Development and Evaluation of North America Ensemble Wildfire Forecast:

Initial Application to the 2020 Western United States "Gigafire"

Peewara Makkaroon[1]

Co-authors: Daniel Tong[1], Yunyao Li[1], Youhua Tang[2], Barry Baker[2], Mark Cohen[3], Rick Saylor[3], Anton Darmenov[4], Alexei Lyapustin[4], Yujie Wang[4], Edward Hyer[5], Peng Xian[5], Ivanka Stajner[6]

[1]Air Quality Group of George Mason University, [2]NOAA Air Resources Laboratory (ARL)/GMU, [3]NOAA Air Resources Lab,[4]NASA Goddard Space Flight Center (GSFC), [5]Naval Research Laboratory (NRL), [6]NOAA National Weather Service.

Background

The 2020 Gigafire in the Western US

- During **August-September 2020 i**n Northern California
- **Burning more than 1-million acres** and destroyed 935 structures.

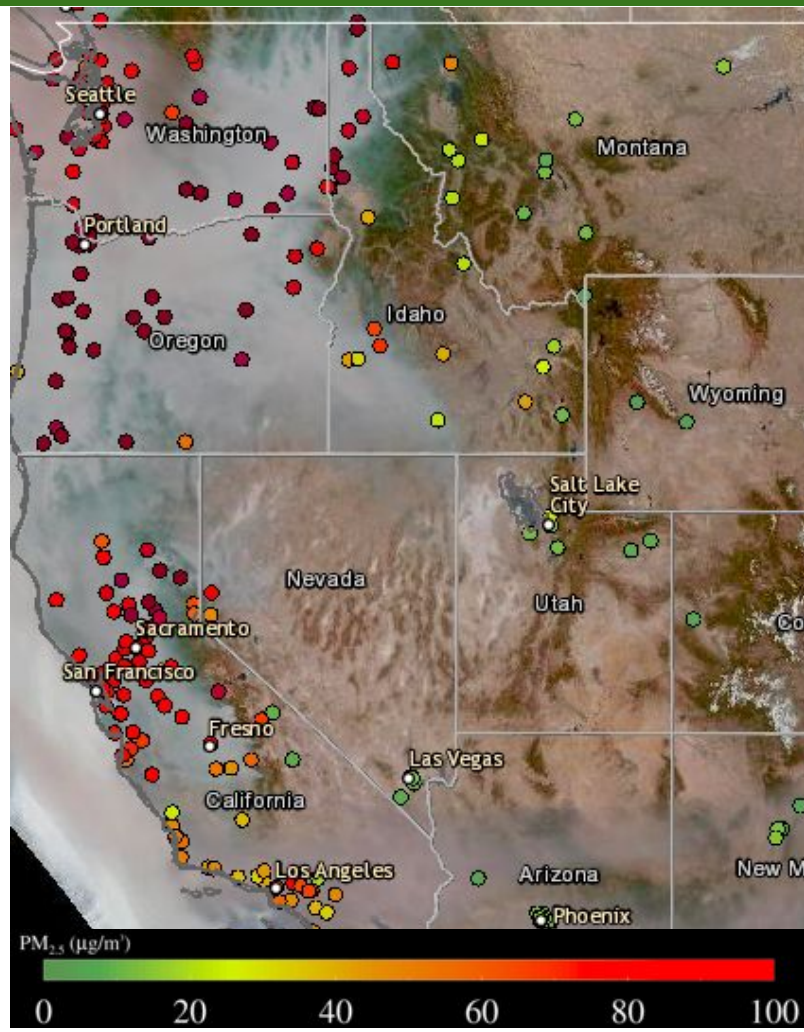
Ensemble Forecast

- Can reduce uncertainties in input emission and meteorological data, and simulation of each model by averaging them out.

Objective

- To improve forecasting performance of $PM_{2.5}$ and AOD during fire events across the US by developing a multi-model ensemble wildfire forecast using regional and global models over the Continental United States (CONUS) domain.

$PM_{2.5}$: Particulate Matter less than 2.5 µm in diameter
AOD : Aerosol Optical Depth



VIIRS true color and EPA AQS daily $PM_{2.5}$ observations (circles) on September 12, 2020 from NOAA AerosolWatch.

Multi-Model Ensemble Forecast

Ensemble Members

Regional models:

- GMU-CMAQ, NACC-CMAQ, HYSPLIT;

Global models:

- GEFS-Aerosols, GEOS, NAAPS;

Global Multi-Model Ensemble:

- ICAP-MME

Ensemble Creation

- Grid resolutions of all models were interpolated to 12km×12km before calculating their mean $PM_{2.5}$ concentrations and AOD values.
- **Ensemble mean** was calculated based on mean $PM_{2.5}$ and AOD of each ensemble member.

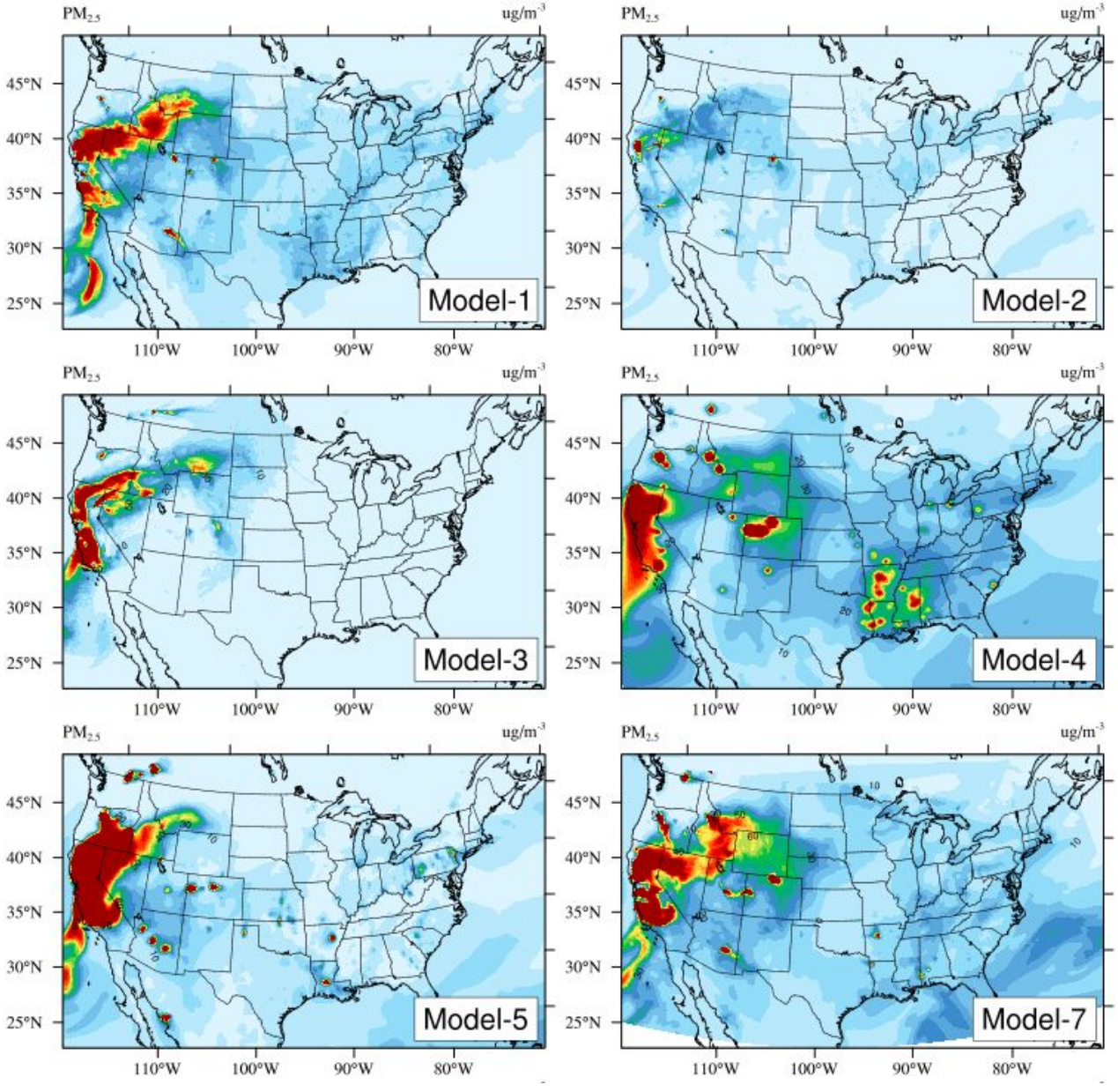
Ensemble Forecast

- Provides 24-hour $PM_{2.5}$ and AOD forecast on a 12km×12km grid resolution over the CONUS domain.
- Forecast time starts from 12 UTC - next day 11 UTC.

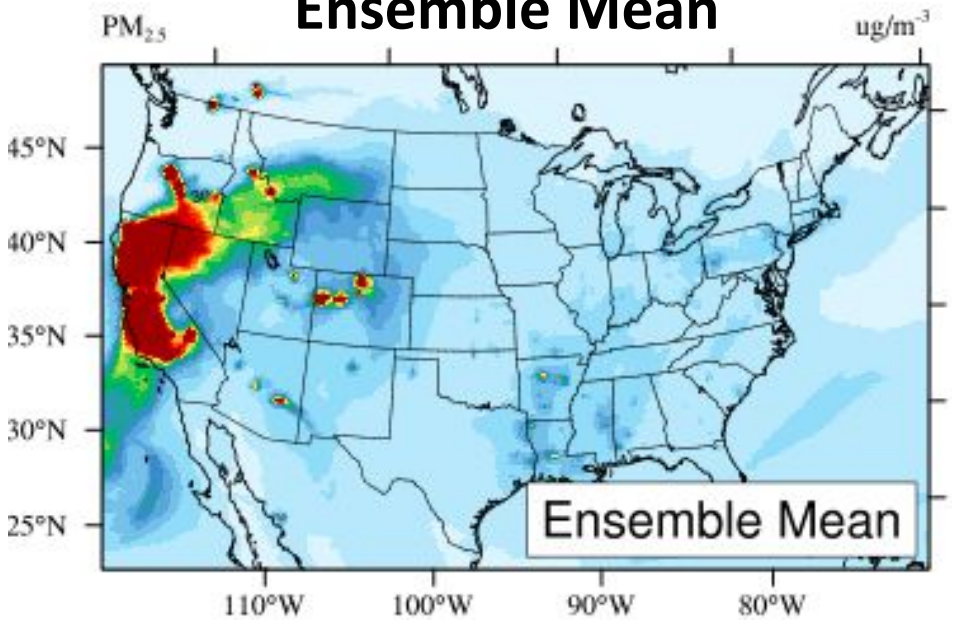
GMU-CMAQ (George Mason University Community Multiscale Air Quality), NACC-CMAQ (NOAA-EPA Atmosphere-Chemistry Coupler Community Multiscale Air Quality), HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory), HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory), GEFS-Aerosols (Global Ensemble Forecast System Aerosols), GEOS (NASA Goddard Earth Observing System), ICAP-MME (International Cooperative for Aerosol Prediction Multi-Model aerosol forecasting Ensemble), NAAPS (Navy Aerosol Analysis and Prediction System)

Surface $PM_{2.5}$ Simulations (Aug 22, 2020)

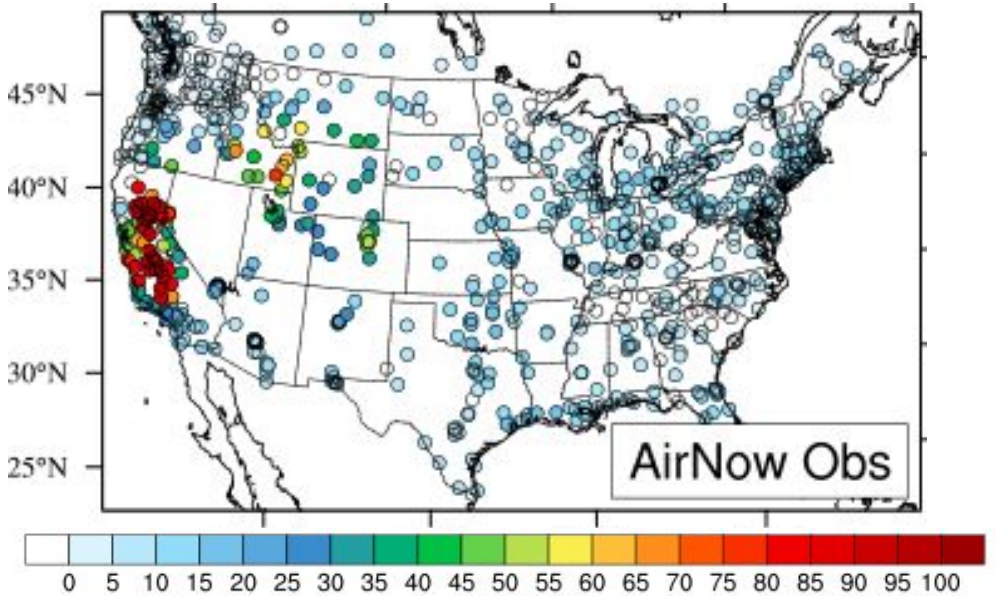
Individual Models



Ensemble Mean

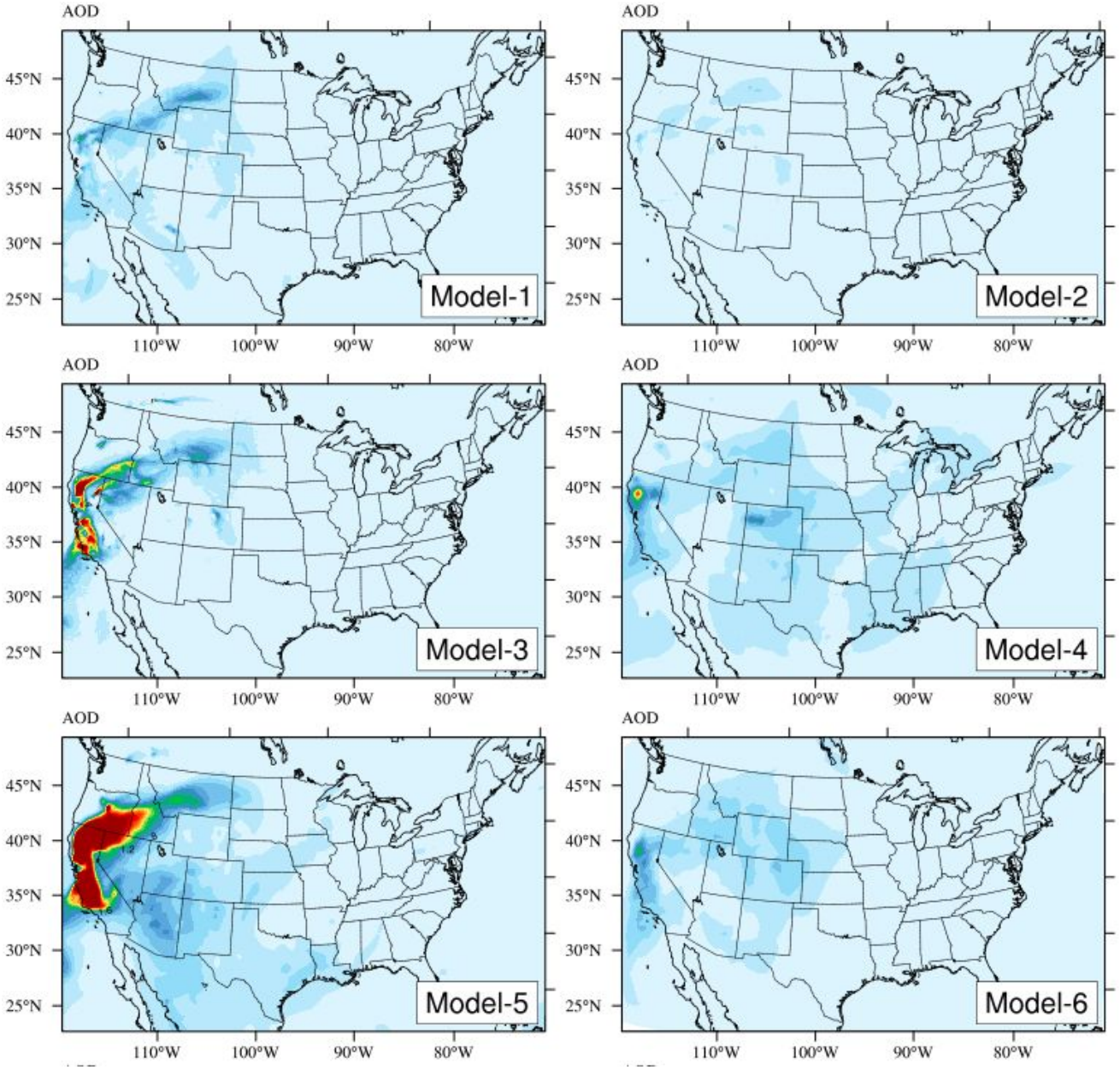


Observation



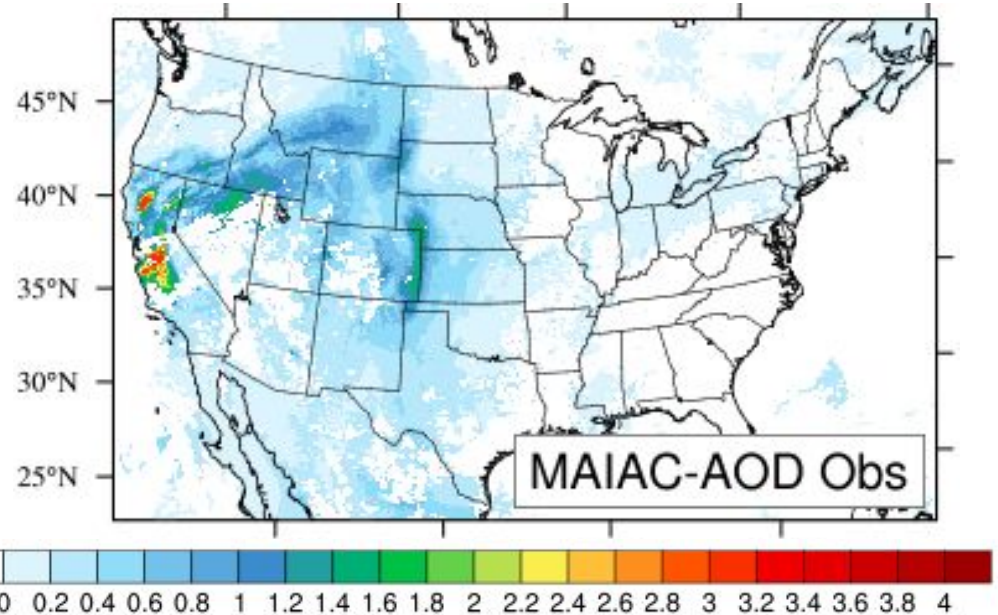
AOD Simulations (Aug 22, 2020)

Individual Models



Ensemble Mean

Observation



Model Evaluation Metrics

1. Traditional metrics

Root Mean Square Error **(RMSE)** Correlation **(CORR)**, Fractional Bias **(FB)**, Mean Bias**(MB)**, Mean Error **(ME)**, Normalized mean Bias**(NMB)**, Normalized Mean Error**(NME).**

2. Categorical metrics (Kang et al., 2007)

Used to measure performance of model in forecasting $PM_{2.5}$ exceedances (>35 µg/m^3) during extreme fire periods.

Area Hit Rate (aH)

Refers to **a number of hits if a forecast exceedance is observed within the designated area** centered at the observation location.

$$aH = \left(\frac{Ab}{Ab + Ad}\right) \times 100\%$$

Area False Alarm Ratio (aFAR)

Refers to **a number of forecast false-alarm ratios if a forecast exceedance is not observed** within the area centered at the observation location.

$$aFAR = \left(\frac{Aa}{Aa + Ab}\right) \times 100\%$$

Aa is the number of forecast exceedances that were not observed, **Ab** is the number of forecast exceedance that were observed, and **Ad** is the number of observed exceedances that were not forecasted.

3. Overall Rating *(RANK)*

Used to determine the overall forecasting performance of ensemble mean.

For $PM_{2.5}$, Rank = sum of Normalized CORR, FB, aH, aFAR, ranging from **0 (worst) to 4 (best)**.

For AOD, **Rank = sum of Normalized CORR, FB,** ranging from **0 (worst) to 2 (best)**.

Note: N=Normalized; NCORR=(CORR+1)/2; NFB=1-(FB/2); NaH=aH/100; NaFAR=1-(aFAR/100)

Model Performance

1. $PM_{2.5}$ Prediction vs. AirNow

Table 1. Overall ensemble mean and individual model performances in forecasting $PM_{2.5}$ concentrations during the 2020 Gigafire event (August-September 2020)

Model	*RMSE*	*CORR*	*MB*	*aH*	*aFAR*	*FB*	***RANK***
Model-1	24.854	0.542	3.107	69.046	44.359	0.548	2.811
Model-2	16.726	0.477	-4.540	39.753	24.121	0.597	2.723
Model-3	19.714	0.432	-4.614	71.658	47.734	1.323	2.375
Model-4	48.878	0.496	17.809	81.245	75.933	0.881	2.402
Model-5	49.388	0.438	11.566	80.438	68.652	0.773	2.493
Model-7	31.054	0.544	10.145	82.195	62.706	0.664	2.673
Ensemble Mean	24.059	0.609	6.537	86.827	60.393	0.530	**2.832**

2. AOD Prediction vs. MAIAC

Table 2. Overall ensemble mean and individual model performances in forecasting AOD values during the 2020 Gigafire event (August-September 2020)

Models	*RMSE*	*CORR*	*MB*	*FB*	***RANK***
Model-1	0.280	0.569	-0.154	0.983	1.293
Model-2	0.296	0.524	-0.182	1.184	1.170
Model-3	0.355	0.368	-0.128	1.345	1.012
Model-4	0.241	0.521	-0.075	0.614	1.453
Model-5	0.858	0.458	0.201	0.773	1.342
Model-6	0.259	0.525	-0.128	0.897	1.314
Ensemble Mean	0.276	0.587	-0.074	0.711	**1.469**

The highest *RANK* is highlighted in **bold red**.

Ensemble Probability Forecast of $PM_{2.5}$ Exceedances

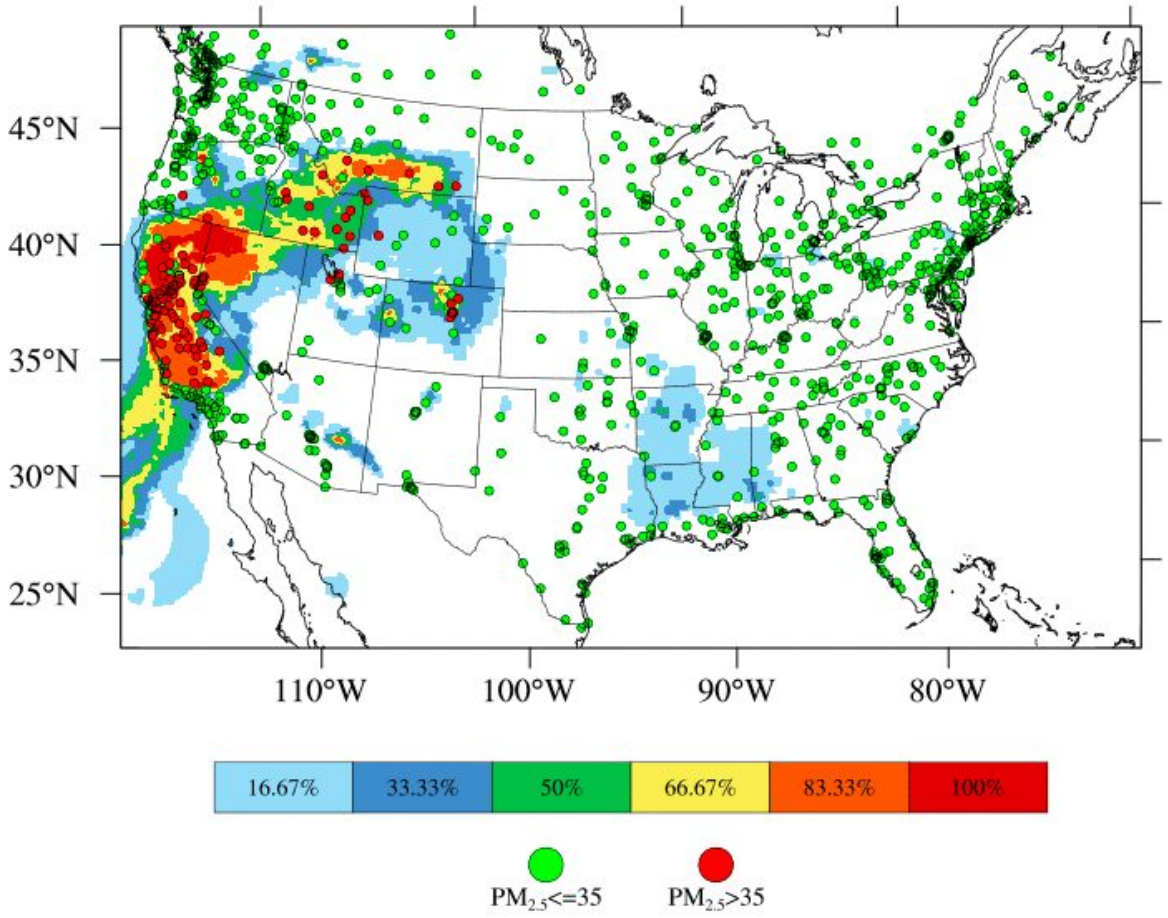
Ensemble forecast can provide a probabilistic forecast based on the spread of the ensemble members.

6 Members: GMU-CMAQ, NACC-CMAQ, GEFS-Aerosols, GEOS, HYSPLIT, and NAAPS.

P(A) = Number of models that forecast the exceedances/ Total number of models

100% : **All models** forecast the exceedances
83.33% : 5 out of 6 models forecast the exceedances
66.67% : 4 out of 6 models forecast the exceedances
50% : 3 out of 6 models forecast the exceedances
33.33% : 2 out of 6 models forecast the exceedances
16.67% : 1 out of 6 models forecast the exceedances
0% : **None of models** forecast the exceedances

Ensemble Probability Forecast - $PM_{2.5}$ Exceedances



Conclusions

1. On average the ensemble forecast can reduce bias and uncertainties in forecasting by averaging them out.
2. In most cases, the ensemble forecast provides superior forecasting performance compared to the individual models.
3. Our next step is to test the multi-model ensemble forecast with the 2021 Spring dust event (January-March 2021) in the Western US.

Acknowledgement

Funding support: NASA Health and Air Quality Program

- **GMU-CMAQ** from Air Quality Group of George Mason University (http://air.csiss.gmu.edu/);
- **NACC-CMAQ** and **GEFS-Aerosols** from NOAA Air Resources Laboratory (ARL)/GMU;
- **HYSPLIT** from NOAA Air Resources Lab;
- **GEOS** from NASA Goddard Space Flight Center (GSFC);
- **ICAP-MME** and **NAAPS** from Naval Research Laboratory (NRL);
- **MAIAC** from NASA GSFC;