

Overview of smoke forecast map

This smoke forecast tool provides a five-day prediction of the amount of smoke that can be expected in 53 zones around the state, from wildfires and other sources. Forecasts can be used to plan outdoor activities, and reduce exposure to air pollution. There may be areas within each zone that experience higher or lower smoke concentrations throughout the day due to various reasons. Health warnings linked to each air quality category [can be found here](#).

The smoke forecast is based on:

- Forecasts issued by local clean air agencies.
- Projections by a smoke forecast model run by Washington State University.
- Several global smoke forecast models.
- Ecology smoke models.
- State and local agency forecasters' best professional judgement.
- Note: No forecasts are made for areas without monitors, or when monitors are down.

The forecast zones shown on the map are built around Ecology's statewide network of air quality monitors, and the air quality management regions of some local air agencies. Ecology analyzed past measured and modeled data to estimate smoke levels for a zone around each monitor, based on terrain, weather patterns and other information. The size of these zones is different during wildfire smoke season and wintertime home heating season.

Some local air agencies issue forecasts from 1 to 5 days. These are not altered in any way. The WSU forecast model uses the previous day's particulate pollution level, along with [forecast weather conditions from a University of Washington computer model](#), to predict smoke for the next two days.

Forecasting smoke for longer durations is very challenging since wildfires and other sources of smoke emissions often change quickly. New fires may start, existing fires can grow or be brought under control by firefighters. Smoke plume behavior can also change significantly. To forecast smoke on days 3 through 5, we use multiple models and calculate their median. Each model makes different assumptions about sources of smoke emissions. Only two-day forecasts are provided during winter months.

For more information

Appendix A and Appendix B on the following pages provide technical descriptions of these procedures. Appendix C provides an evaluation on the accuracy of forecasts during the 2020 wildfire season. The evaluation was not repeated for the 2020- 2021 winter since air quality was "Good" or "Moderate" 99.5% of the time between November and March, and the forecast closely followed that.

Still have questions?

Email: airemissions@ecy.wa.gov

Phone: 360-407-7658

Appendix A: Technical description of smoke forecast tool development

Local air agency (LAA) forecasts

Four of the seven LAAs in Washington submit daily forecasts to EPA's AirNow system. These forecasts range from 1 to 5 days in length. Daily forecasts are not issued in areas within the jurisdiction of Ecology and the other three LAAs. Further, AirNow simply interpolates all available point forecasts using inverse distance weighting. This technique does not respect natural airshed boundaries or differing pollutant emissions in areas without forecasts.

LAA- issued point forecasts are retrieved several times a day from AirNow. PSCAA and SRCAA provided Ecology with maps (see Figure A.1) and shapefiles of their air quality management zones, so their forecasts are assigned to these areas. Where the LAA boundaries intersect the polygons described in Appendix B (hereafter referred to as "ECY polygons"), best professional judgement was used to smooth out discrepancies and preserve the LAA boundaries. ORCAA and NWCAA have not yet developed such maps, and requested Ecology to assign their forecasts to ECY polygons.

Two-day forecasts from WSU Machine Learning (ML) model

Following the successful implementation of a Machine Learning model to [forecast ozone in the Tri Cities](#), the same approach was pursued to [build a more reliable forecast product](#) for all PM_{2.5} and ozone monitoring sites in Washington, Oregon and Idaho. Site-specific O₃ and PM_{2.5} versus meteorology relationships were developed using 4-kilometer WRF forecasts archived since 2017. These learned relationships describe the how the airshed's PM_{2.5} (or ozone) levels have responded in recent years under different meteorological conditions. The relationships are then applied to the forecast ensemble mean meteorological parameters from UW, to construct a two day forecast. No forecasts can be made if the monitor didn't report data the previous day. More details [can be found here](#).

The Ecology polygons outside the local air agency polygons in Figure A.1 are used to map these PM_{2.5} forecasts.

Extending the summer wildfire smoke forecast up to 5 days

Because of the inherent uncertainty of forecasting wildfire smoke for longer than 2 days, an ensemble modeling approach was pursued to account for emissions uncertainty. We calculated the median of a dozen different simulations. These consist of four global smoke models and eight HYSPLIT ensemble members simulated at Ecology.

Global, long range smoke models

We selected four global models which account for emissions from all sources and atmospheric chemistry, and simulate surface level PM_{2.5} concentrations at a horizontal resolution no coarser than 0.5 degrees x 0.5 degrees. The gridded model results must be available early in the morning PDT, and cover the next 5 days. The models are:

- [ECMWF CAMS](#) (same day's 0Z run, 0.4 degree resolution)
- [NASA GEOS-CF](#) (previous day's 12Z run, 0.25 degree resolution).
- [NOAA GEFS-Aerosol](#) (same day's 6Z run, 0.25 degree resolution).
- [Finnish SILAM](#) (same day's 0Z run, 0.5 degree resolution).

The ECMWF-CAMS and NOAA GEFS- Aerosol models are often the data sources behind smoke forecasts offered by many websites.

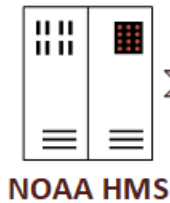
The model forecasts for days 3-5 are bilinearly interpolated to a common 0.25 degree grid. The concentrations for each day at each PM_{2.5} monitor in Washington is extracted. These data are not subjected to the bias correction procedure that is applied to HYSPLIT results.

Ecology's HYSPLIT modeling scenarios

Multiple simulations of wildfire smoke for the next 5 days are performed by running the NOAA [HYSPLIT](#) model at Ecology. Meteorology is obtained from the UW WRF 12km extended forecast. Eight different HYSPLIT ensemble members are simulated by assuming four different emissions scenarios and two vertical plume velocity parameterizations. See the HYSPLIT Ensemble Forecast Emissions Scenario figure for more information.

HYSPLIT does not include emissions from other anthropogenic or natural sources, nor does it consider any chemical processes in the atmosphere. The WSU ML model implicitly accounts for these and is thought to offer a better forecast for the first two days. Therefore a bias correction is derived by comparing the first two days of each of the 8 HYSPLIT simulations, to the WSU ML forecast at the same location. This simple additive bias correction constrains resulting concentrations between 5 and 500 µg/m³. After correcting each of the HYSPLIT estimates, they and the coincident global model estimates are aggregated to produce a monitor- specific ensemble median. This is considered our best estimate of expected smoke levels after reasonably attempting to account for uncertainties associated with long range smoke forecasting.

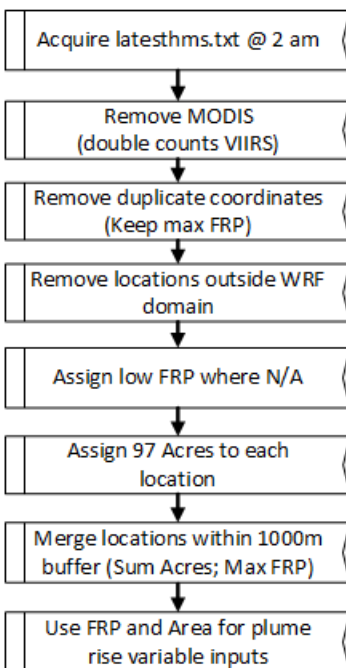
HYSPLIT Ensemble Forecast Emissions Scenarios



FRP Emissions Rate (g. PM2.5/hr) =
 $0.003 \text{ (ton veg./MJ)} * 3600 \text{ (sec/hr)} * 6123.6 \text{ (g. PM2.5/ton veg.)}$
66,134 g PM2.5 / HMS FRP MW
 (HMS FRP is constrained from 0.3 to 100 MW)

Area Emissions Rate =
300,000 g PM2.5 / HMS Detect
 (Equivalent to a 4.5 MW FRP Detect)

HMS Preprocessing:



Scenario Details:

FRP = Fire Radiative Power DF = Diurnal Factor
 FPF = Fire Potential Factor LCF = Land Cover Factor

Scenario 1 Emissions = FRP Rate x DF
Scenario 2 Emissions = FRP Rate x DF x FPF x LCF
Scenario 3 Emissions = FRP Rate x DF x FPF x LCF x 1.25
Scenario 4 Emissions = Area Rate x DF x FPF x LCF

Fire Scaling Factors:

Fire Potential

USA → US NIFC 7-Day Fire Potential Forecast (Day Specific)

- Moist = 0.25
- Dry = 1.0
- Very Dry = 1.25
- Windy = 1.25
- Hot = 1.25
- Lightning = 1.0
- Burn Env. = 1.25
- Hot & Dry = 1.5

CAN → Canada Fire Danger Rating (Most Recent)

- Low = 0.25
- Moderate = 0.5
- High = 0.75
- Very High = 1.0
- Extreme = 1.25

Land Cover

Global LCLU

- Non-Veg = 0.01
- Shrub/Grass = 0.4
- Crops = 0.2
- Forested = 1.0

Diurnal Factors:

Emissions Rate Factor

Hour (PST)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Rate	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.3	0.6	1	1.5	1.9	2.3	2.4	1.7	1	0.6	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Plume Rise Forcing Factor

Hour (PST)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Heat	0	0	0	0	0	0	0	0.5	1	1	10	20	20	20	10	1	0.5	0	0	0	0	0	0	0

It is important to note that if a PM_{2.5} monitor did not operate the previous day, the WSU ML model cannot produce a forecast for days 1-2. In that case, or if the WSU model failed to run, the previous day's measured PM_{2.5} concentration (if available) is replicated as a persistence forecast for day 1. The day 3 forecast issued the previous day is slid over to today's day 2 forecast. These two days form the basis of the bias correction for new HYSPLIT results. However this backup procedure is not implemented if the monitor was down for more than one day. No forecasts are produced in such instances.

If forecasts for days 3-5 are missing from the local air agency's AirNow forecasts, the ensemble median is substituted there. However no substitutions are made if days 1-2 are missing in AirNow.

Manual forecasts by Ecology

This option is used to override the WSU ML forecast only if a compelling reason exists, such as a large new fire that isn't handled well by the forecast model. No changes to LAA forecasts are made.

Submission to AirNow

Forecasts produced for Ecology's jurisdictional areas are uploaded to EPA's AirNow and EnviroFlash systems. At the request of BCAA and SWCAA, those forecasts are also included in this upload.

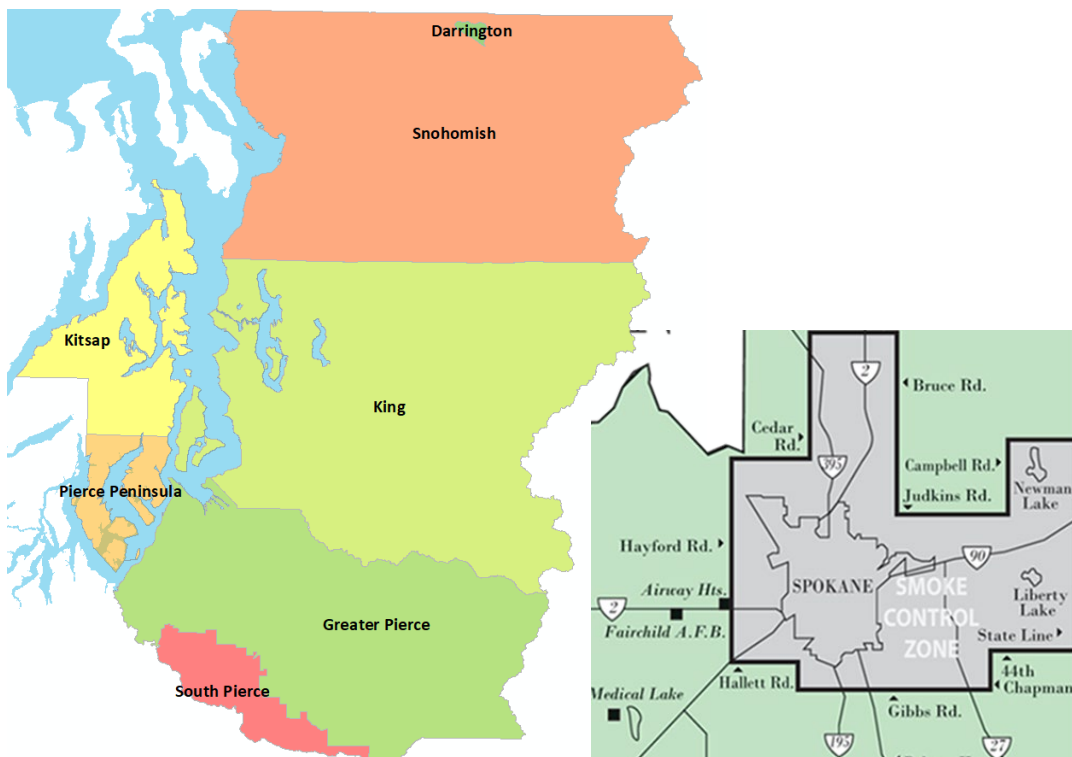


Figure A.1: PSCAA (left) and SRCAA (bottom right) air quality management zones.

Appendix B: Zones Attributed to Monitors

Questions? Contact Jill Schulte, WA Ecology Air Quality Program, jils461@ecy.wa.gov, 360-407-6877

Summary

The zones attributed to each monitor are nearby areas that are estimated to have similar smoke conditions based on maps of interpolated PM_{2.5} across Washington during the summer season (June – Sept).

Where local air agencies provide forecasts according to their own geographical units, these units were used instead of the process described below. Zones selected by local air agencies are located in the Spokane and Puget Sound areas.

For the remainder of the state, summer forecast zones show areas with similar interpolated PM_{2.5} concentrations during the years 2015-2020. These zones represent Ecology's best understanding of the area represented by each monitoring site during warmer months when smoke from wildland fires is likely to be the dominant source of PM_{2.5}.

Summer 98th percentile PM_{2.5} was interpolated at 4km resolution for each year (2015-2020) using output from the AIRPACT-4 forecast model and data collected at agency monitoring sites.

Input datasets

1. Summer (June – Sept) 98th percentile daily PM_{2.5} concentrations at monitoring sites by year from 2015-2020.
2. Summer (June – Sept) mean modeled PM_{2.5} at 4km grid cells from AIRPACT-4 from July 2014-June 2017. This is the most recent continuous 3-year period available with minimal wildfire influence, which was chosen due to inconsistent model performance during wildfire periods. The summer mean modeled PM_{2.5} was used to represent the typical spatial variation in PM_{2.5} concentrations in unmonitored areas.

PM_{2.5} interpolation

For each year (2015-2020), summer 98th percentile PM_{2.5} concentrations were interpolated at 4km resolution using [Empirical Bayesian Kriging Regression Prediction](#) in ArcGIS Pro 2.6.3. The summer 98th percentile PM_{2.5} concentration at each monitoring site was treated as the dependent variable, and the modeled summer means at each grid cell were treated as the independent or explanatory variable. This process yielded maps that show the upper end of the distribution of PM_{2.5} concentrations at each 4km grid cell in each year, which are reflective of spatial patterns in wildfire smoke impacts. Using monitored values as the dependent variable ensures that the resulting spatial layer accurately reflects monitored concentrations, while using modeled means as inputs ensures that the results reflect the influence of topography, meteorology and emissions in areas without monitoring sites.

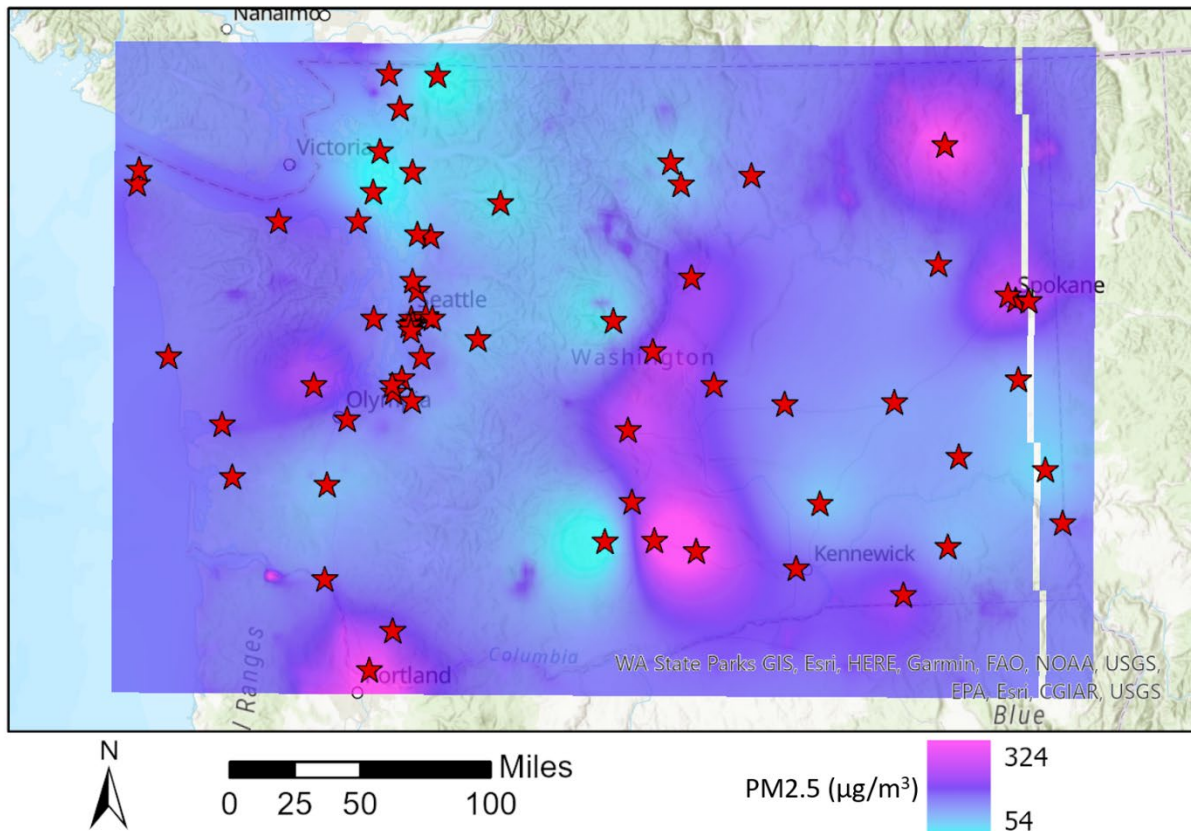


Figure B.1. Interpolated 98th percentile PM_{2.5} in 2020

Forecast zone selection

Forecast zones are based on clusters of grid cells with similar 98th percentile PM_{2.5} concentrations in multiple years. These represent areas that typically experience similar smoke impacts and can be represented by a single monitor.

Grid cells were aggregated into clusters using the [Spatially Constrained Multivariate Clustering](#) (SCMC) tool in ArcGIS Pro 2.6.3 with the six interpolated 98th percentile PM_{2.5} concentration layers (2015-2020) as the input features. This tool identifies distinct clusters of contiguous grid cells based on similarities in the input data. The tool selects clusters to minimize variability in PM_{2.5} concentrations within clusters and maximize variability between clusters.

First, the SCMC tool was applied statewide with 100 iterations, using the highest calculated pseudo-F statistic to select the optimum number of clusters. Three of the resulting ten clusters did not contain a monitoring site and therefore are not considered represented by a monitor: northeast Washington between Colville and Omak, the Columbia River Gorge east of the Tri-Cities, and the high-elevation band of the Cascade Range between the Columbia River and the Canadian border.

In order to further refine the forecast zones, the remaining areas were split into regions: southwest (the Olympic peninsula and all areas between the Cascade Range and the Pacific Ocean south of Pierce County), northwest (Skagit, Whatcom, Island and San Juan Counties), and central/eastern (all areas east of the Cascade Range). The SCMC process was repeated within each region with the number of clusters specified to match the number of monitoring sites within that region.

Within these regions, clusters that contained only one monitoring site were assigned to that site. Clusters that contained more than one monitoring site were split by distance to those sites, and areas were assigned to their closest monitoring site within that cluster. Within-region clusters that did not contain monitoring sites were split according to topography and distance to the nearest sites, using best professional judgement to assign areas to nearby sites.

Some additional refinements were made to reduce the size of forecast zones in areas lacking sufficient monitoring data. These areas include the Olympic Range and portions of the Cascade Range west of Leavenworth and Winthrop, which are considered unmonitored.

The final forecast zone boundaries were smoothed using the [Polynomial Approximation with Exponential Kernel](#) method.

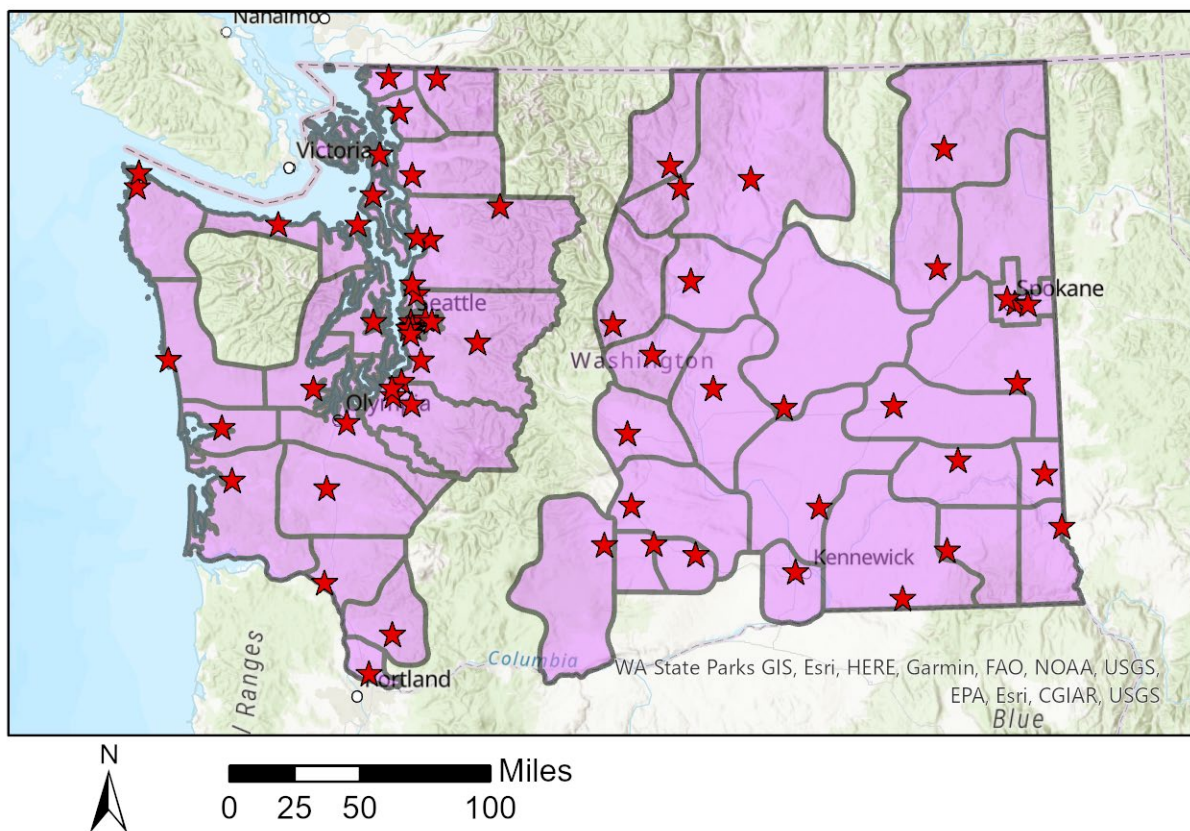


Figure B.2. Summer smoke forecast zones

Appendix C: Evaluation of model performance in summer 2020

The evaluation was limited to August 24- October 8 2020, since several monitors recorded Moderate or worse air across Washington State (Figure C.1).

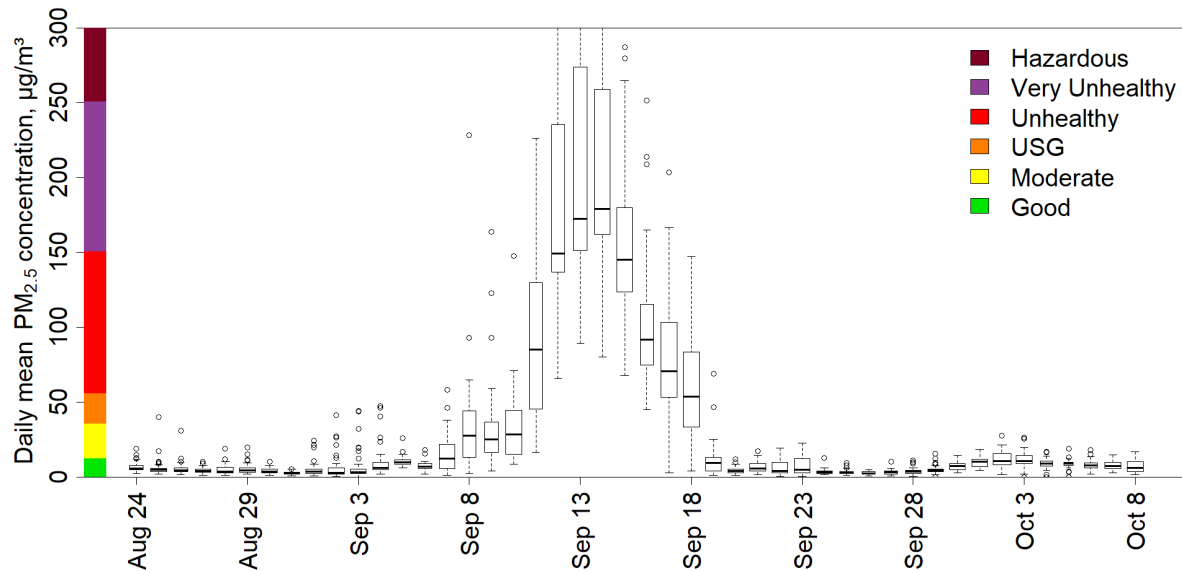


Figure C.1: Boxplot of statewide PM_{2.5} concentrations by day

Figure C.2 shows forecast accuracy at different air quality levels. Most forecasts on days 1 and 2 proved correct. A larger fraction of the poor predictions (i.e. off by more than 1 category) occur at higher concentrations, seen at the top right area of the plot, underscoring the challenges that ML models and human forecasters face with extreme pollution levels.

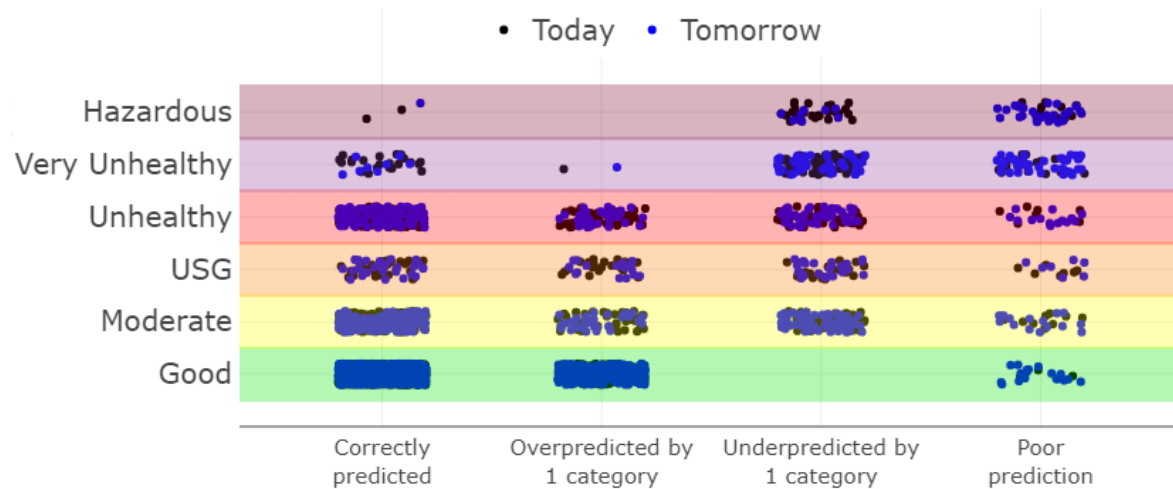


Figure C.2: Forecast accuracy for each measured air quality category

Figure C.3 aggregates all day 1 and 2 forecasts and shows the fraction of time forecasts fell into each of the performance categories.

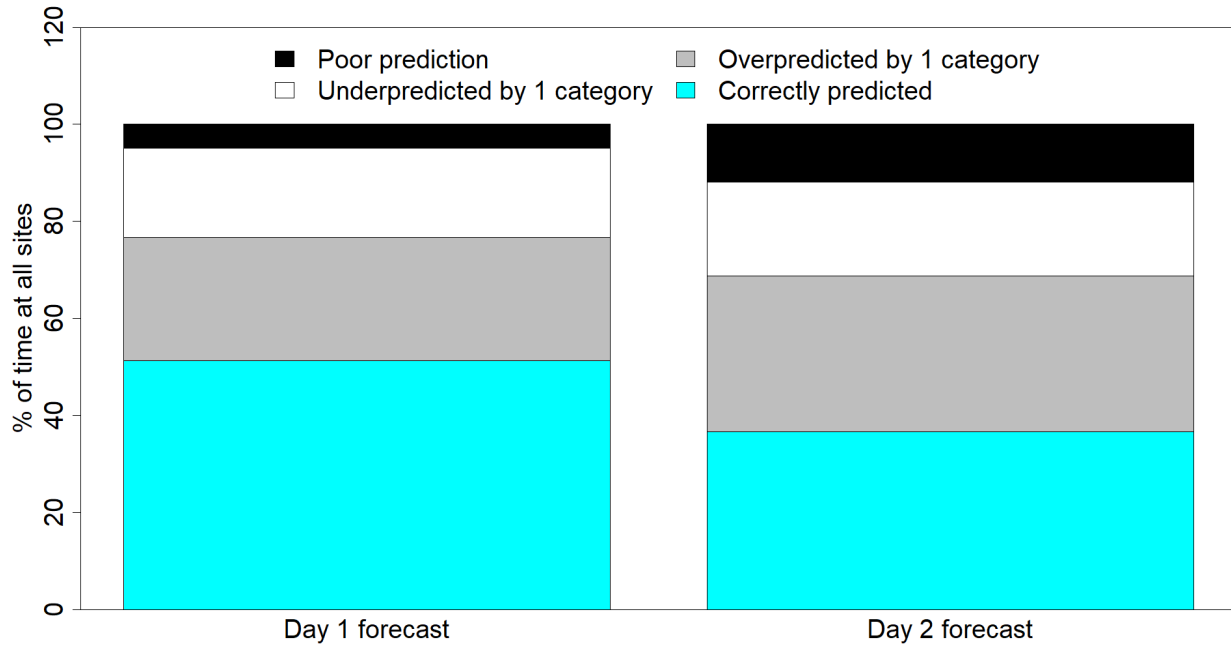


Figure C.3: Smoke forecast accuracy at all sites. Omitted days when Good air was forecast and measured

Disregarding days when “Good” air was accurately forecast:

- Day 1 forecasts at most sites, whether produced by LAA forecasters or the ML model, appear reasonably accurate at least three quarters of the time.
- Part of the reason for lower forecast skill on day 1 is when the monitor did not operate (i.e. fewer days to compare against). It must be noted that if a valid 24-hr average measurement is not available for the previous day, the ML model cannot produce a forecast for that site.
- Day 2 forecast skill is slightly lower than day 1.
- On occasion, the ML model handled some instances of heavy smoke rather poorly and had to be manually overridden by Ecology forecasters. However this option was not exercised at sites where LAAs submitted forecasts to AirNow.
- At sites where both were present, the LAA- issued forecasts were usually more accurate than the ML model.
- Overall, the ML forecast performance was good or acceptable about two thirds of the time, when considering days 1 & 2. Poor predictions (i.e. more than one AQ category off) happened just under one fifth of the time.
- Overall, this is considered satisfactory forecast skill although improvements are possible.