CAM; Process Oriented	between mandatory levels		soundings, analyses	may need additional pre-proc

Suggested by Workshop

Suggested by V&V WG

“Important to Most” Tab

Upper Air					
Geopotential Heigh	Global	Mandatory Levels +	24-hr instantaneous	Soundings, UPA	Y
Temperature	Global	Mandatory Levels +	24-hr instantaneous	Soundings, UPA	Y
Wind Speed	Global	Mandatory Levels +	24-hr instantaneous	Soundings, UPA	Y
U & V Wind	Global	Mandatory Levels +	24-hr instantaneous	Soundings, UPA	Y
Vector Wind	Global	Mandatory Levels +	24-hr instantaneous	Soundings, UPA	Y
Specific Humidity	Global	Mandatory Levels +	24-hr instantaneous	Soundings, UPA	Y
Jet Stream Position	Process Oriented	250mb		Analysis	Y
Highs and Lows	Suggested by V&V WG - would this be meaningful as well??				
Mandatory Levels + means much finer resolution requested (i.e. every 25-50mb)					

Suggested by Workshop

Suggested by V&V WG

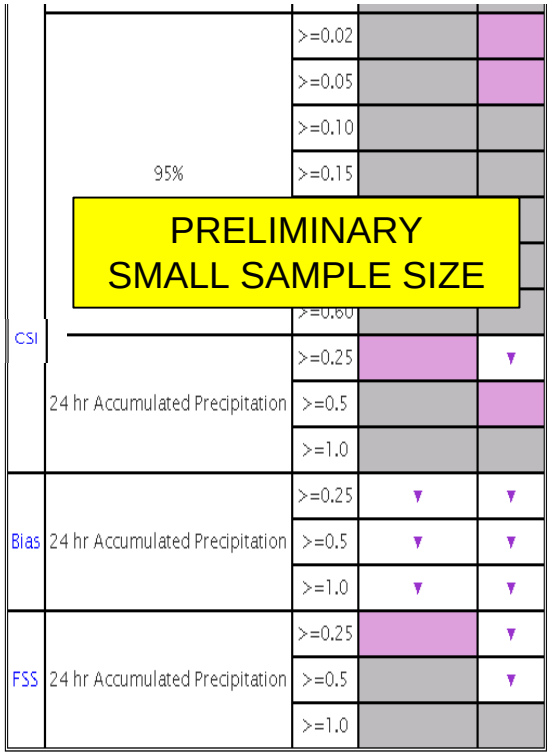
What else is missing?

CAM Metrics

https://hwt.nssl.noaa.gov/sfe_viewer/2019/verification/scorecards.php

Forecast Field	Application	Vertical Attribute	Temporal Attribute	Validation
Downward Shortwave Radiation	Air Quality/Energy Land Surface	Surface	Instantaneous/Average	ARM, S… (Oak ridge, ameriflux), USCRN
Ceiling	Aviation	Column	Instantaneous	METARs
Echo Top Height	Aviation	Column	Instantaneous	MRMS Echo Top
Visibility	Aviation	Surface	Instantaneous	METARs
CAPE/CIN	Environmental	Mixed, Most-Unstable,	Instantaneous	RAOB
Dewpoint	Environmental	2-m	Instantaneous	METARs
Specific Humidity	Environmental	Column	Instantaneous	RAOB
SRH	Environmental	0-1, 0-3 km AGL	Instantaneous	RAOB
Temperature	Environmental	2-m	Instantaneous	METARs
Temperature	Environmental	Column	Instantaneous	RAOB
Wind	Environmental	10-m	Instantaneous	METARs
Wind	Environmental	Column	Instantaneous	RAOB
Wind Gust	Environmental	10-m	Instantaneous	METARs
Precipitation	Precip	Surface	1-hr, 6-hr, 24-hr Totals	Stage IV 1-hr, 6-hr, 24-hr Precip
Simulated Reflectivity	Severe	Composite	Instantaneous	MRMS Mosaic Composite
Updraft Helicity	Severe	2-5, 0-3 km AGL	Hourly Maximum	Storm Reports

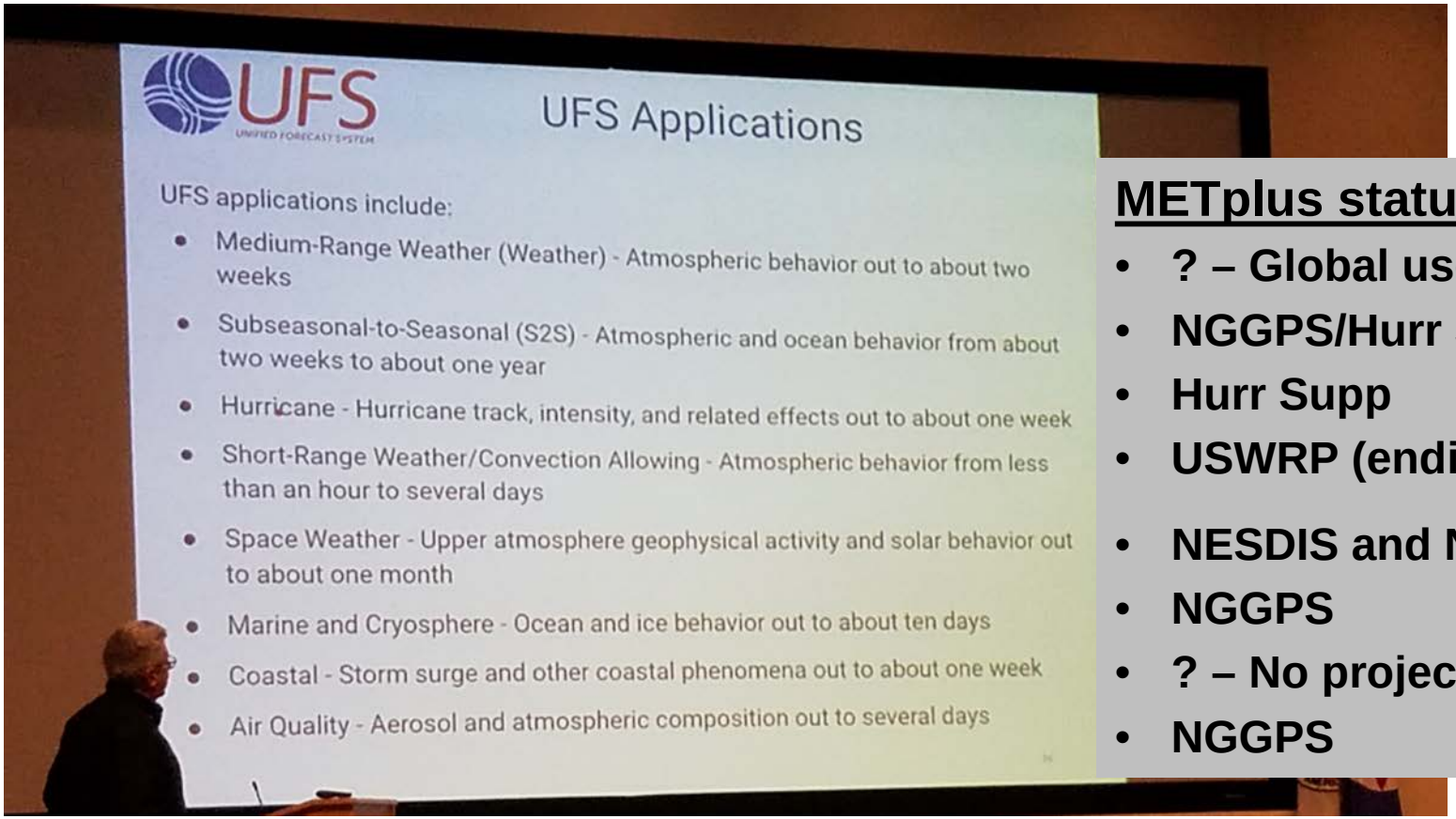
METViewer Surrogate Severe HREFv2/HRRRe CAM Scorecard



▲	HREFv2_cluegrid is better than HRRRe at the 99% significance level
	HREFv2_cluegrid is better than HRRRe at the 95% significance level
	No statistically significant difference between HREFv2_cluegrid and HRRRe
	HREFv2_cluegrid is worse than HRRRe at the 95% significance level
▼	HREFv2_cluegrid is worse than HRRRe at the 99% significance level
	Not statistically relevant

Rest of Metrics Worksheet

Google Doc – Anyone with Link Can View/Edit:
https://docs.google.com/spreadsheets/d/1NLX6Z_Ir6ue_aNMI4u30YrfmzfeifNQ2-Qhxpv2I_d4/edit#gid=1536770974



METplus status

- **? – Global use-cases?**
- **NGGPS/Hurr Supp**
- **Hurr Supp**
- **USWRP (ending); JTTI**
- **NESDIS and NASA**
- **NGGPS**
- **? – No project now**
- **NGGPS**

Discussion

MET v8.1:

- https://dtcenter.org/community-code/model-evaluation-tools-met/download (links to other github repos)
- https://github.com/NCAR/MET (MET v9.0 dev and beyond)

METplus v2.1: https://github.com/NCAR/METplus

METviewer 2.10:

- https://github.com/NCAR/METviewer

Help: met_help@ucar.edu