Supplementary material for "Interpretation of Probabilistic Surface Ozone Forecasts: A Case Study for Philadelphia"

Nikolay V. Balashov,[a,b] Amy K. Huff,[c] Anne M. Thompson[d]

[a]*NASA Global Modeling and Assimilation Office (GMAO), NASA Goddard Space Flight Center, Greenbelt, MD, 20771, USA*

[b]*Visiting Assistant Research Scientist, ESSIC, University of Maryland, 5825 University Research Court, College Park, MD, 20740, USA*

[c]*I.M. Systems Group, College Park, MD, 20740, USA*

[d]*Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, MD, 20771, USA and University of Maryland-Baltimore County, JCET, Baltimore, MD, 21228, USA*

The supplementary material explains how 56 weather regimes contribute to the REGiS O_3 MDA8 forecasts used in the article. It also demonstrates how REGiS chooses representative probabilistic forecasts from the five O_3 monitoring stations based on the highest forecasted exceedance probability.

In the training stage REGiS analyzes all the daily (0000 UTC) 500 hPa geopotential height and 850 hPa temperature fields from the MJJAS 2000-2011 timeframe generating 7 by 8 (total of 56) SOM identifying representative weather patterns that are going to be used to produce 56 individual stepwise weighted quadratic regressions. The 56 patterns are designated using both data fields 500 hPa geopotential height (Figure S1) and 850 hPa temperature (Figure S2).

The next step is to identify O_3 MDA8 values corresponding to the classified weather patterns from each of the five Philadelphia area AQ monitors. An MDA8 is computed by scanning averaged 8-hour segments of the 24-hour period and identifying the segment with the highest value. Figure S3 shows O_3 MDA8 values for each of the 56 regimes at the NEA monitoring site. For each regime a stepwise weighted quadratic regression is constructed where O_3 MDA8 values are predictands and variables of surface temperature, relative humidity, wind speed, wind direction, solar zenith angle, and previous day O_3 MDA8 are predictors. The process of generating such regressions is described in detail in Balashov et al. (2017).

Finally, to produce a probabilistic O_3 MDA8 forecast for Philadelphia area predicted fields of 500 hPa geopotential heights and 850 hPa temperatures are required along with forecasted predictors (representative of Philadelphia area) needed to run regressions mentioned above. Once the spatial fields of 500 hPa geopotential heights and 850 hPa temperatures are provided for a day of interest (Figures S4a and S4b) the most relevant patterns are identified from the previously classified 56 regimes. The more similar a pattern is to the forecasted fields the more weight it gets and so the weights are distributed for each of the 56 patterns. Then solutions of regressions are combined and weighted in such a way that produces a probabilistic forecast depicted in Figure S4d.

Figure S5 illustrates daily REGiS probabilistic forecasts of exceeding O_3 MDA8 assuming EPA exeedance standard of 70 ppbv at the five O_3 monitoring stations in the Philadelphia area for MJJAS 2018. The station with the highest probability of an exceedance is chosen as a representative one for a given day.

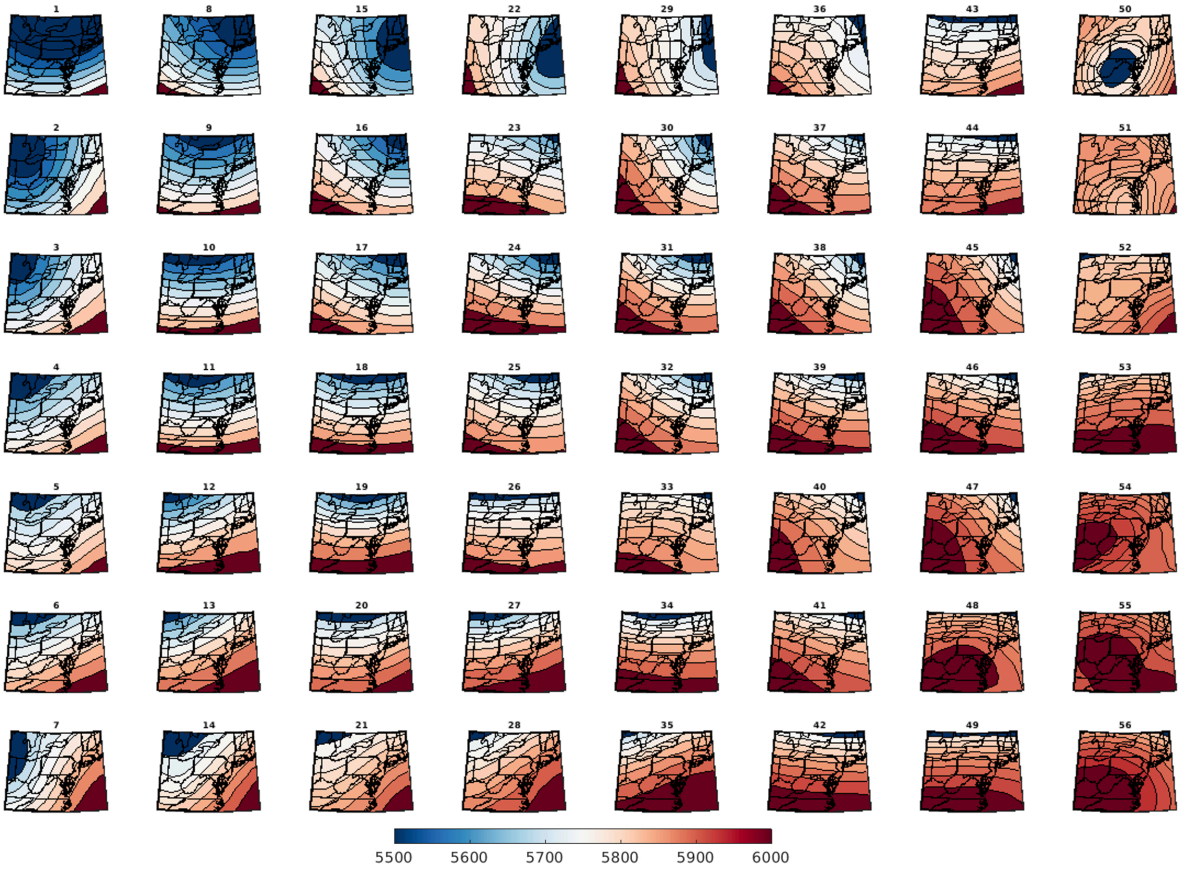


Figure S1. The 7 by 8 SOM of daily 500 hPa geopotential heights (in meters) from MJJAS over 2000-2011 generated for the northeastern part of the US. The daily fields of 850 hPa temperatures are also used in this SOM training process.

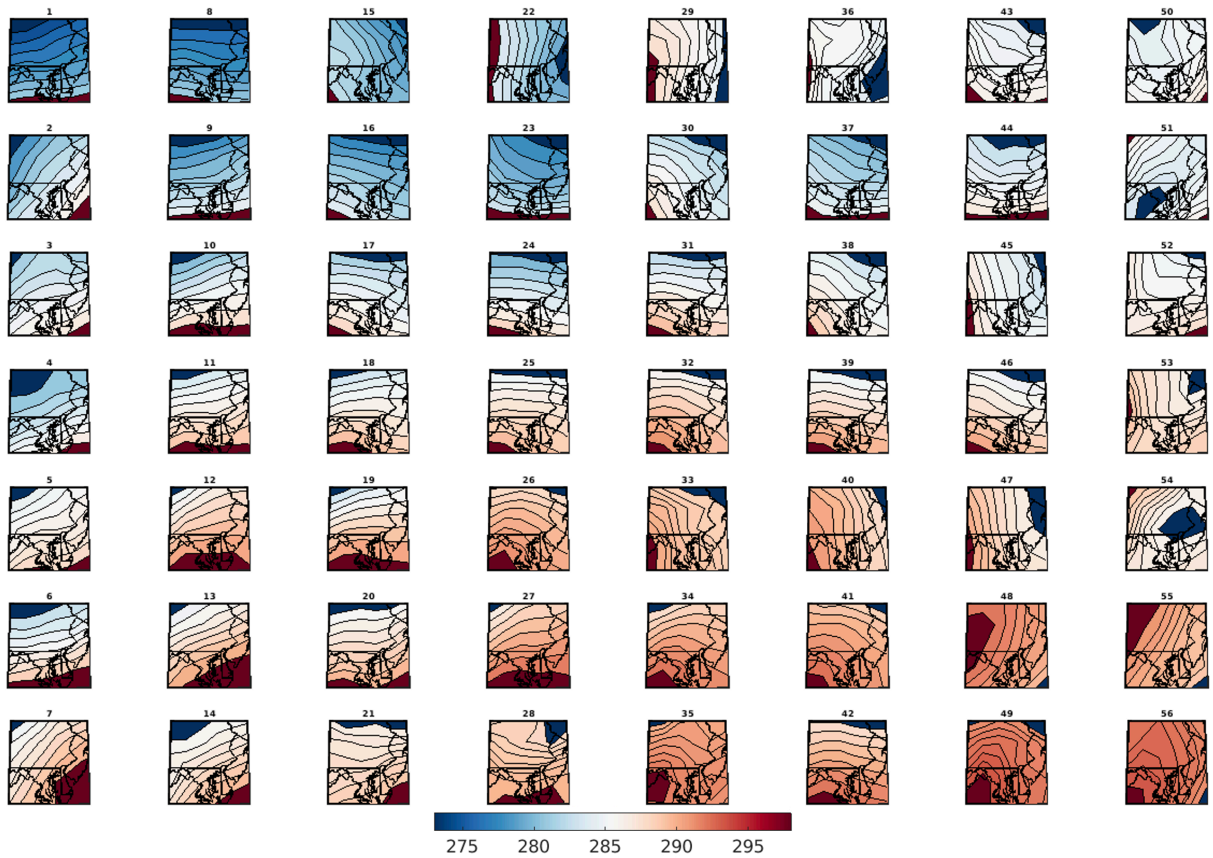


Figure S2. The 7 by 8 SOM of daily 850 hPa temperature fields (in K) from MJJAS over 2000-2011 generated for the northeastern part of the US. The daily 850 hPa temperature fields are also used in this SOM training process.

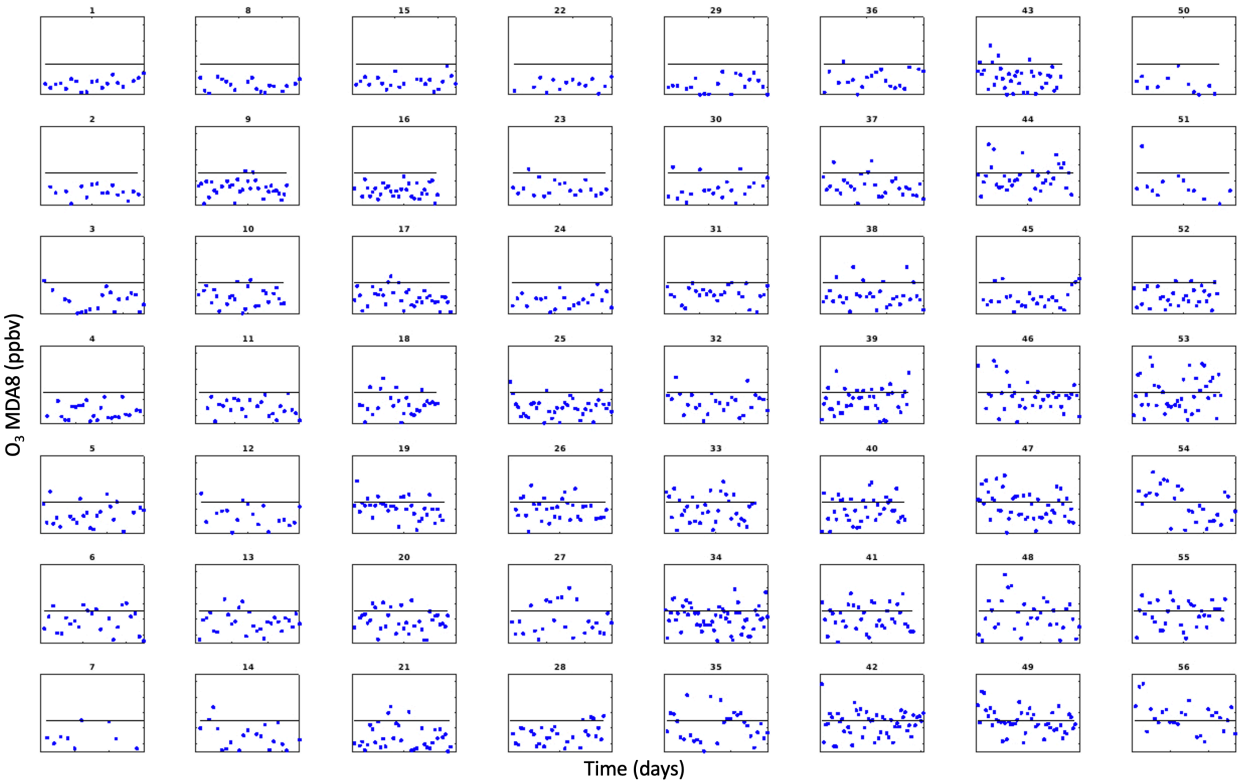


Figure S3. O_3 MDA8 from the NEA station corresponding to the 7 by 8 SOM in Figures S1 and S2 during MJJAS 2000–2011. Y and X ticks along with their labels are removed for clarity, Y limits range from 30 to 130 ppbv and X axis represents occurrence of the MDA8 values in time (days) as they relate to the weather regimes shown in Figures S1 and S2. The horizontal line indicates EPA exceedance O_3 MDA8 standard of 70 ppbv.

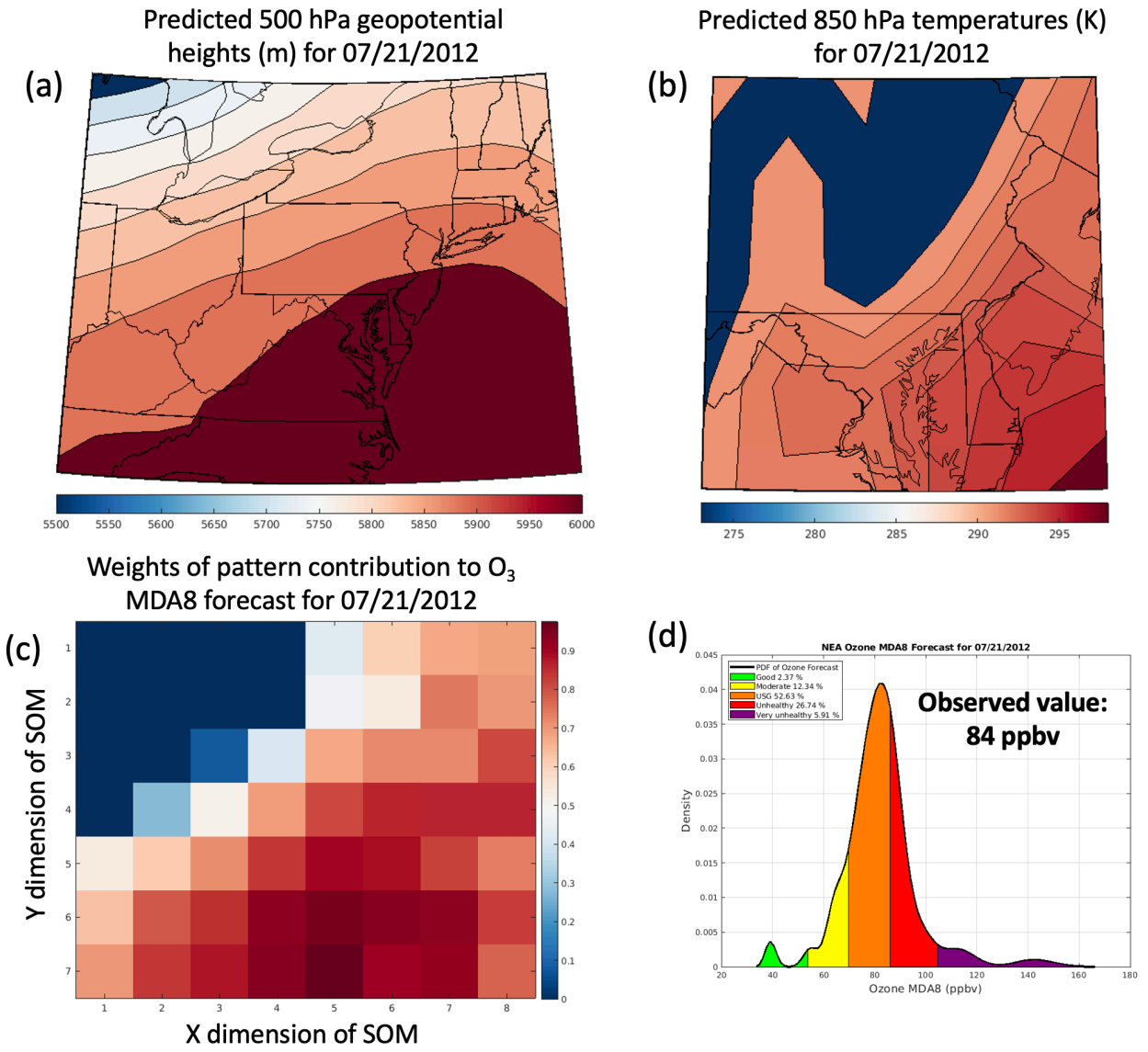


Figure S4. An example of how a probabilistic forecast is generated by REGiS. First, predicted fields of 500 hPa geopotential heights (in meters) and 850 hPa temperatures (in Kelvin) are provided as shown in panels **(a)** and **(b)**. Then these fields are compared to the 56 weather regimes shown in Figures S1 and S2 generating a distribution of weights **(c)** that indicate which of the 56 weather regimes are most similar to the forecasted weather pattern presented in **(a)** and **(b)**. Finally, a probabilistic forecast is generated **(d)** by combining forecasts from weighted regressions (weighted as shown in **(c)**) that are based on O_3 MDA8 values illustrated in Figure S3. This example forecast is for the Philadelphia area valid for 07/21/2012.

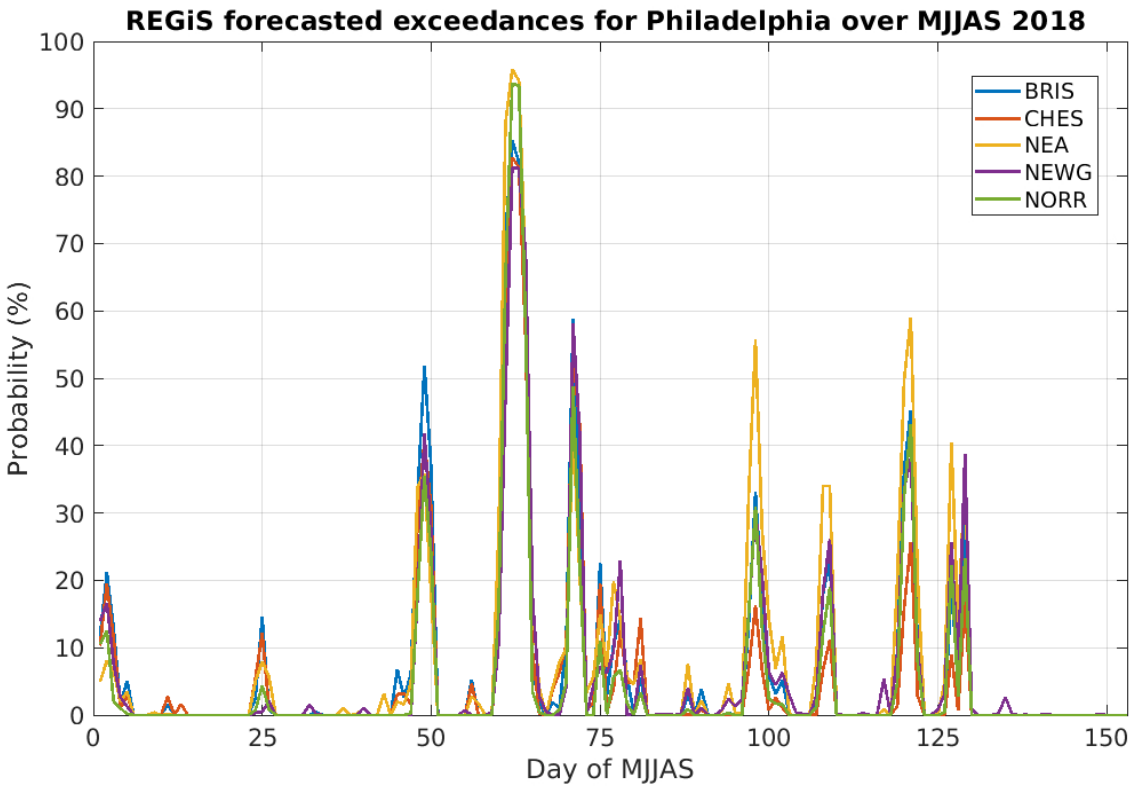


Figure S5. Probabilities of an exceedance at 5 of the monitoring stations in Philadelphia area for MJJAS 2018 generated by REGiS, where an EPA exceedance O_3 MDA8 standard is 70 ppbv.

Reference:

Balashov, N.V., Thompson, A.M. and Young, G.S., 2017. Probabilistic forecasting of surface ozone with a novel statistical approach. Journal of Applied Meteorology and Climatology, 56(2), pp.297-316.