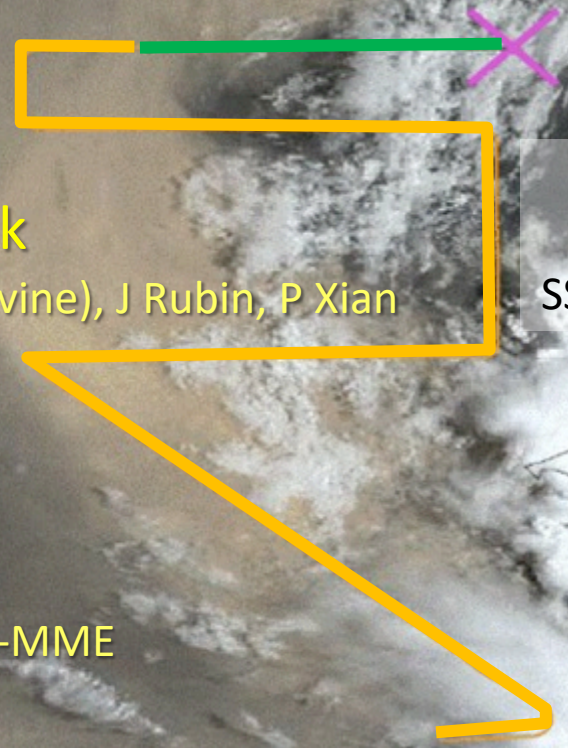
Update on ICAP product distribution: Field campaigns as examples of the need for interoperability

Jeffrey S. Reid, US Naval Research Laboratory
jeffrey.reid@nrlmry.navy.mil



NASA DC 8
CPEX-CV
22SEP2022
SSEC Geoworldview

Product Contributors to this Talk

NRL: L Frost (SAIC) JS Reid, J Roetman (Devine), J Rubin, P Xian
SSEC: A Gumber, RE Holz, W Marias
GSFC: PR Colarco, A da Silva, TF Eck
LARC: G Chen, A Nehrir
Copernicus: M Aides
JMA: T. Sekiyama
And all other dat contributors to the ICAP-MME

U.S. NAVAL RESEARCH LABORATORY

13th ICAP Working Group
Darmstadt, Germany
November, 2023

Todays Talk
Going about ICAP interoperability a little differently

Typically at ICAP meetings NRL gives an update on ICAP-MME systems developments. E.g., Peng Xian spoke earlier on a new monthly reanalysis product; Juli showing vertical verification.

Also at meetings NRL gets requests to expand ICAP-MME and display capabilities. We do our best, but it's a lot to do. Especially maintenance. Oh, the maintenance.

Consequently, we don't like to reinvent products, and if we can rely on an operational partner, so much the better.

But when does the rat race of data production, evaluation, posting, visualization, analysis, and storage end? Well, never really so long as there is development in the community. And the number of useful products keeps growing.

Proper data integration is hard, time consuming, and hence costly. Little things can make all of the difference in utility.

Today's talk we will go through operations to analysis in the context of recent field deployments to help address community wide desires of better interoperability. This afternoon, we can even discuss what interoperability even means.

But first: Current data products

ICAP Consensus:

- Generated at NRL with ~18 hr latency from 0Z
- 8 multi species, +1 dust only global model
- Have three consensus aggregations: “Core 4-C4” , all models (ICAP-MME), and dust.
- Products: Total, fine, coarse, dust, smoke AOD; $PM_{2.5}$ and PM_{10}
- Available on https://usgodae.org/pub/outgoing/nrl/ICAP-MME/ Free to use, just please acknowledge NRL and ICAP contributors

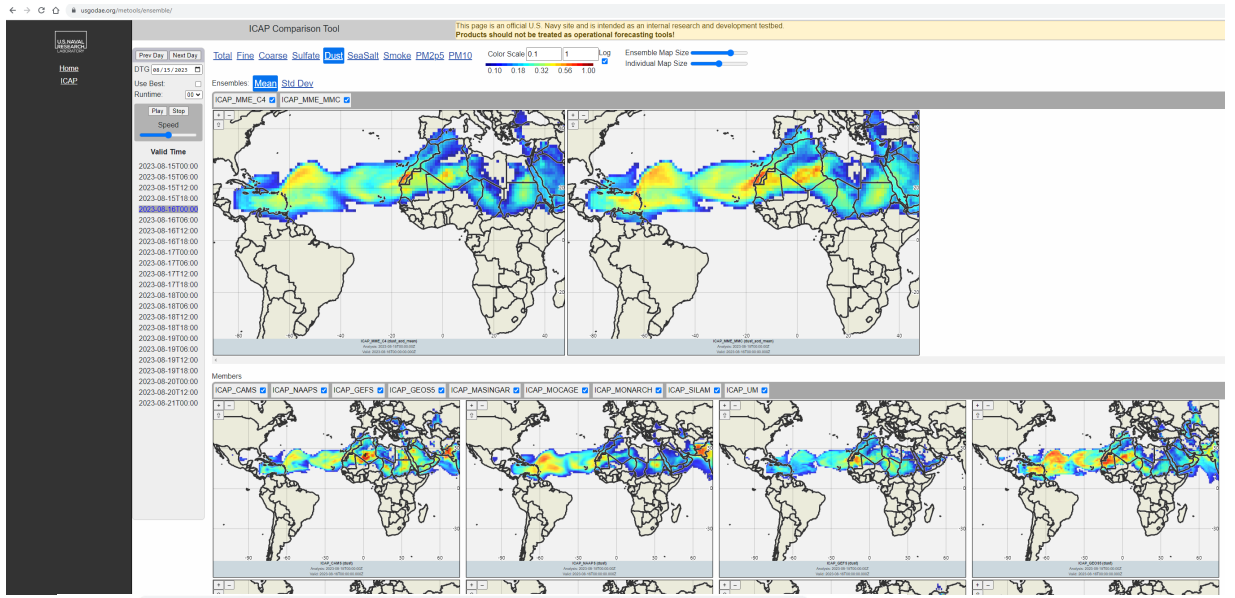
ICAP Assimilation & Verify:

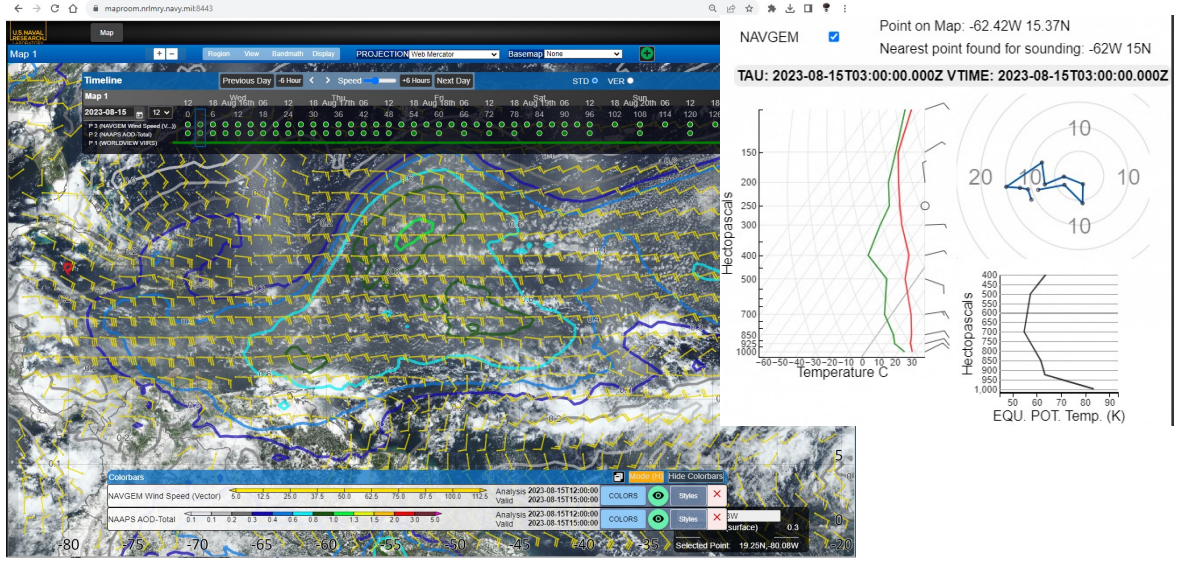
- Generated at University of Wisconsin SSEC w/12 hr latency and 2 day archive sweep.
- Combines into one 1x1 gridded 6 hr nc file: AOD: MODIS, SNPP VIIRS)NASA and NOAA); Surface: AERONET, visibility, AirNow Other: OMI UV products. Know if another good real time feed of L3?
- Also a 6 hr surface met csv from McIDAS
- Available for the last month on ftp://ftp.ssec.wisc.edu/nrl_aod_l3/ Feel free to use, just please acknowledge SSEC and product contributors

And: Current visualization

- ICAP Map Room Quickdraw
  - Our old visualization machinery was becoming difficult to maintain. So a new more flexible system was built. We still have IT access issues from time to time.
  - https://usgodae.org/metools/ensemble/

- Map Room Public Site (beta)
  - A more limited version of the Navy site now hosts a small set of Distribution A (public release) data. This includes the ICAP-MME product line and a subset of the SSEC L3 product, including AERONET.
  - In beta, so working out some kinks-especially in color bars.
  - Looking for a few good beta testers
  - https://maproom.nrlmry.navy.mil:8443/





OK, back to the story line: Why we do this.

The interoperability process for models, remote sensing, and field data alike.

- Operational and quasi-operational data centers are there to provide the most meaningful data guidance to some defined set of customers on a set schedule without interruption.
- Developers need to provide the best possible system and configuration, and ensure fixing one thing does not break another
- Assessments need to inform developers and centers how they are doing, and what can be fixed without breaking something else.
- Data needs to be ingested from multiple sources and quality assured.

"Data should be easily accessible, publicly available, reasonably well documented, and, for baseline quantities, encoded into a similar format. Currently data distribution is diffuse and potential users have difficulty maintaining and evaluating global-scale data outside of the largest and most consistent networks", Benedetti et al. 2018.

Help us help you!

Capital Field Missions: The ultimate soup pot

Each year hosts many field campaigns related to aerosol science, often combining airborne, ground, satellite remote sensing, and model assets.

Forecast models always play a significant role in mission execution, and their analyses provide crucial information for planning and assessment.

The Good: Field missions often make a diverse enough set of observations to largely close the system of study.

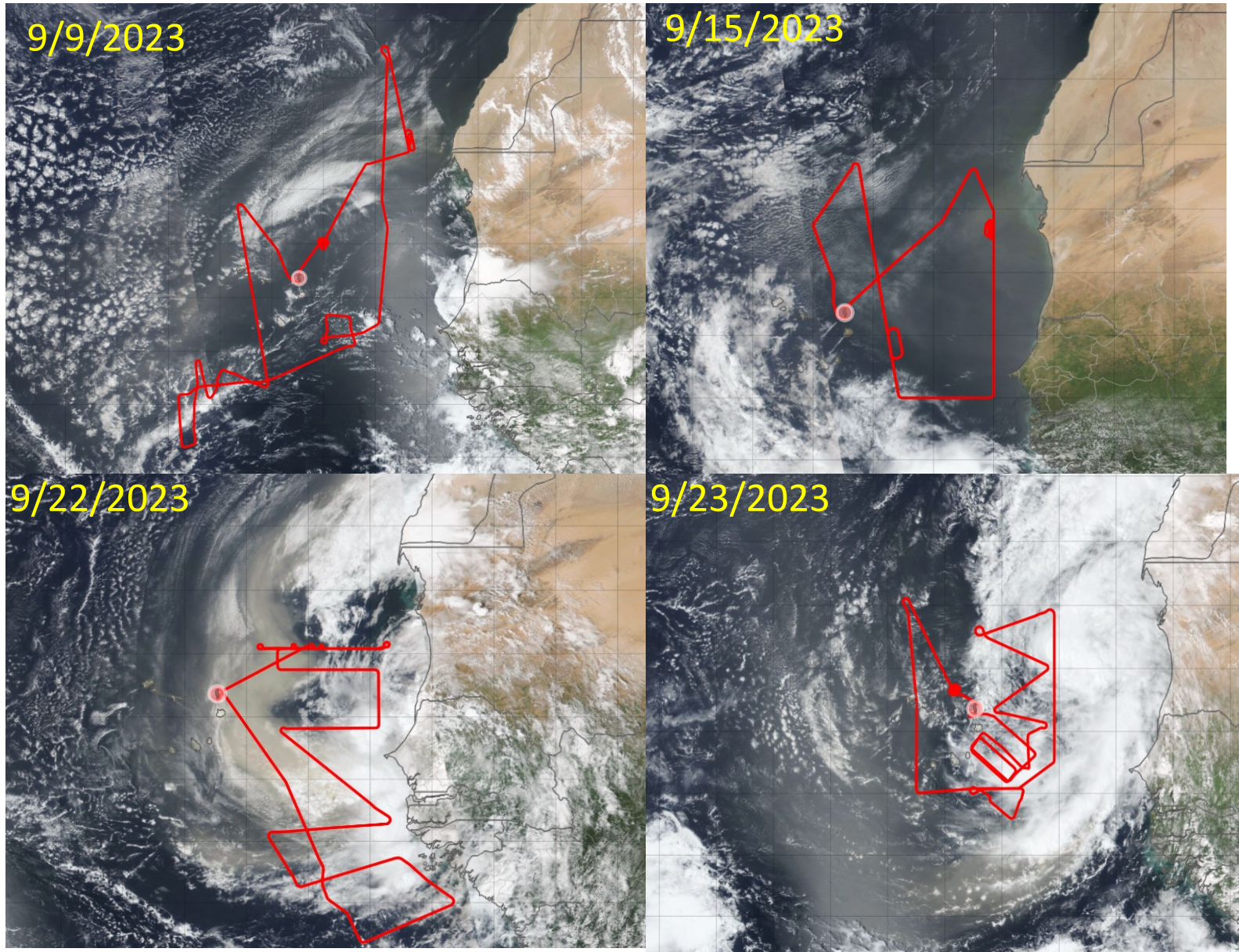
The Bad: There are never enough observations. And It's a lot of work to integrate in situ observations, sat obs, and models that we already have for just a snapshot as to what is going on. And then apply to lasting systems.

The Ugly: Data from different providers is just different enough to require lots of hand analysis. LOTS. Data is thus underutilized. When it is, often it misunderstood.

September 2022 CPEX-CV as an example

- CPEX-CV was a NASA mission operating out of Sal Island, Cabo Verde September 2022.
- 3 dust events were well sampled: Sept 9, 15, 22-23.
- Onboard was HALO aerosol/H_20; DAWN wind lidar, & HAMSR microwave remote sensors. Plus dropsondes and a CAPS probe.
- Small differences in ob location resulted in vastly different sampling to models.

Lesson 1, 9/23-you cant always get an observation where you want it. Thanks ATC strike…. Nevertheless, **think hard about observation placement and influence when you can. Especially long term sites** And how to hedge when you can't.



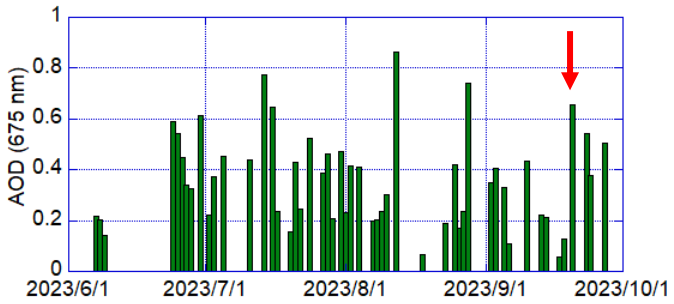
Start with AOD. Why? Because its easy and everyone has it. And Ang has her wish; but it all needs to be there

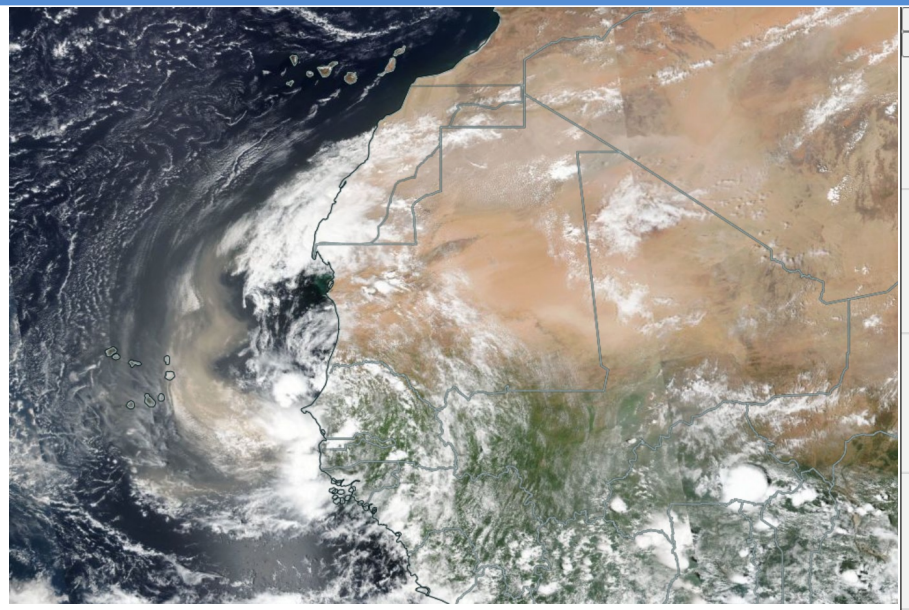
U.S. NAVAL RESEARCH LABORATORY

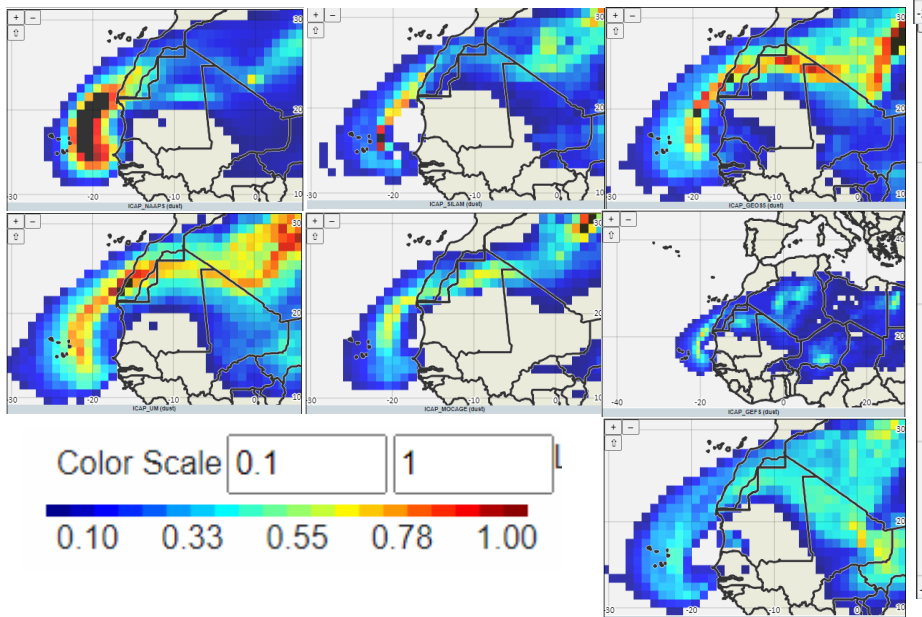
- Sept 22,2023 was our 1.5 sigma dust event as dust wrapped around a tropical wave soon to be a tropical storm

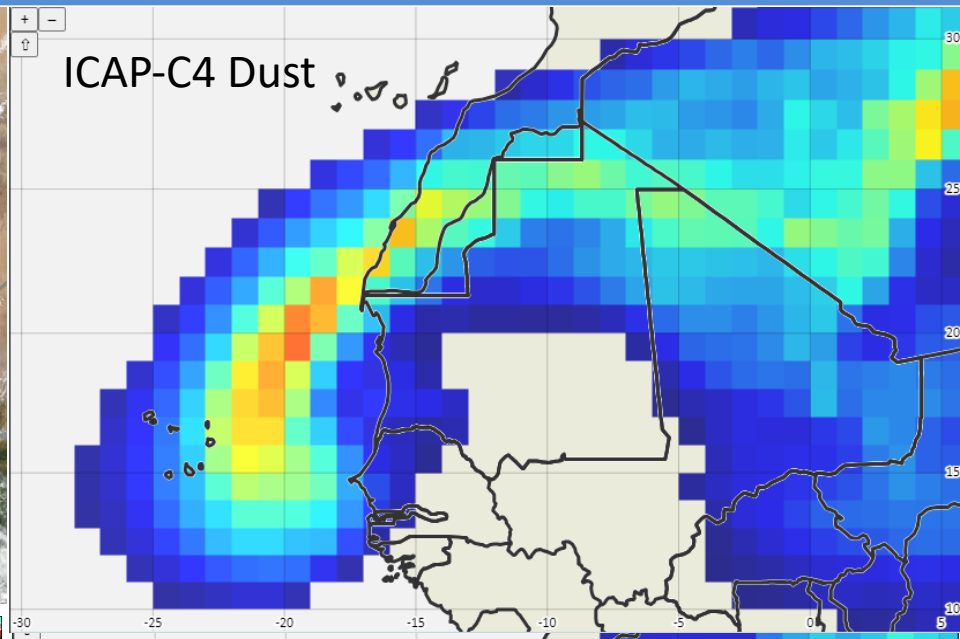
- Shown are 0Z runs, 12Z VT. 7 of 9 models reporting. All models showed a dust event that run. But all underestimated the magnitude and the spread was a factor of 3.

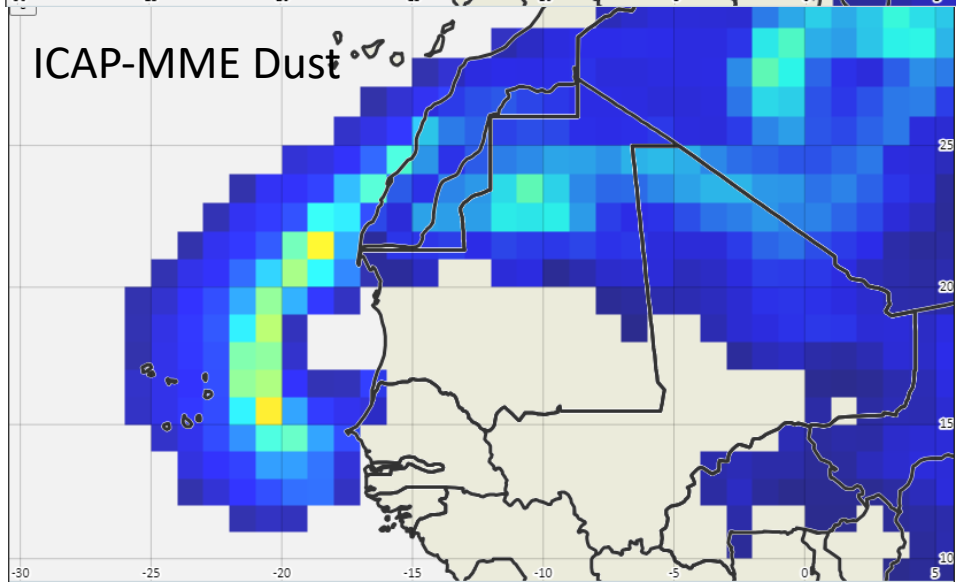
Lesson 2: Ang's dream fulfilled of all models on the same color bar, and then some thanks to the MME. Interoperability in data exchange is good. **Please conform to conventions and be verbose in netcdf files.**











How about the forecast?

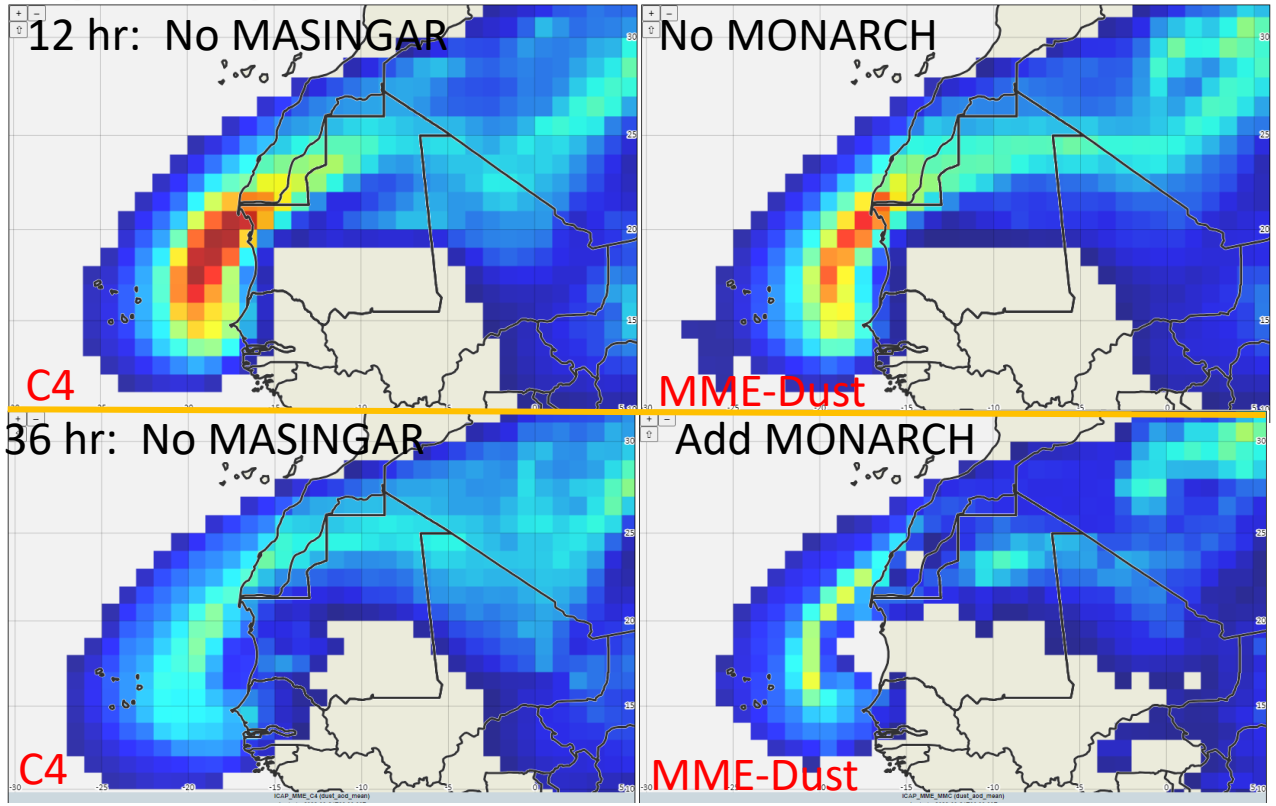
Color Scale 0.1 1

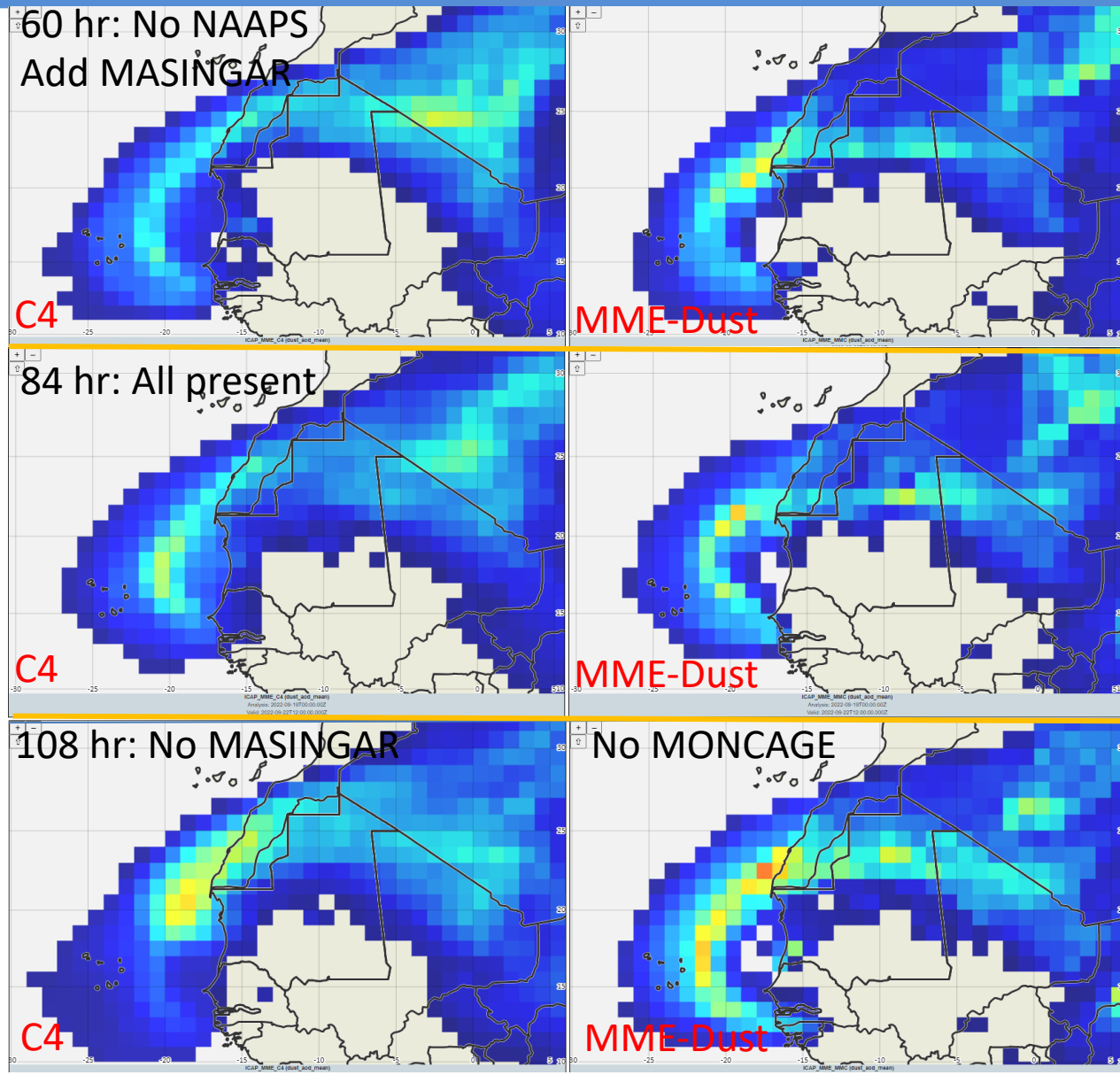
0.10 0.33 0.55 0.78 1.00

Small differences in which models posted for the day made a noticeable change in magnitudes of the consensus.

The ”best model” by tau changed frequently. Some did better with longer lead times.

Lesson 3: Consistency in product availability is good. Probably a lesson for met forecasters is in here too. We know IT systems have an ecosystem of their own, but please **take the time to help make systems robust.**



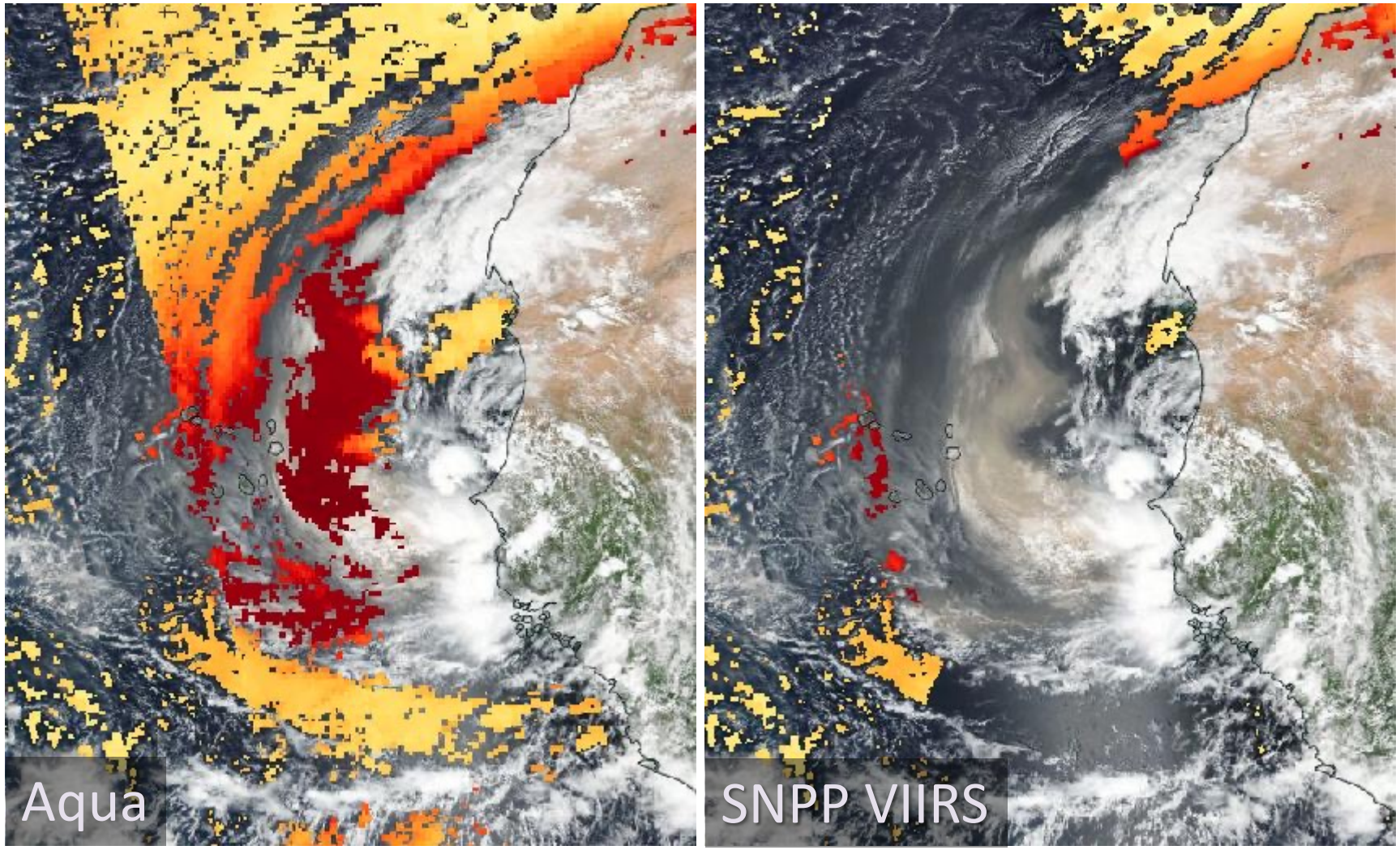


Capability and interoperability

Interoperability is not just about formats, but also coverage/capability.

But at least in this case things match pretty well for what overlap there is

Lesson 4: We need constellations to ensure continuous observations to support operational modeling. But this means more products to manage. And bias correct. Please **take the time to characterize products with long time series** else we will have a hot mess.



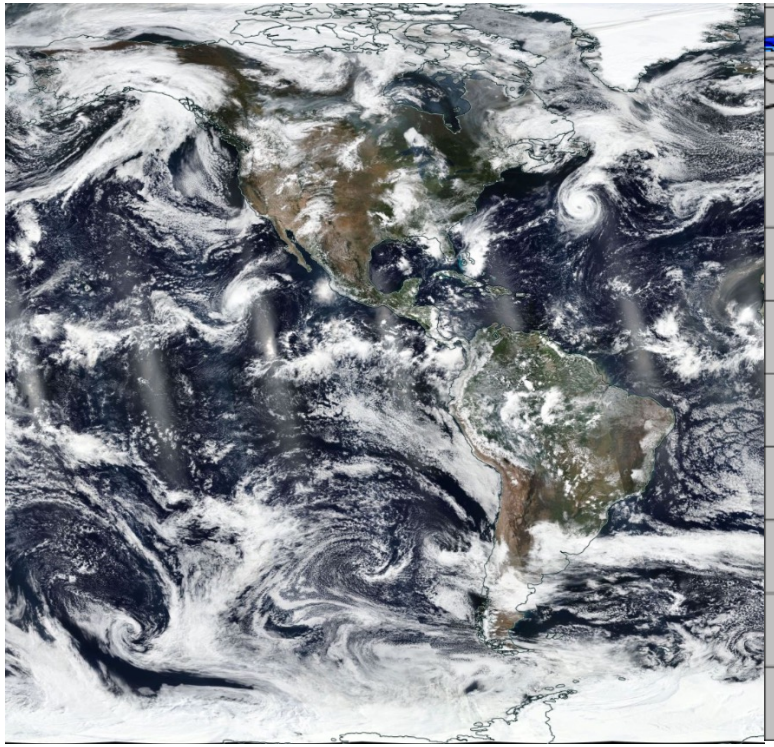
OK, so how do similar “verification” products compare?

MODIS has been a good ride and is now drifting, so bias corrections will likewise drift.

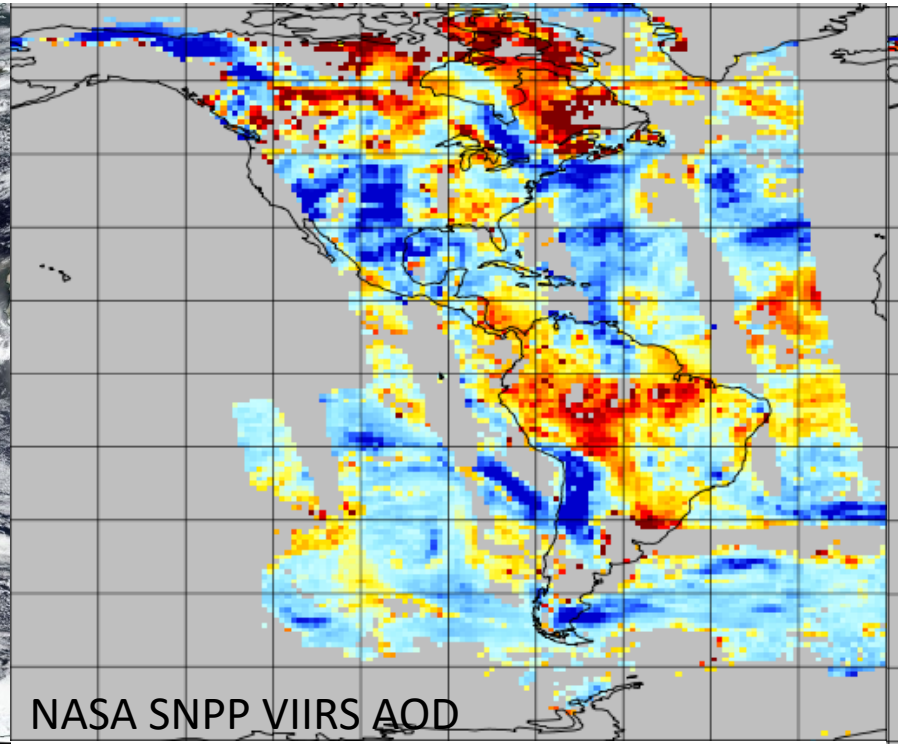
How is the VIIRS transition? NASA and NOAA products have significant differences for the same sensor.

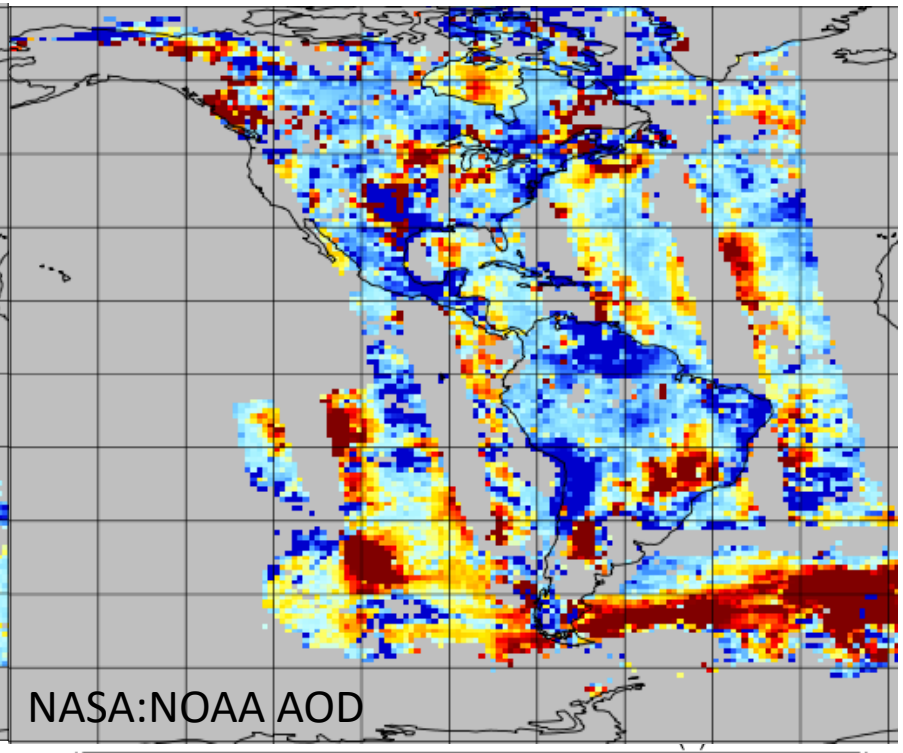
Obvious spatially correlated biases related to scan angle, glint, aerosol specie.

Lesson 5: Even using the same sensor, there are widely varying capabilities between products. And this is only going to get worse. **Determine what your niche is**. And put yourself in the position of the user.



Sep 20, 2023

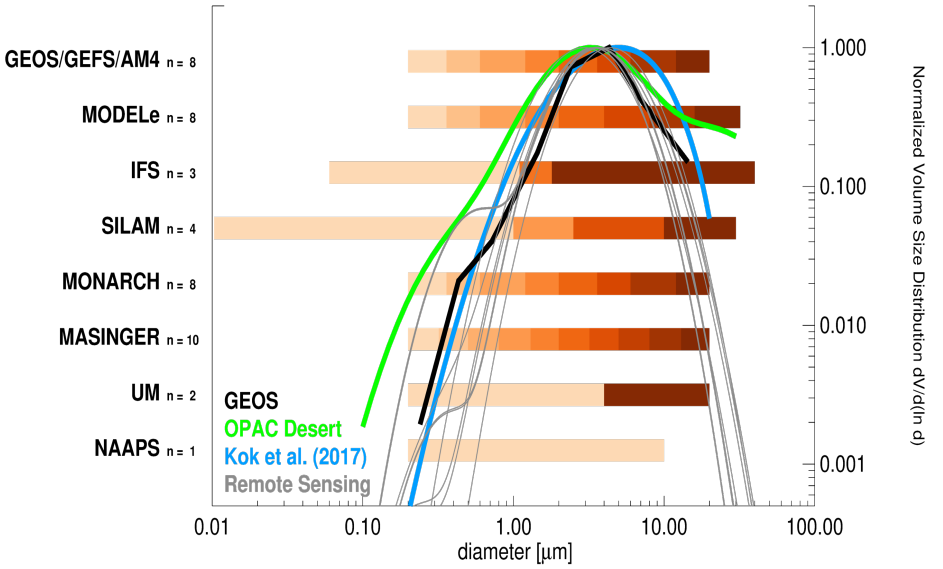


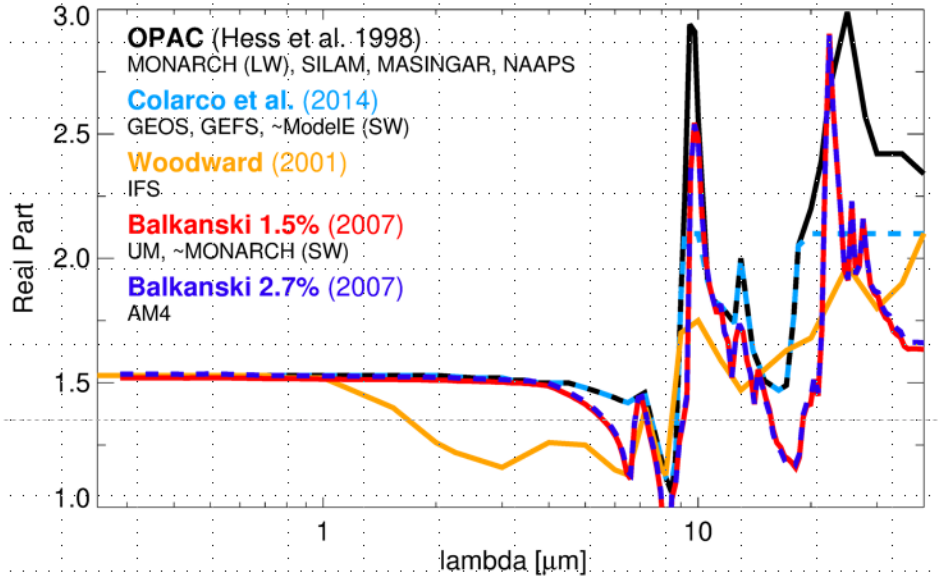


Great, we have vertical extinction! Now what? Lets review Peter Colarco's microphysics findings

- Choice of size section placement and resolution has large impact on AOD
  - AOD computed by the individual models can vary by 40% depending on the assumptions of the dust PSD
  - This is not a consequence of the assumption of the refractive index
  - Dust absorption variability due PSD is somewhat smaller (<30%), with intermodal variability driven by refractive choice
- Models with the finer resolution of the size distribution have greater sensitivity in computed optical properties to variation in the PSD
  - Implications for observational need to constrain fine mode portion of dust
  - Implications for estimates of radiative forcing from models
  - Implications for data assimilation -> possible error in AOD to mass translation

Lesson 6: Diversity is good, but don't be arbitrary. Decisions can be impactful, and ***it is hard to go back and change things in operational systems***. Please clearly document rationales and outcomes.

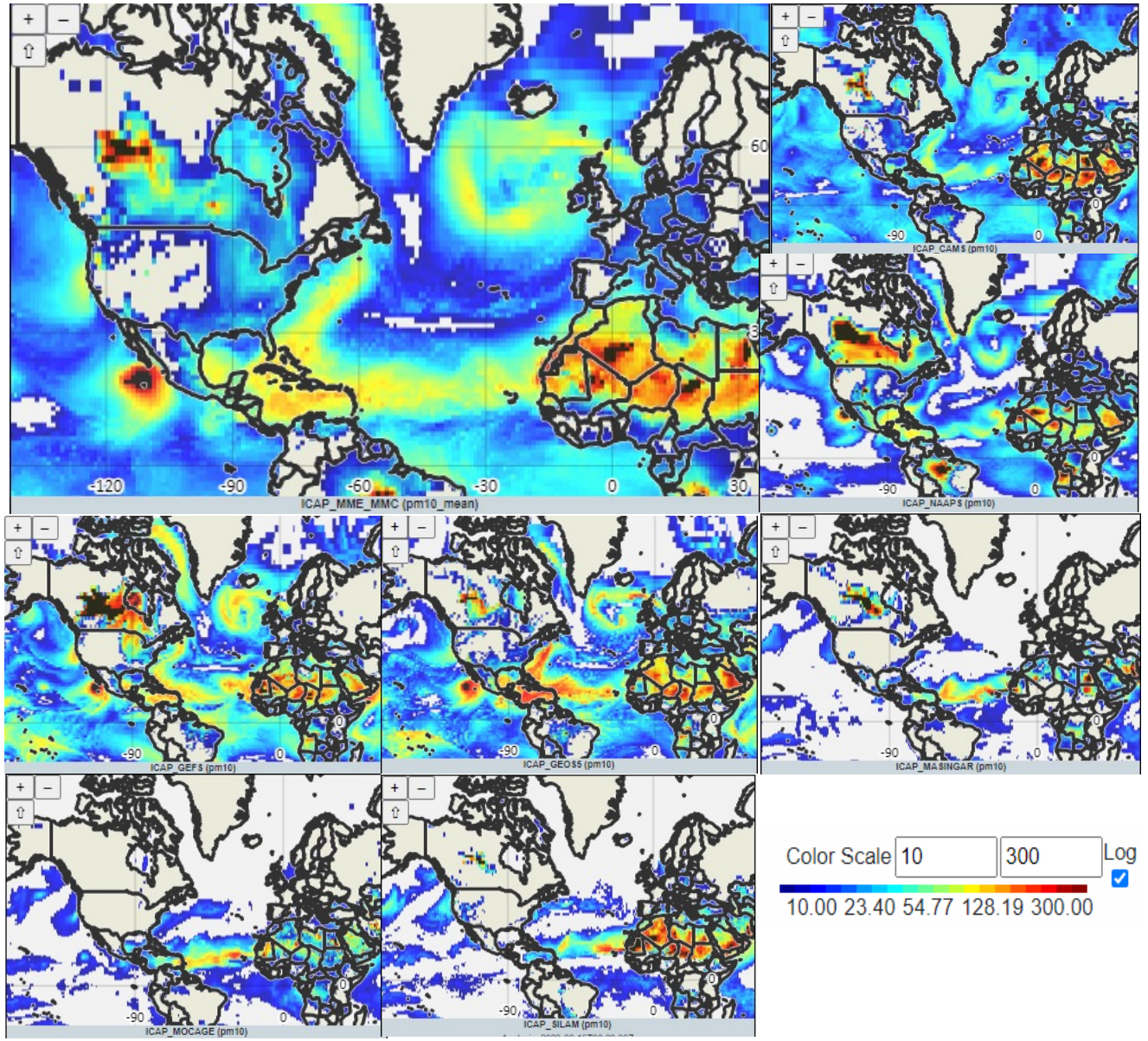




ICAP now has $PM_{2.5}$ and PM_{10}. Shouldn't this be easier right compared to vertical profiles of extinction? Aug 15, 2023. Dust in the Caribbean, smoke in Canada, and TCs. Oh My..

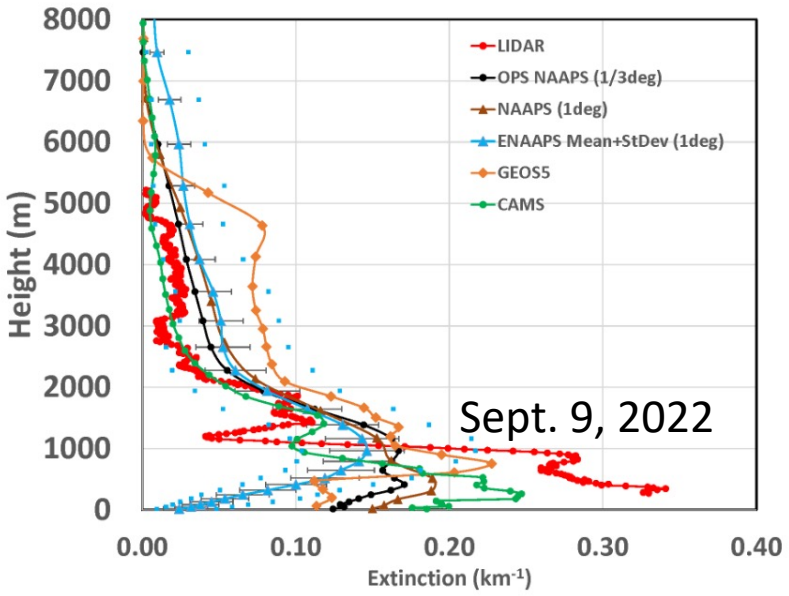
- So we start with Colarco's assessment on the sensitivity to radiation to mass.
- It only becomes more complex because:
  - Now we care about a single level so there is cancellation of errors in the vertical.
  - PBL-Free troposphere exchange is tricky.
  - Surface is where the source is. The lowest levels are also tricky numerically.
  - Lots of AOD and even lidar data over the glove. Systematic and QA'ed $PM_{2.5}$ and PM_{10} data is hard to come by outside of where it is clean. Oh, and they are often located next to point sources.
- Most models don't have $PM_{2.5}$ and PM_{10} products per say, just sum of species. And there is a little accounted for hygroscopicity term for the inlets.

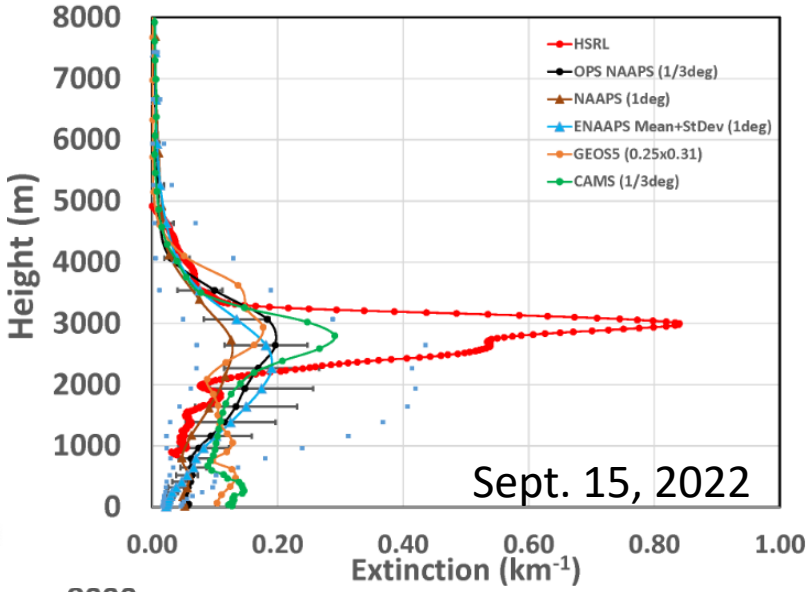
- Lesson 7: All of this means we need more data to understand the outcomes. ***And that is more data that needs to be organized***. See lesson s 1&2.

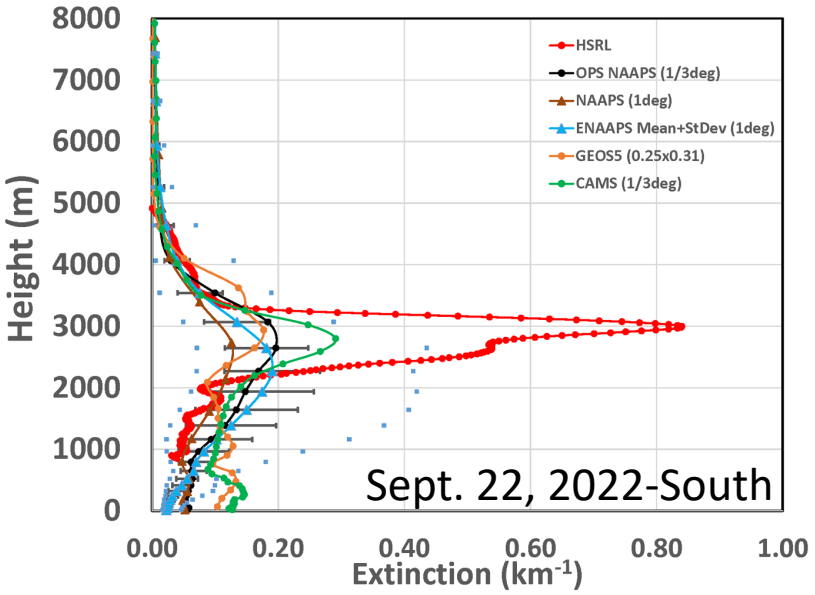


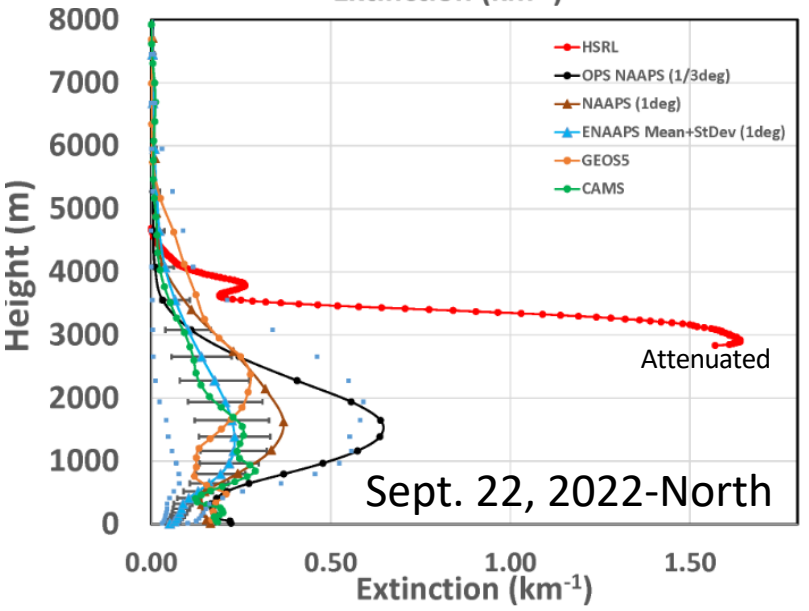
OK Back to models: Going "deeper" than AOD and the surface

- The AOD consensus and NASA satellite products are pretty straightforward but time consuming to track the different product production times and changes.
- Lets add "The Other" dimension, the vertical. Fortunately, the rise of quality lidar data in the air and on ground is making it much easier to score the vertical.
- What do we find? A lot more of what we see in AOD. Correlation, but spatially diffuse and big differences in magnitude.
- Indeed, each model can declare victory in some aspect for some event.
- But it is a slippery slope-with this we can ask lots more questions about intensive and extensive errors.
- Lesson 8: Going beyond AOD means accounting for orders of magnitude of more data to manage-and everything that goes with it on the obs and model side. **Lots more of lessons 1-6 AND 7.**







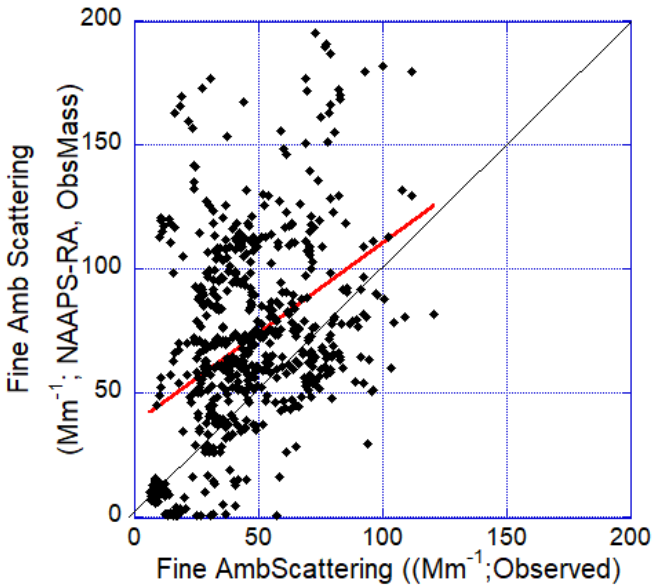


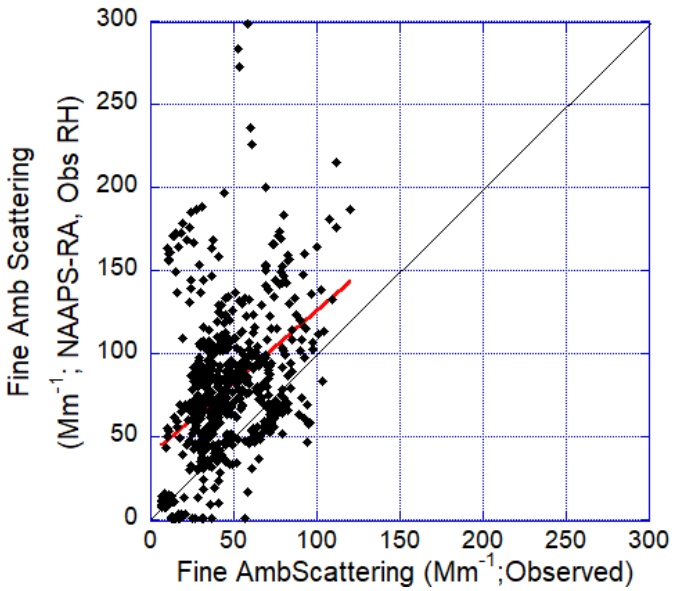
What does interoperability buy us: Vectorized error propagation for all. With mostly obvious results.

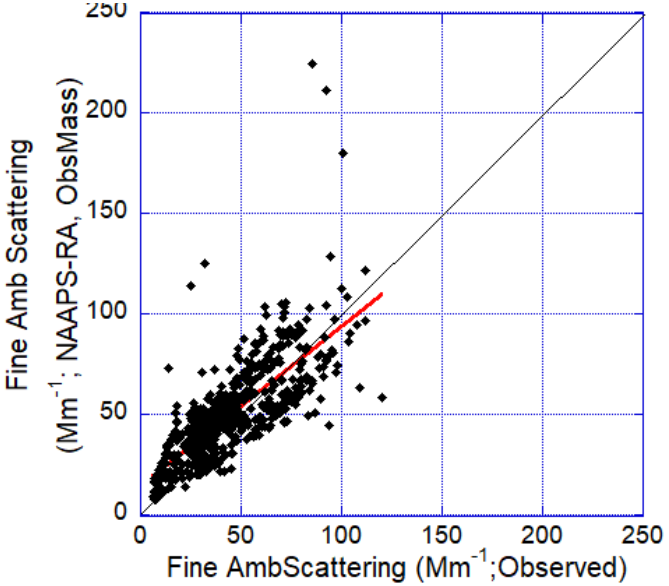
- Interoperability is not the same thing as harmonization. Diversity in products and ideas is good, as long as it move the field forward.
- We all want to know where the bodies are buried in the models so we can improve.
- This requires vectorized error models, where we can systematically swap out observed versus modeled variables.

Lesson 9: Obviously, errors are not local, so not so easy to bias correct. ***We need to work on interoperability of meteorology too.***

How much better would ambient extinction prediction be if we had perfect relative humidity? Not as much as you would think. SEAC4RS PBL light extinction. From Victoria Lang (AGU)







In closing: How do we get up the power curve? By making it easier I suppose.

- If the ICAP Consensus has taught us anything, everybody does everything just different enough to make interoperability a big last mile problem.
- This is easier (if not easy) doing with a few reanalysis products. In NRT, interoperability becomes time critical. And it is harder to backfill.
- Once you get the data into the same format and matrix, the rest is much easier (if not easy. Again).
- If we want to make sense of model diversity, we need to fully understand static versus dynamic dependencies (e.g., size and refractive index treatment versus RH and wind).
- From there we can put it all into a data store to mine it and eventually vectored error propagation.
- To get there, we really need an interchange nc format. Yes it is another product line, but it will pay you and the community dividends.