Supporting Information

XIS-$PM_{2.5}$: A daily spatiotemporal machine-learning model for $PM_{2.5}$ in the contiguous United States

Allan C. Just[1,2,3*], Kodi B. Arfer[1,3], Johnathan Rush[3], Alexei Lyapustin[4], Itai Kloog[3,5]

[1]Department of Epidemiology, Brown University School of Public Health, Providence, RI, USA
[2]Institute at Brown for Environment and Society, Brown University, Providence, RI, USA
[3]Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[4]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[5]The Department of Geography and Environmental Development, Ben-Gurion University of the Negev, Beer Sheva, Israel

Corresponding Author: allan_just@brown.edu
Address: Allan Just, Box G-S121-2, Providence, RI 02912 USA

Cross-validation results stratified by meteorological season and for isolated sites by year, as well as scatterplots showing the association of relative topography and isolation with a summary of the contribution to predicted $PM_{2.5}$

Table S1: 2019 cross-validation results ($\mu g/m^3$) broken down by meteorological season. Results for December are taken from the 2018 model so that a contiguous winter is analyzed.

Season	Observations	Sites	MAD	MAE	Bias
all	350,051	1,284	3.09	1.69	-0.19
Spring	88,942	1,257	2.92	1.52	-0.20
Summer	88,946	1,254	2.80	1.48	-0.13
Fall	87,022	1,246	3.06	1.65	-0.19
Winter	85,141	1,284	3.55	2.12	-0.25

Table S1 shows cross-validation results by season for one year. Compared to the whole one-year period, MAD and MAE are lower in spring and summer and higher in winter.

Table S2: Results from each yearly cross-validation among isolated sites, in $\mu g/m^3$.

Year	Observations	Sites	MAD	MAE	Bias
2003	25,913	236	4.67	1.92	-0.03
2004	29,202	231	4.43	2.01	0.10
2005	30,520	225	5.02	2.09	-0.00
2006	34,457	233	4.38	2.07	0.00
2007	37,396	225	4.78	2.21	-0.08
2008	38,129	226	4.24	2.04	-0.20
2009	40,172	223	3.83	1.94	-0.19
2010	44,479	230	3.98	1.98	-0.06
2011	45,310	223	4.20	2.22	-0.26
2012	47,121	222	3.86	2.17	-0.28
2013	49,856	224	3.71	2.12	-0.16
2014	54,379	238	3.59	2.05	-0.27
2015	55,638	242	3.58	2.04	-0.17
2016	57,603	241	3.03	1.86	-0.18
2017	60,450	239	3.49	1.99	-0.15
2018	62,863	239	3.54	1.91	-0.26
2019	63,539	242	3.00	1.68	-0.18
2020	66,022	239	3.39	1.91	-0.21
2021	67,509	243	3.86	1.95	-0.27
2022	69,192	247	3.04	1.71	-0.23
2023	71,329	251	3.97	1.92	-0.18

In addition to the weighted analyses using all stations, we wished to evaluate performance where ground networks were especially sparse. Thus, Table S2 shows unweighted MAE from cross-validation among the sites that were particularly isolated, defined as being more than 50 km from all other sites available in the same year.

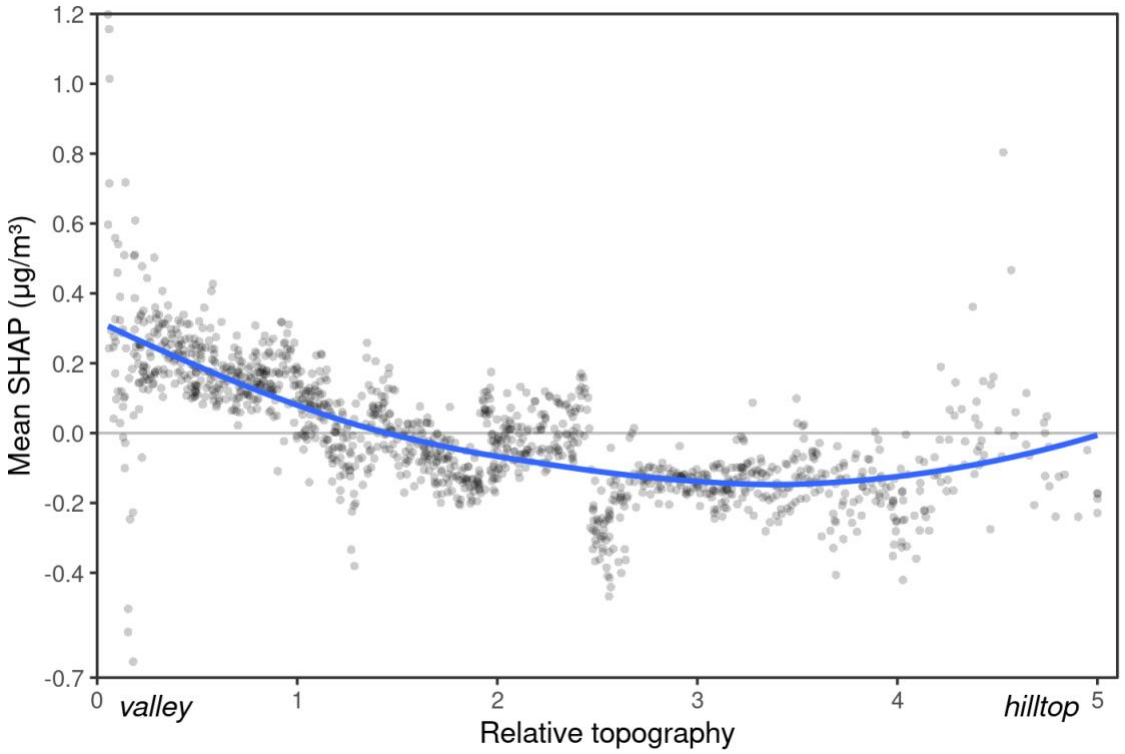

Figure S1: SHAP of hilliness as a function of hilliness.

Figure S1, similar to Figure 3, plots the mean SHAP of the hilliness feature for each site. We see higher average predicted $PM_{2.5}$ at sites in a valley versus on a hill.

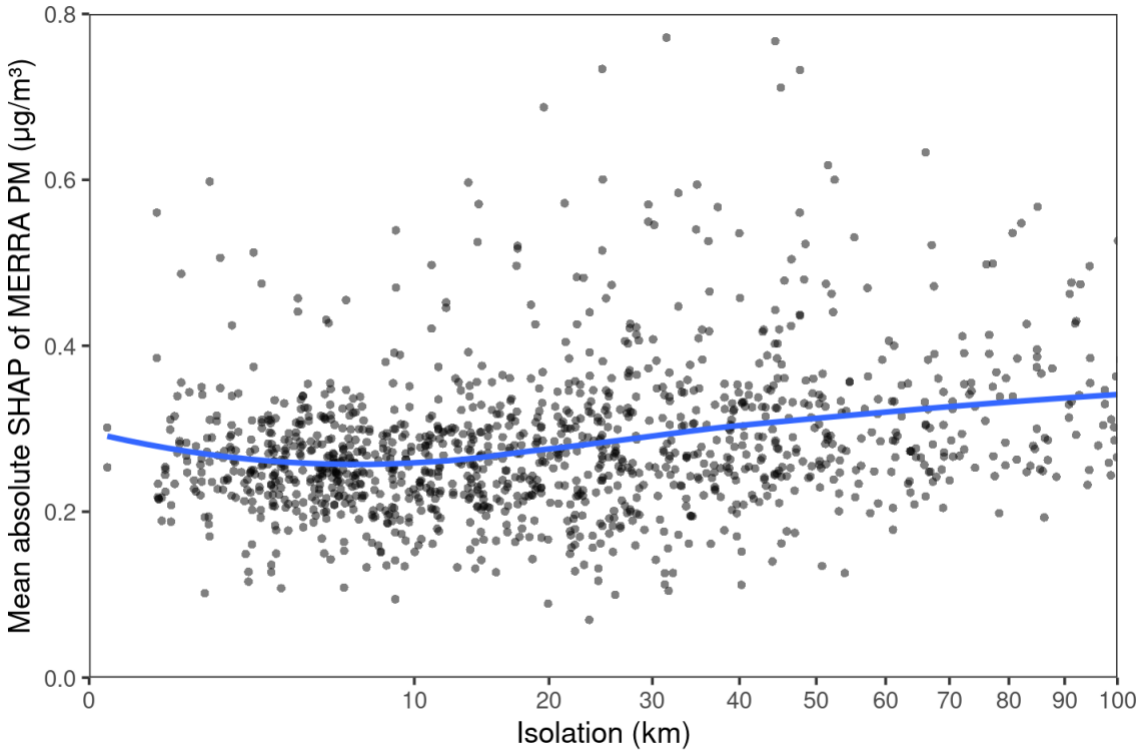

Figure S2: SHAP of modeled surface $PM_{2.5}$ concentrations from MERRA-2 as a function of site isolation. The *x*-axis is on a square-root scale.

Figure S2 is an example of the relationship between the SHAP of one variable and a different quantity. The *y*-axis shows the per-site mean absolute SHAP of MERRA-2 modeled $PM_{2.5}$, but the *x*-axis shows the site's distance from its nearest neighbor; that is, its degree of isolation. Similarly, we examined how the SHAP of the IDW feature varied according to isolation. The per-site mean absolute SHAP for IDW in 2010 was Kendall-correlated -0.17 with the distance to the nearest other site, meaning that the IDW is less influential on predictions for more isolated sites.