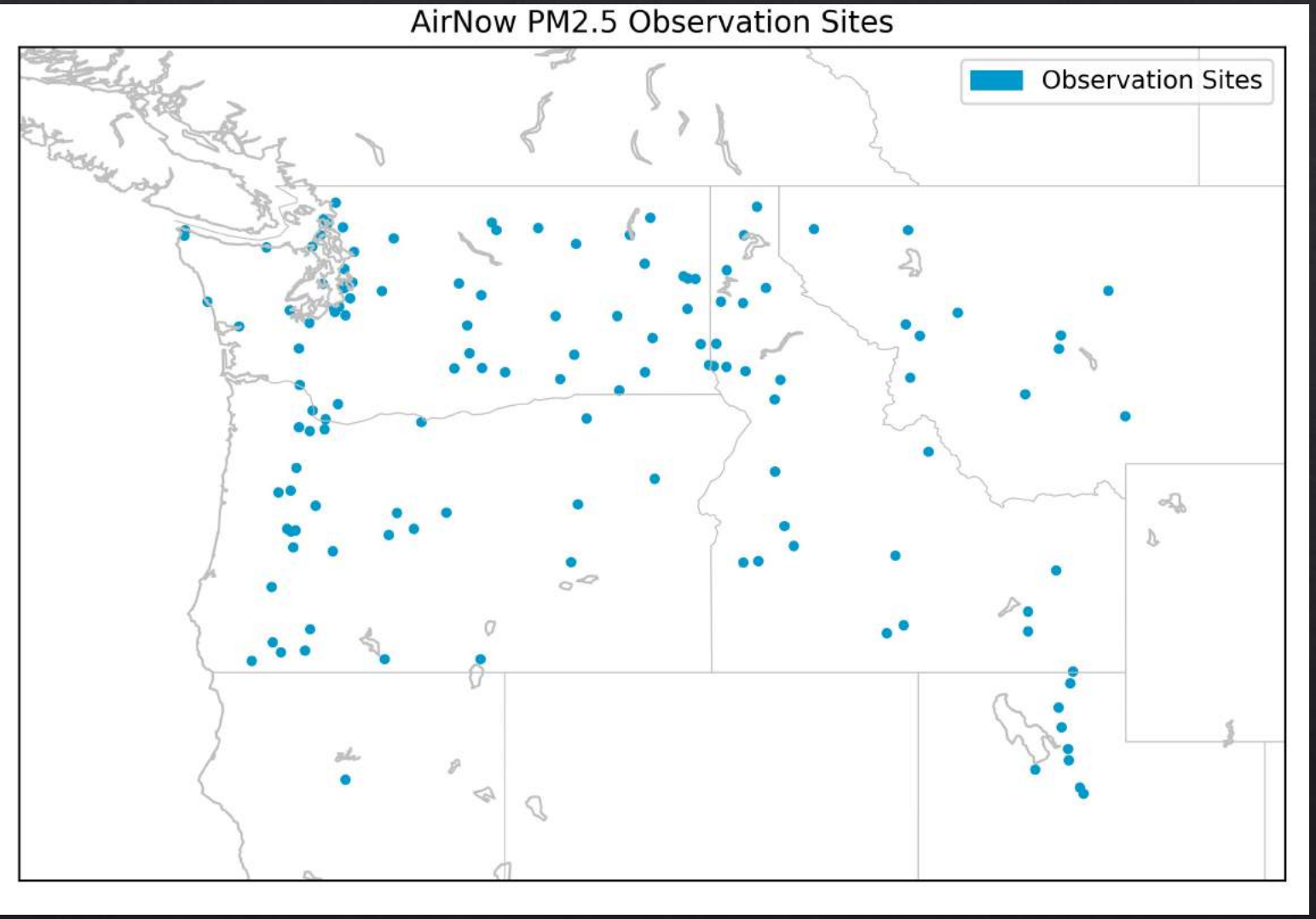
A Kalman Filter Bias Correction for 24-hr PM2.5 Forecasts

Nicole June (PSU),
Joe Vaughan, Yunha Lee, Brian Lamb (LAR)

Laboratory for Atmospheric Research, Pennsylvania State University

Methodology for correcting 24-hr Average PM2.5

- A Kalman Filter (K-F) using matched forecast and observed values at PM2.5 monitoring sites from a multi-day training period computes model bias statistics for that time period.
- The K-F statistics are used to estimate the bias at each site for the next 24-hours of the model run.
- The K-F-estimated bias values are used to correct the model results at the monitoring sites.
- Initial analysis done for all of 2017 AIRPACT Model and AirNow Observations, an then the method was applied to the 2018 data.
- *What minimum number of observations is needed for the post-processing to add skill to the forecast?*
- *How and Why does the performance change on a site by site basis?*

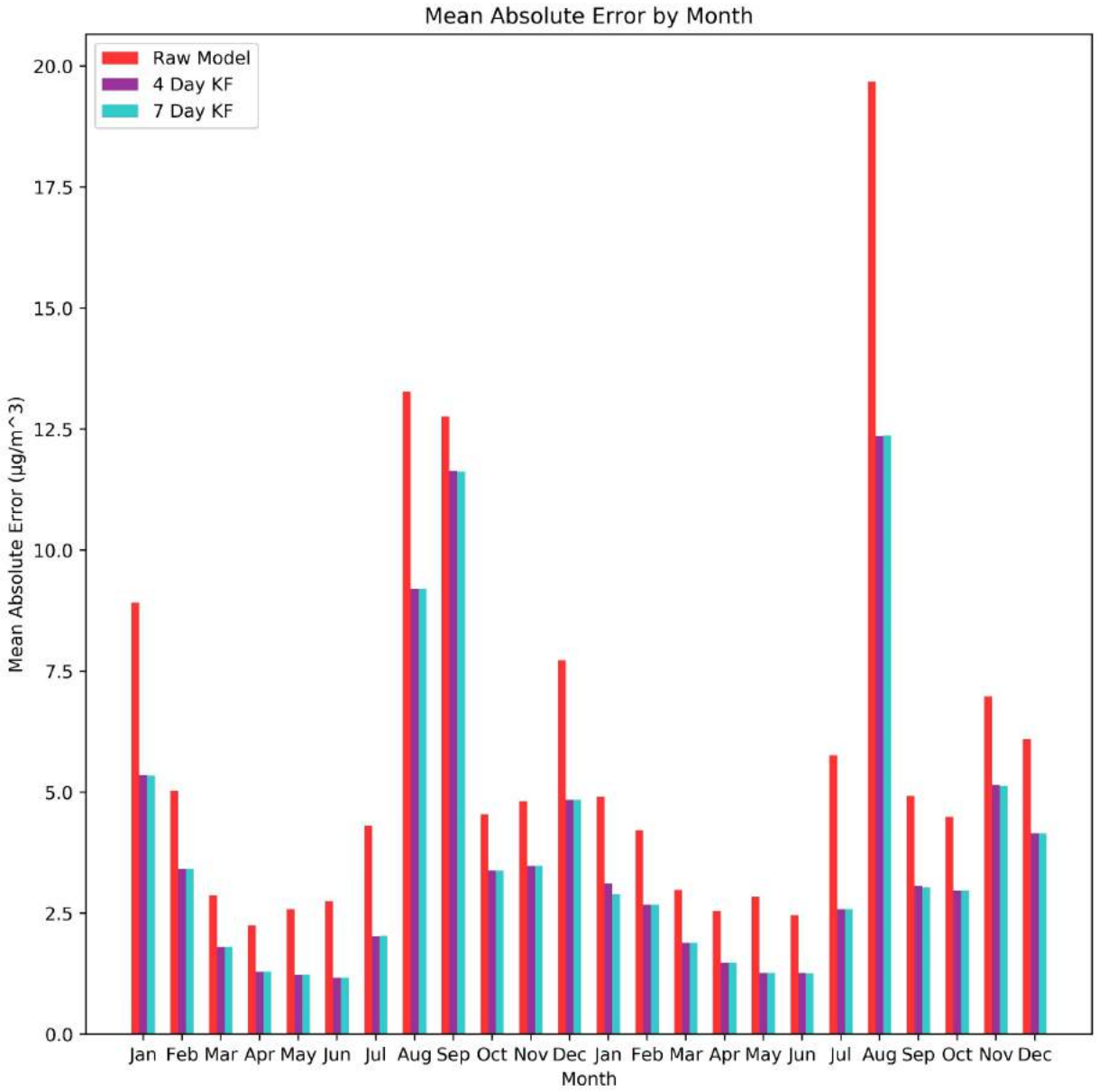


- Locations of the AirNow Observing Sites for $PM_{2.5}$ in 2017
- Sites used in analysis were required to submit reports (either concentrations or NAN's) for 8736 hours of the year
- The AIRPACT Model results were extracted at these observing sites
- The Kalman Filter is then applied at the observing sites to create a post-processed model at the observation site

Results: Error and Bias for AIRNow monitoring sites Raw Model vs Kalman-Filter Bias Corrected 4-Day & 7- Day 2017 and 2018

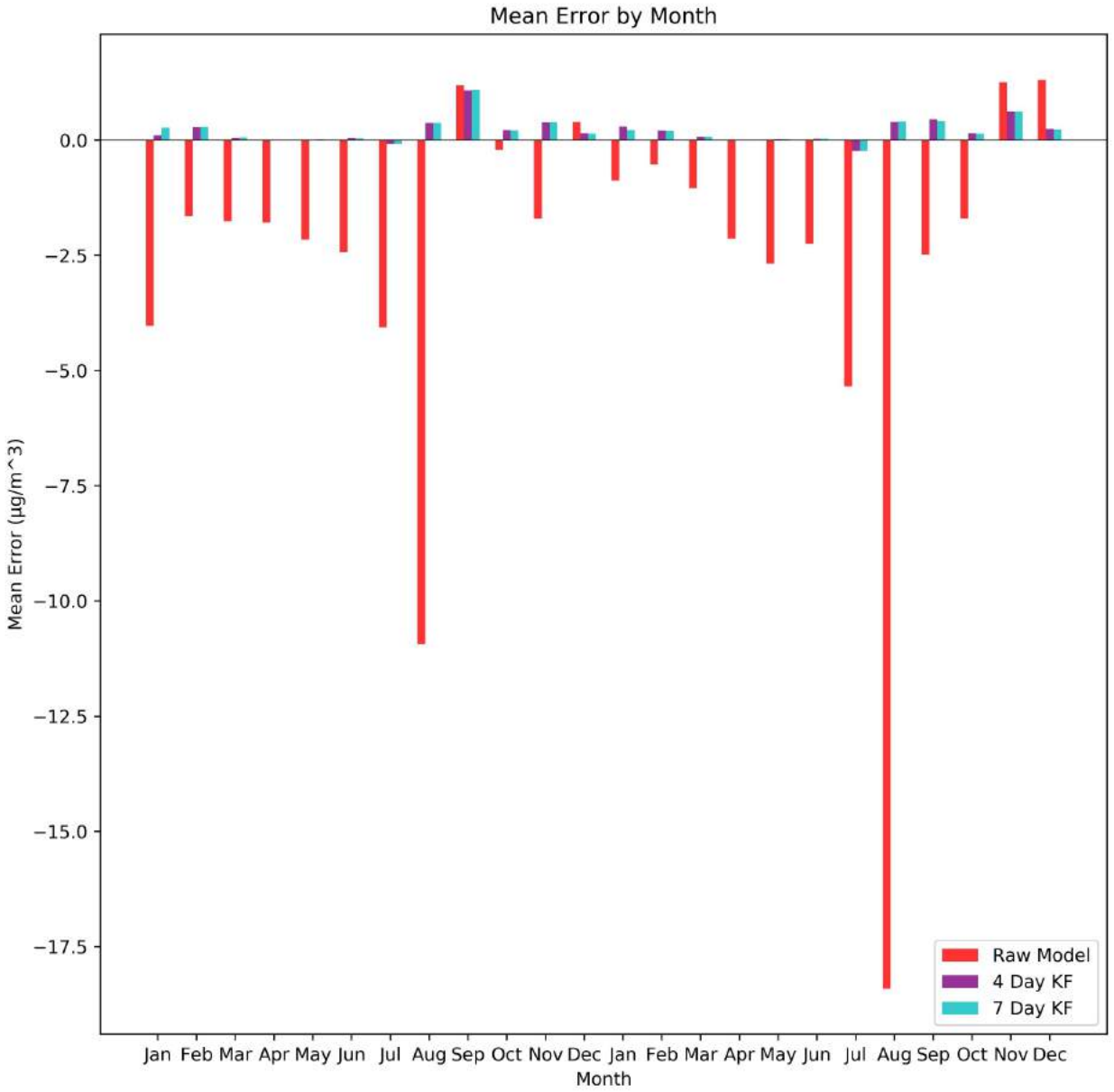
Monthly Mean Absolute Error

	RAW MODEL		7 DAY KALMAN FILTER		4 DAY KALMAN FILTER	
	2017	**2018**	**2017**	**2018**	**2017**	**2018**
January	8.909718	4.902606	5.345553	5.345553	5.355218	3.107664
February	5.030652	4.211946	3.41293	3.41293	3.412707	2.672424
March	2.868526	2.980642	1.80533	1.80533	1.803031	1.89064
April	2.247444	2.546284	1.290926	1.290926	1.286126	1.469767
May	2.582014	2.840444	1.226644	1.226644	1.221099	1.266152
June	2.747072	2.462374	1.164501	1.164501	1.164483	1.263965
July	4.302645	5.762219	2.031474	2.031474	2.028798	2.582874
August	13.26847	19.67328	9.202931	9.202931	9.209404	12.36353
September	12.76633	4.92689	11.6195	11.6195	11.63737	3.061026
October	4.538321	4.488383	3.381086	3.381086	3.386001	2.974732
November	4.811376	6.978717	3.483787	3.483787	3.481842	5.144407
December	7.728302	6.087963	4.845655	4.845655	4.845086	4.150401
Year	5.983406	5.65515	4.067526	4.06753	4.069264	3.495632


Mean Absolute Error by Month
Raw Model
4 Day KF
7 Day KF
Mean Absolute Error (µg/m^3)
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Month

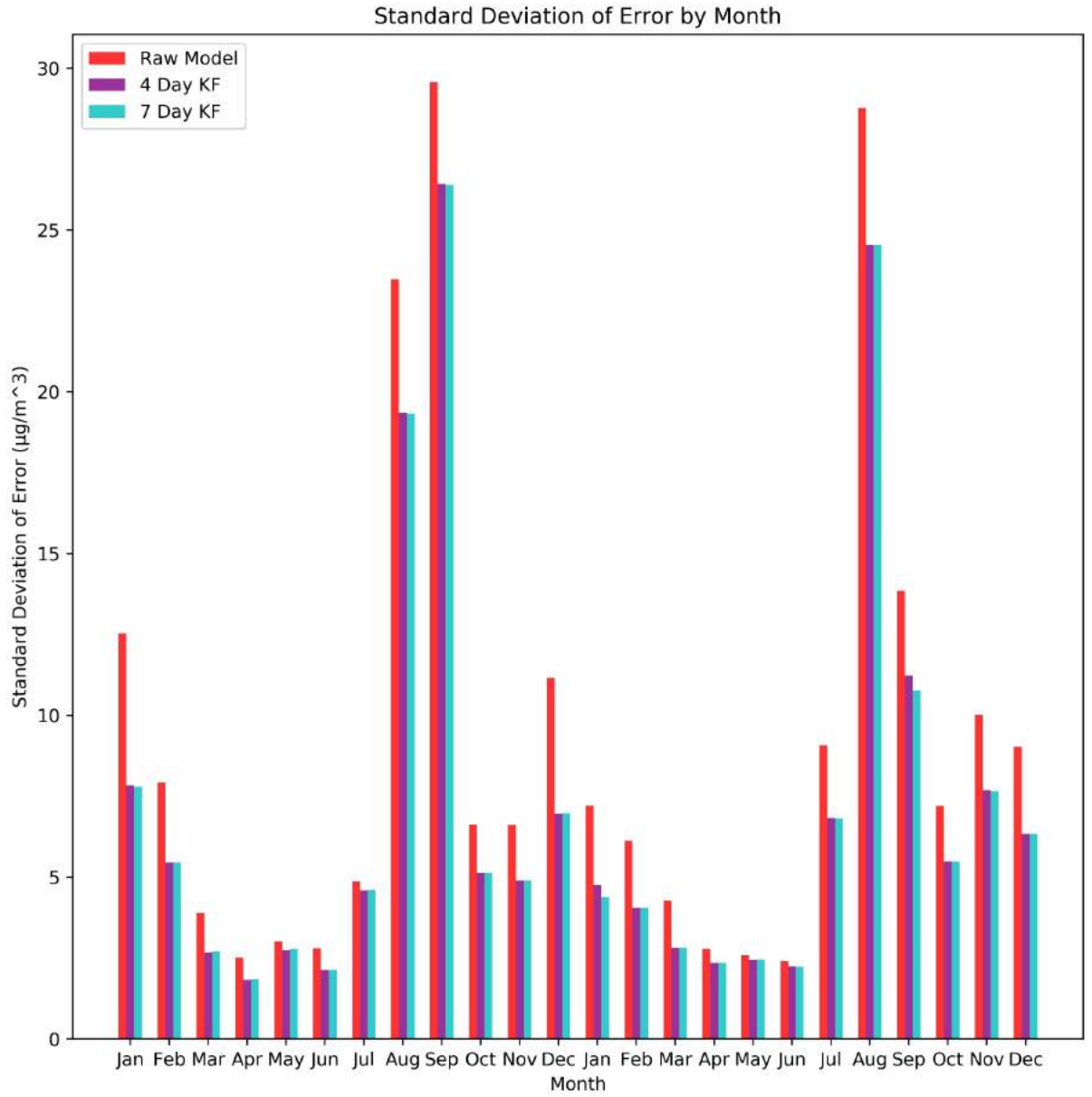
Monthly Mean Error

	RAW MODEL		7 DAY KALMAN FILTER		4 DAY KALMAN FILTER	
	2017	**2018**	**2017**	**2018**	**2017**	**2018**
January	-4.03456	-0.87876	0.266227	0.21502	0.105593	0.295566
February	-1.64901	-0.52835	0.283226	0.198302	0.27584	0.203463
March	-1.75375	-1.0445	0.058805	0.071086	0.050871	0.069188
April	-1.79186	-2.14071	0.001174	-0.00652	-0.00532	-0.00717
May	-2.16074	-2.68045	0.016113	0.019649	0.006375	0.017672
June	-2.4346	-2.24341	0.044194	0.032258	0.047619	0.025783
July	-4.05949	-5.3408	-0.08322	-0.23097	-0.08861	-0.23104
August	-10.9338	-18.4137	0.367399	0.401315	0.368247	0.391398
September	1.185136	-2.4903	1.083831	0.416971	1.075255	0.446879
October	-0.2072	-1.70084	0.212423	0.142986	0.21521	0.147929
November	-1.70543	1.251453	0.385801	0.616739	0.380805	0.619598
December	0.388847	1.309756	0.141992	0.235877	0.151664	0.244148
Year	-2.4297	-2.90838	0.231497	0.176059	0.215296	0.185285


Mean Error by Month
Mean Error (µg/m^3)
0.0
−2.5
−5.0
−7.5
−10.0
−12.5
−15.0
−17.5
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Month
Raw Model
4 Day KF
7 Day KF

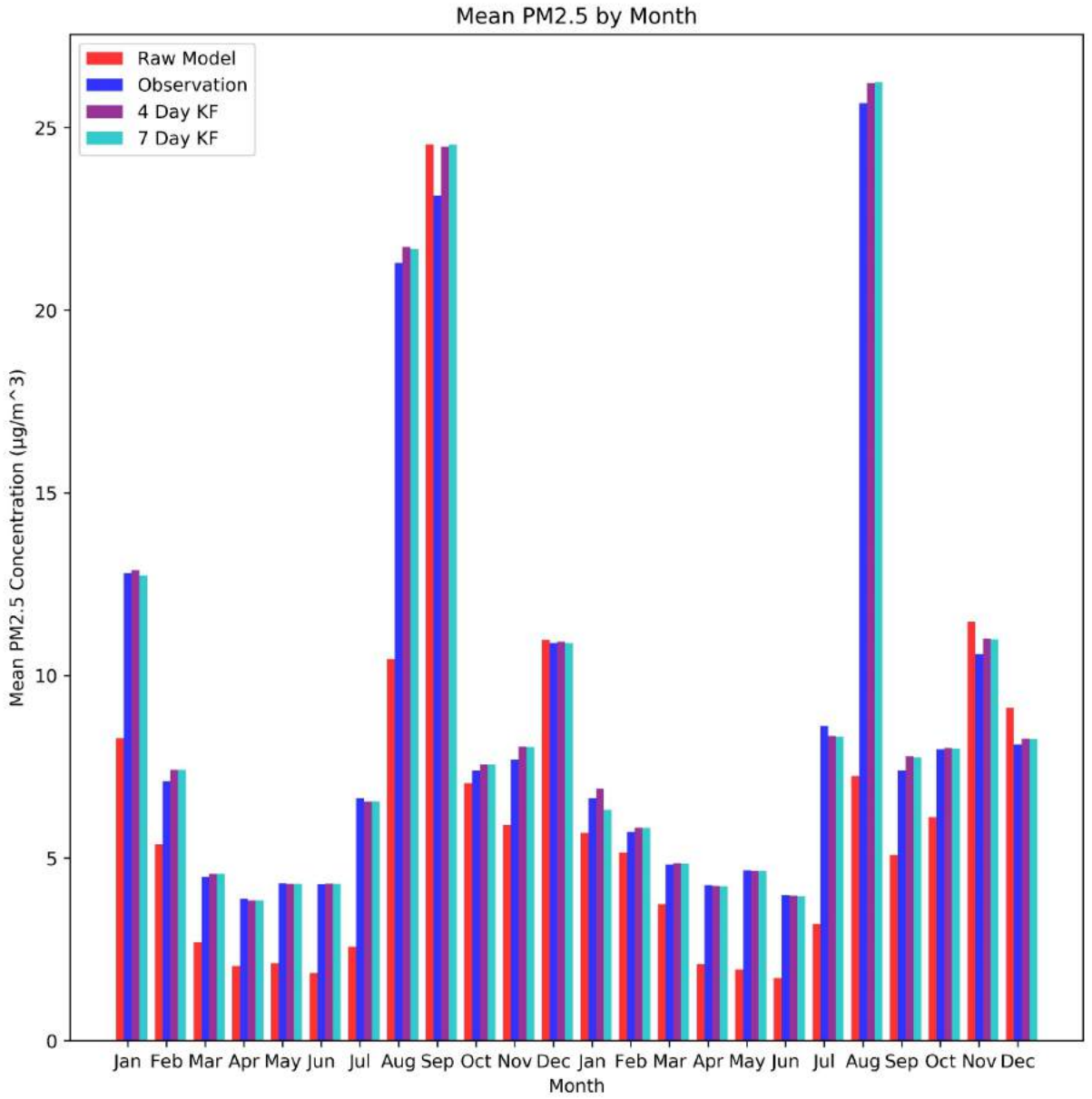
Monthly Standard Deviation

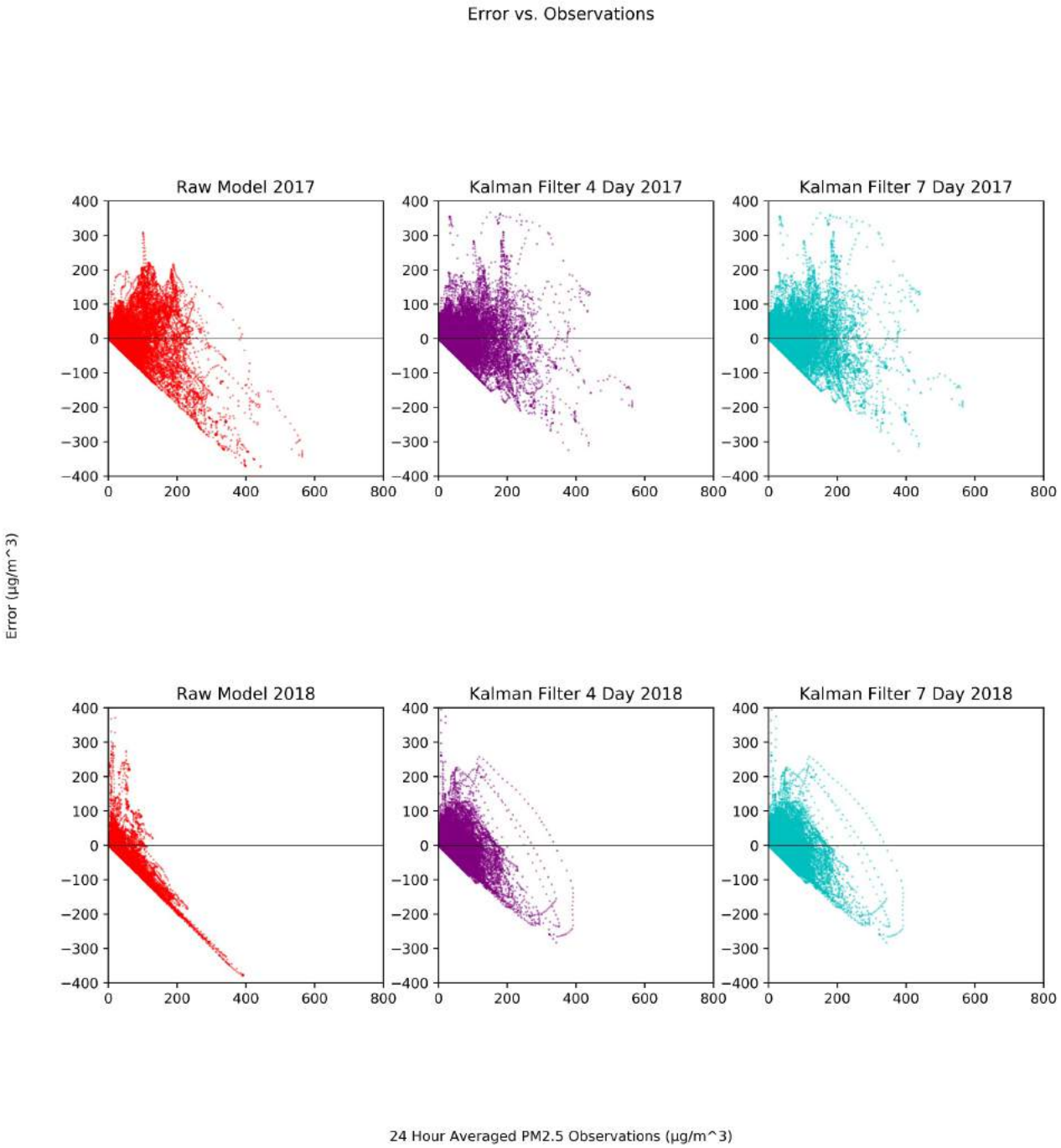
	RAW MODEL		7 DAY KALMAN FILTER		4 DAY KALMAN FILTER	
	2017	**2018**	**2017**	**2018**	**2017**	**2018**
January	12.52634	7.208809	7.794546	4.389188	7.838665	4.759086
February	7.924141	6.11841	5.461062	4.045212	5.461634	4.046892
March	3.883929	4.275076	2.698284	2.817455	2.673962	2.818172
April	2.516202	2.787169	1.837595	2.342173	1.825135	2.343662
May	3.007754	2.588562	2.78095	2.449165	2.733097	2.445553
June	2.808765	2.41205	2.125693	2.230989	2.129681	2.24302
July	4.870946	9.072871	4.600086	6.810187	4.59476	6.821267
August	23.47832	28.76912	19.31684	24.53417	19.35229	24.53079
September	29.5621	13.84742	26.3945	10.77054	26.42952	11.23037
October	6.62626	7.197671	5.128417	5.483366	5.131415	5.489233
November	6.616605	10.01746	4.895887	7.662619	4.896157	7.692192
December	11.15786	9.025256	6.963094	6.336501	6.959912	6.335457
Year	13.050249	12.453806	10.537621	9.216128	10.535945	9.254659


Standard Deviation of Error by Month
Raw Model
4 Day KF
7 Day KF
Standard Deviation of Error (µg/m^3)
0
5
10
15
20
25
30
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Month

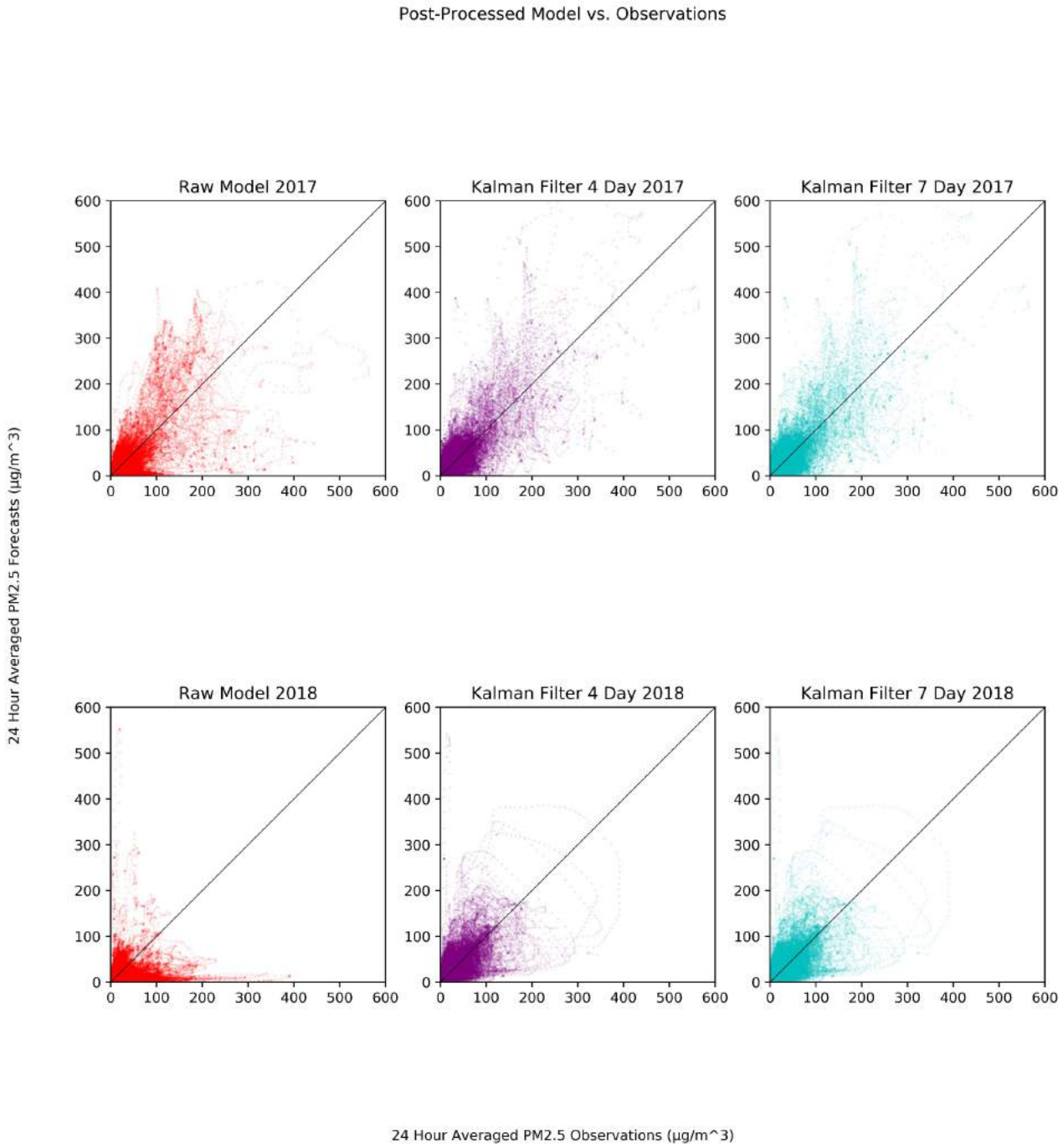
Monthly Mean $PM_{2.5}$ Concentration

	RAW MODEL		Observations		7 DAY KALMAN FILTER		4 DAY KALMAN FILTER	
	2017	2018	2017	2018	2017	2018	2017	2018
January	8.28851	5.699426	12.80267	6.639265	12.7471	6.328584	12.87994	6.908358
February	5.378852	5.148707	7.10567	5.718682	7.41026	5.82682	7.410439	5.829945
March	2.707802	3.741924	4.489457	4.825785	4.569638	4.858179	4.570515	4.863318
April	2.048695	2.09358	3.895503	4.259541	3.851614	4.225601	3.848973	4.234428
May	2.12638	1.955234	4.314627	4.665373	4.303919	4.65052	4.298386	4.649957
June	1.855296	1.715088	4.295029	3.987191	4.297427	3.954125	4.307701	3.970056
July	2.570795	3.200261	6.639754	8.619082	6.552481	8.319887	6.548544	8.343664
August	10.45323	7.246531	21.30007	25.66879	21.67699	26.23761	21.73793	26.21552
September	24.54133	5.095789	23.141	7.406608	24.54028	7.759717	24.47912	7.783826
October	7.05642	6.126661	7.399625	7.980833	7.56121	7.99821	7.565725	8.016393
November	5.912977	11.47736	7.710274	10.58856	8.035108	10.99049	8.046656	11.00996
December	10.96958	9.115511	10.88879	8.112236	10.8809	8.259551	10.9288	8.265057


Mean PM2.5 by Month
Raw Model
Observation
4 Day KF
7 Day KF
Mean PM2.5 Concentration (µg/m^3)
0
5
10
15
20
25
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Month



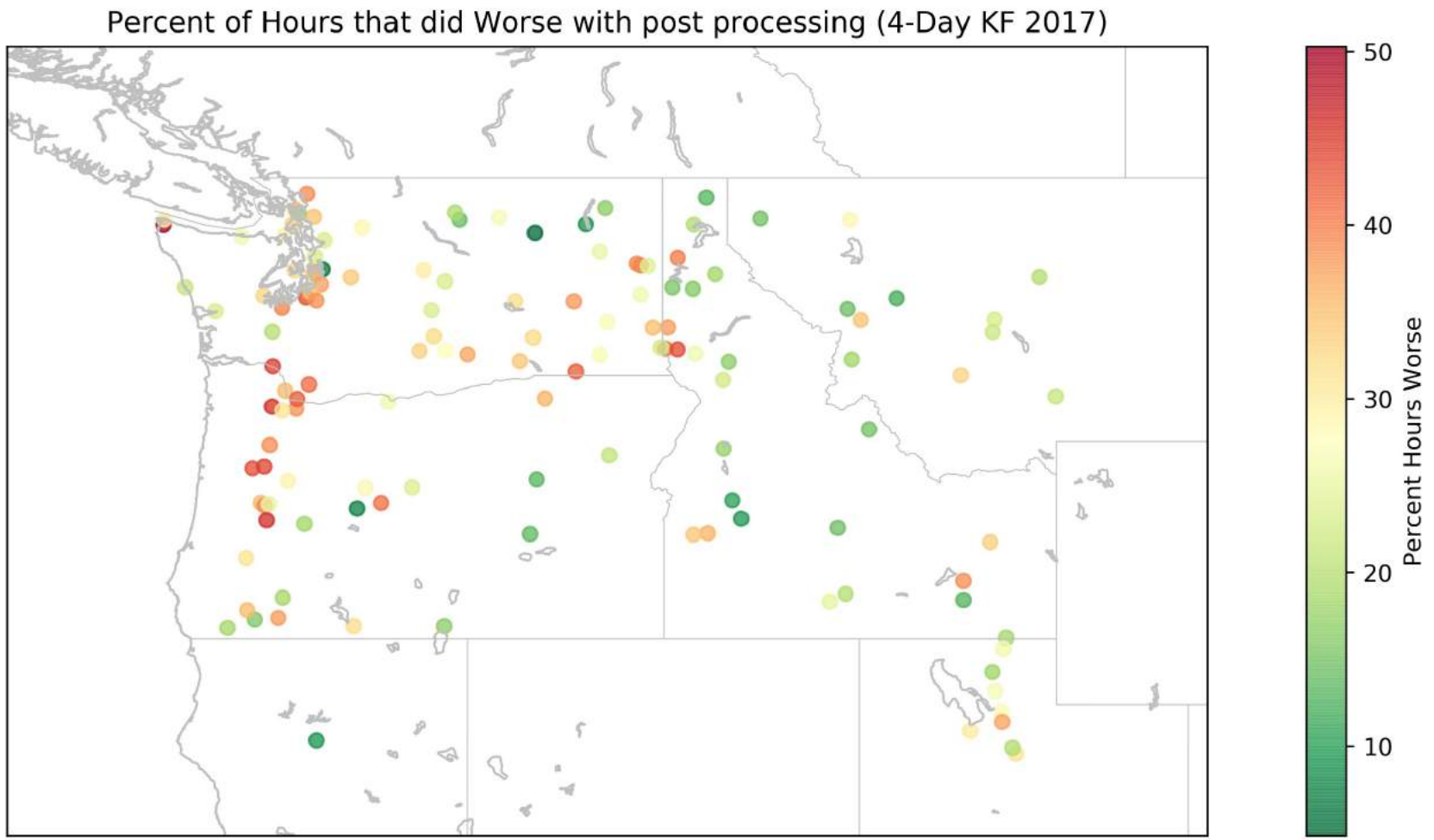
- This shows the error versus observations for 2017 and 2018
- The general trend for the raw model and the two Kalman filter techniques is as the observed concentration of $PM_{2.5}$ increases the error becomes more negative



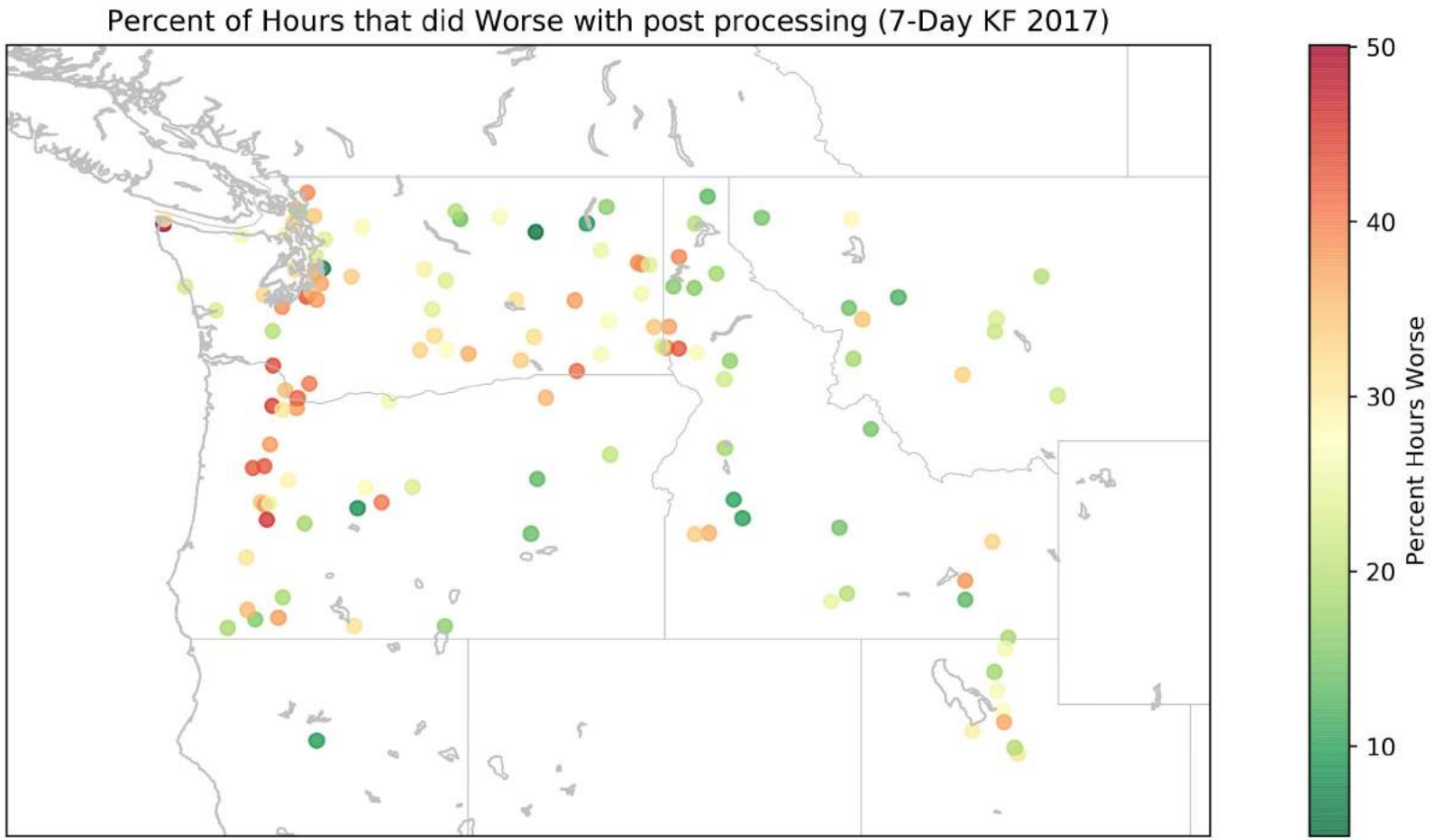
- This plot shows the raw model/ post-processed model versus the observations
- The y=x line is plotted to represent the line that would be the perfect model

Method	Pearson Correlation Coefficient	
	2017	**2018**
Raw Model	0.71	0.31
4-Day Kalman Filter	0.82	0.73
7-Day Kalman Filter	0.82	0.74

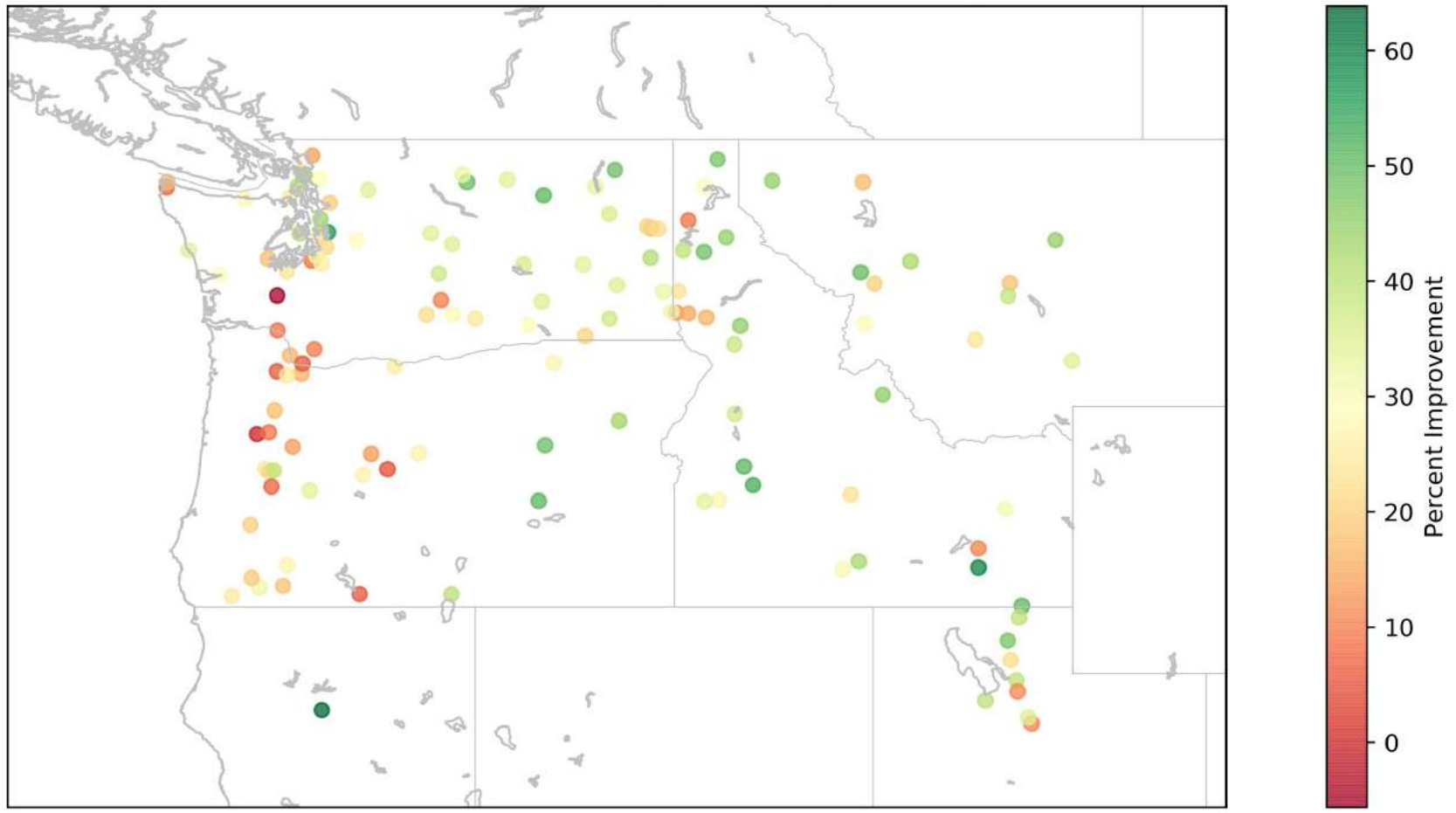
Additional Graphs/Maps for 2017



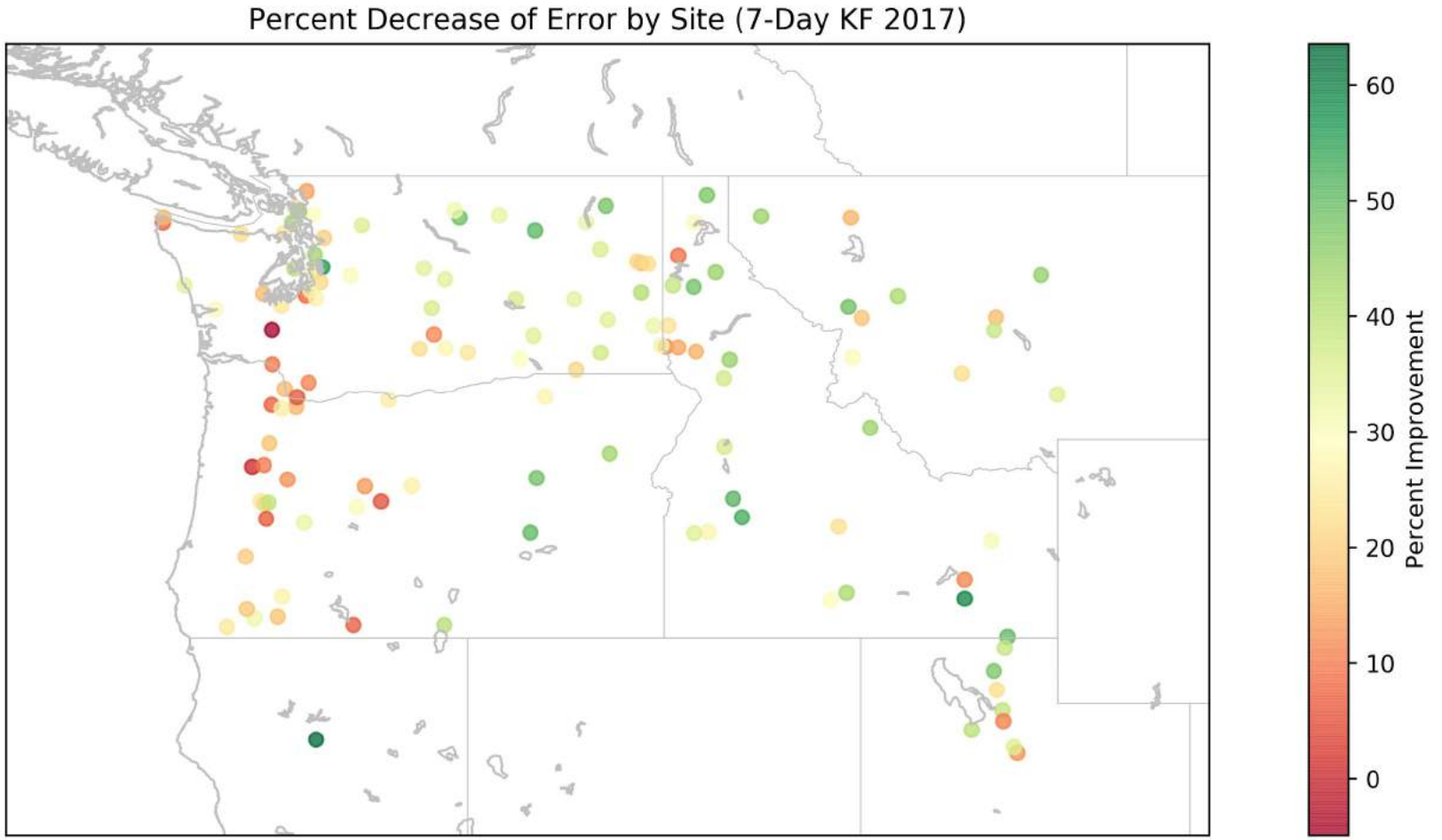
- This figure shows the percent of hours in 2017 where the post-processing did worse than the raw-model
- Worse is quantified as where the absolute value of the error of the post-processed model is greater than that of the raw model
- If the post-processed model does worse than the raw model its typically by not much more than 0.01 µg/m^3



- Same as previous slide just for the 7-Day Kalman Filter



- This figure shows the percent decrease of error by site for the 4-day Kalman filter



- Same as previous slide just for the 7-Day Kalman Filter

Above work was all at monitoring points!

Now, how to use these results to correct the forecast?

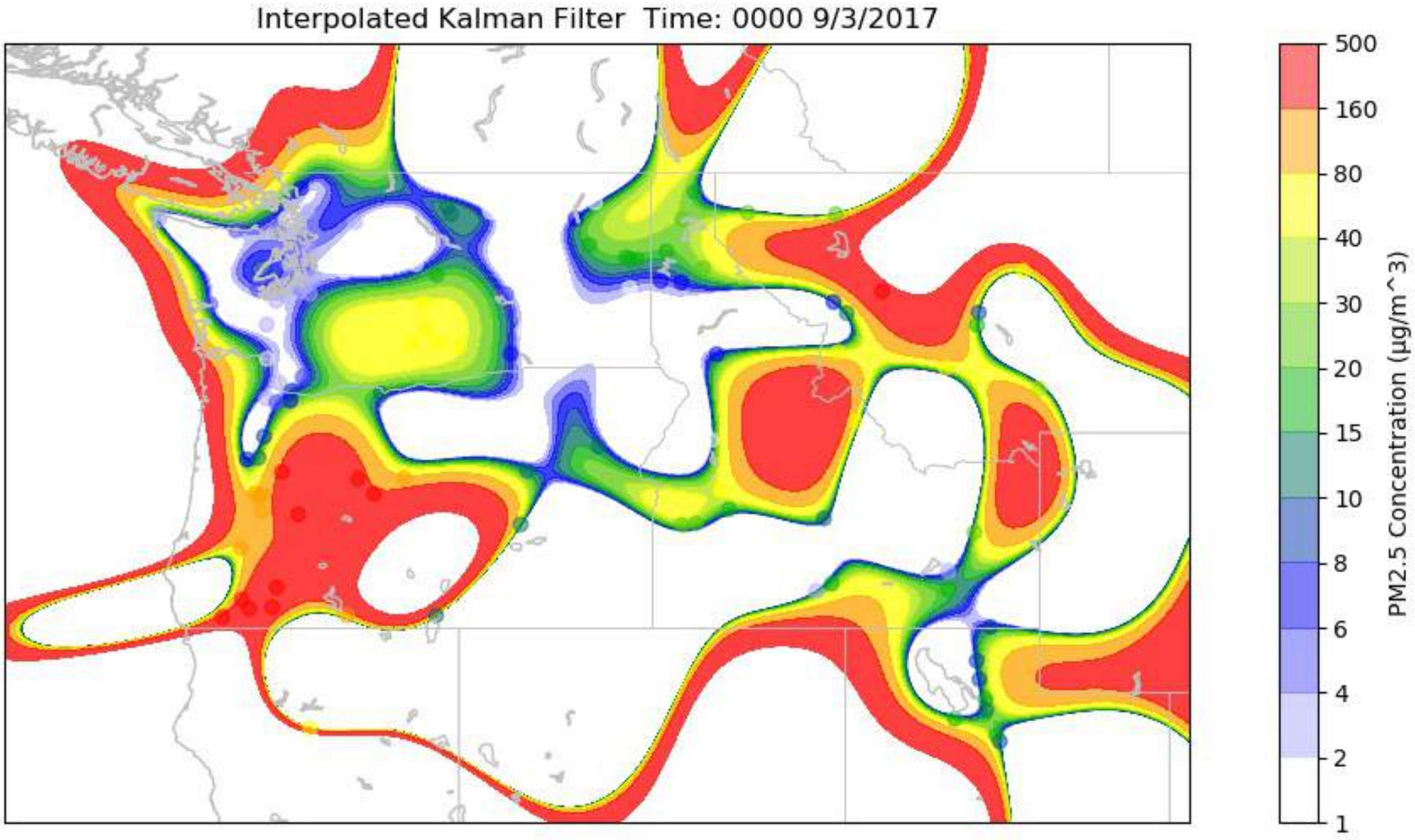
1) Apply corrections to model forecast at points and interpolate results thru domain.

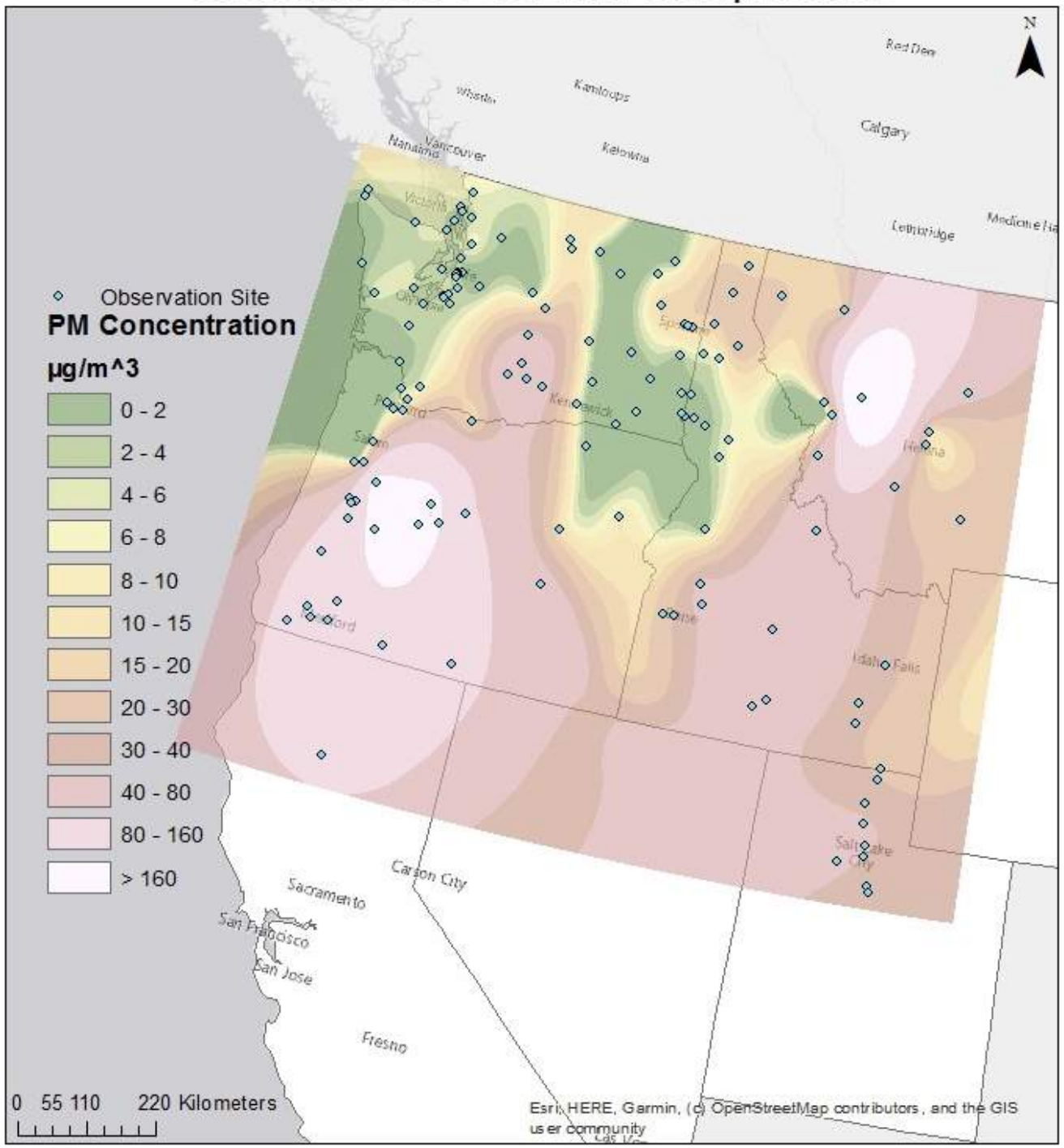
2) Interpolate bias correction thru domain and apply as correction to forecast.

Interpolation Methods

- Barnes Interpolation Method
  - 200 km Search Radius for Observation Sites
- Cressman Interpolation Method
  - 200 km Search Radius for Observation Sites
- Cubic Interpolation Method
  - 200 km Search Radius for Observation Sites
- Linear Interpolation Method
- Kriging (still being debugged – **see next two slides**!!!)

All tested for getting domain-wide results from point values of bias-corrected forecast.


Interpolated Kalman Filter Time: 0000 9/3/2017
500
160
80
40
30
20
15
10
8
6
4
2
1
PM2.5 Concentration (µg/m^3)



This is Kriging interpolation from ArcGIS for the same time as the image on the previous slide.

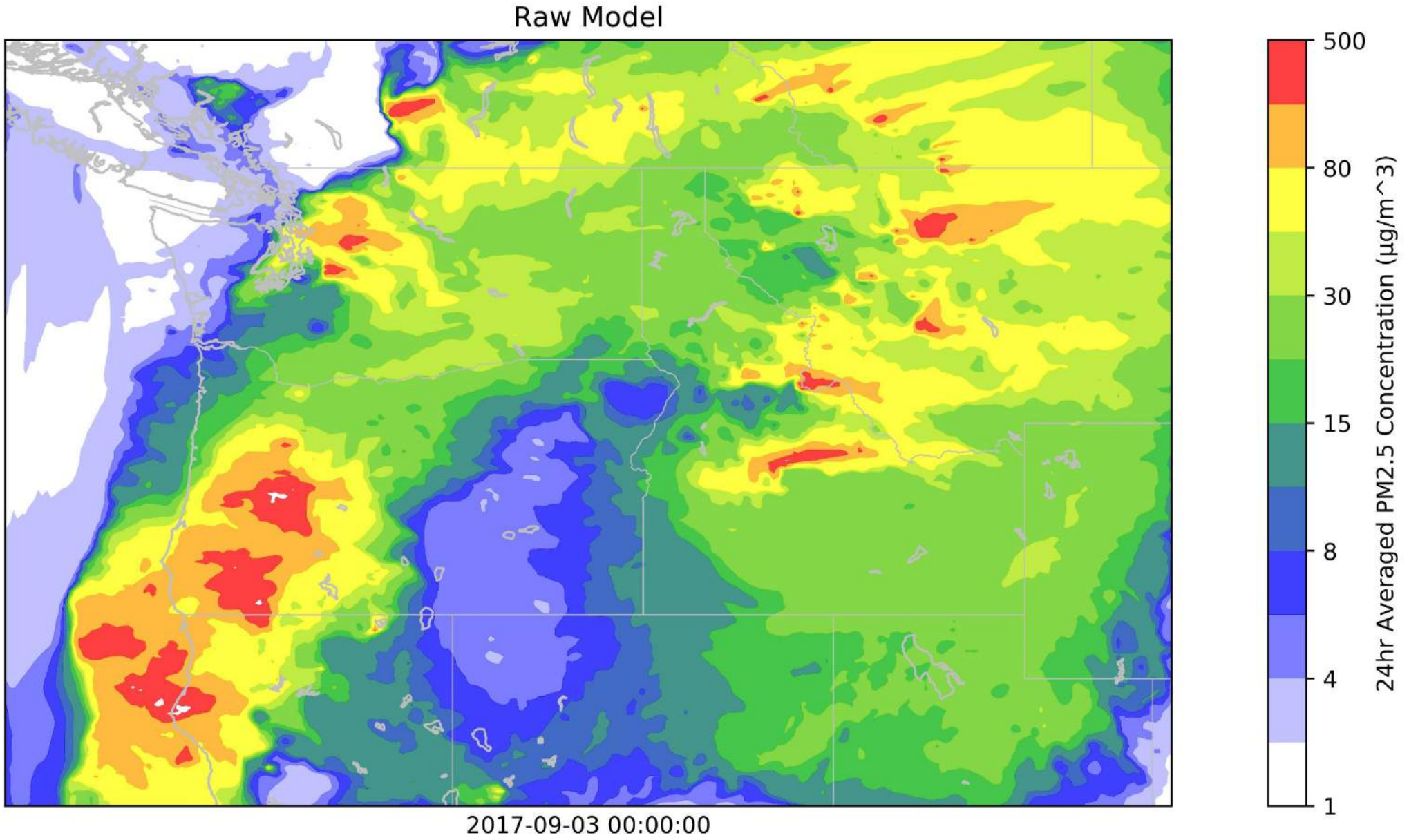
Interpolation Methods

- Barnes Interpolation Method
  - 200 km Search Radius for Observation Sites
- Cressman Interpolation Method
  - 200 km Search Radius for Observation Sites
- Cubic Interpolation Method
  - 200 km Search Radius for Observation Sites
- Linear Interpolation Method
- Kriging (still being debugged)

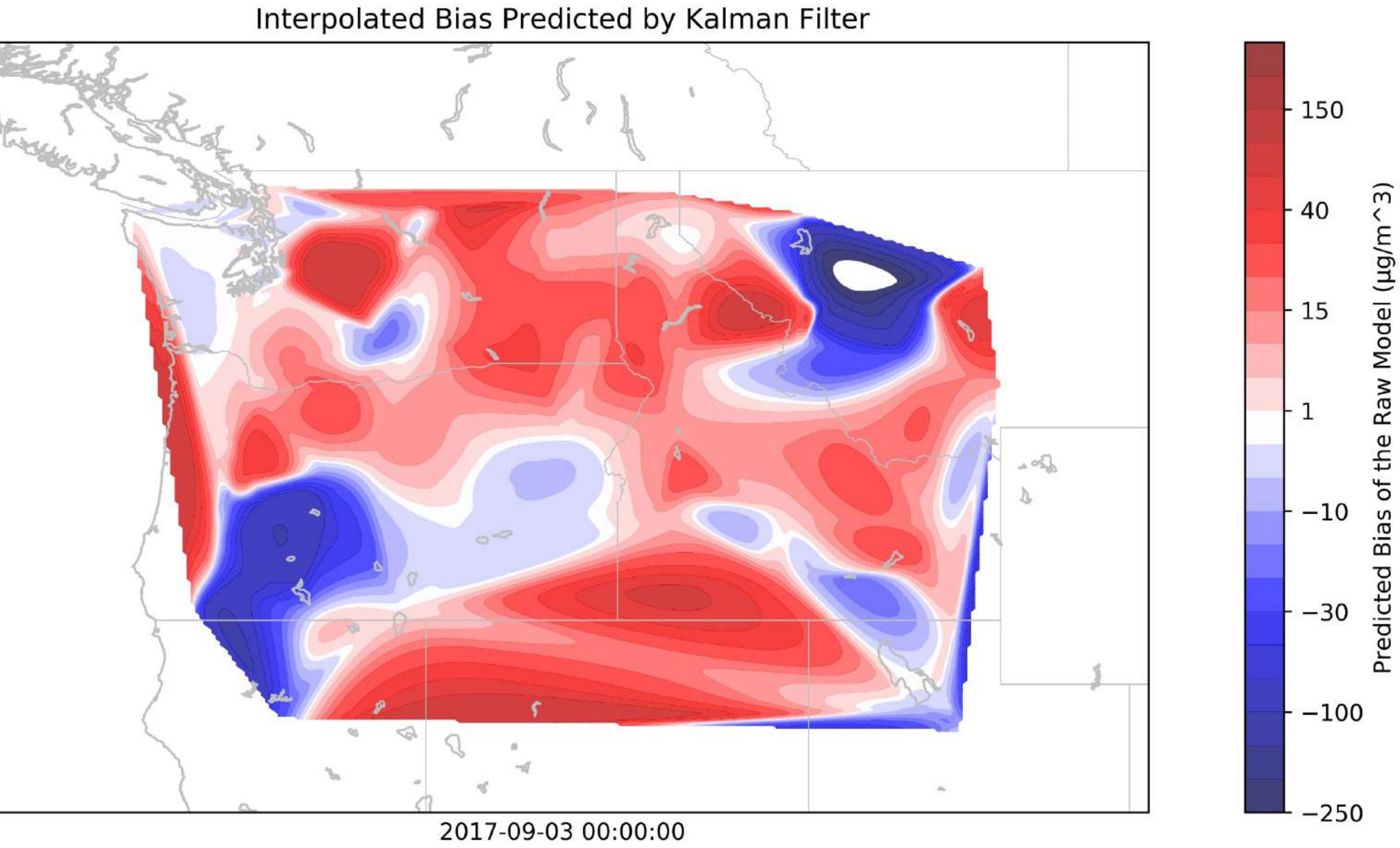
***NEXT: test getting domain-wide results for* bias for correction of model forecast.**

Second Approach for applying the Kalman Filter: Steps

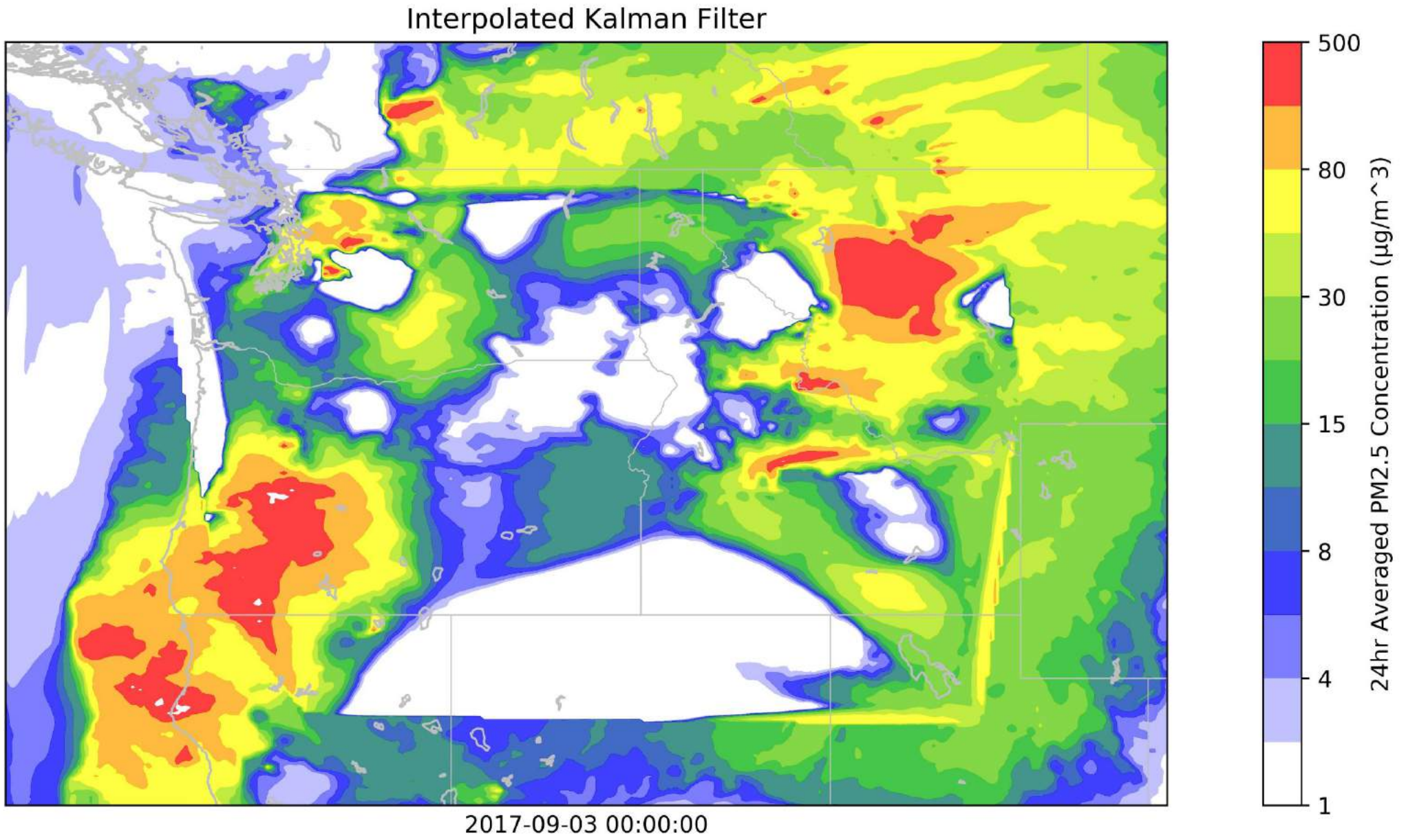
1. Use the Kalman Filter to Predict the bias for a given 24hr averaged period at each observation site within the AIRPACT domain
2. Interpolate the Kalman Filter bias using the Cubic method
3. Obtain the gridded 24 hour raw model PM2.5 concentrations for the same 24hr averaged period as the predicted bias
4. Subtract the interpolated Kalman Filter bias from the gridded model values



- The Raw Model for the 24hr period beginning at 9/3/2017 00:00:00 PDT



Interpolated Kalman Filter bias for the 24hr period beginning at 9/3/2017 00:00 PDT
Red indicates that the original model is overestimating the PM2.5 concentration
Blue indicates that the original model is underestimating the PM2.5 concentration
The extent of this interpolation is limited by the locations of AirNow sites observing PM2.5



- The gridded 24 hr PM2.5 Concentration obtained by doing the subtraction of the gridded model and the interpolated Kalman Filter Bias
- Unlike with the approach previously used this allows the forecast to extend to the boundaries of AIRPACT

Acknowledgements

We wish to acknowledge and thank for their input:

Sara Strachan (IDEQ),

Jill Shulte (ECY),

Brian Lamb (LAR and

Yunha Lee (LAR).