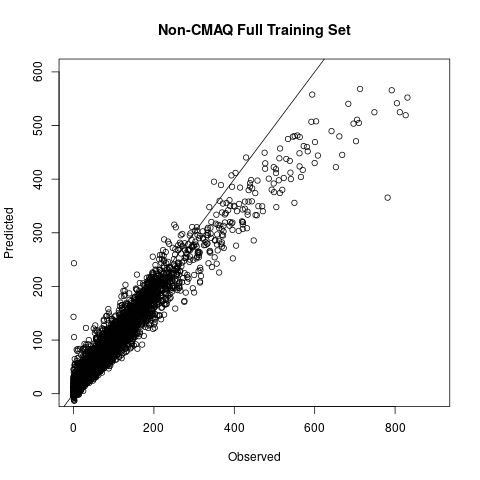
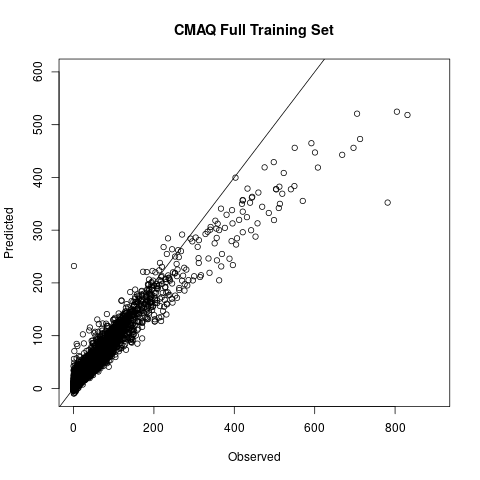
**Supplementary Material**



Variable Importance:

Non-CMAQ: CMAQ:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Importance** |  | **Variable** | **Importance** |
| NLCD\_5km | 100.00 |  | Population density | 100.00 |
| NLCD\_10km | 66.68 |  | Elevation | 54.63 |
| NLCD\_1km | 12.21 |  | Longitude | 47.74 |
| Elevation | 11.19 |  | NLCD\_10km | 43.22 |
| Latitude | 11.07 |  | NLCD\_5km | 42.40 |
| Population density | 10.24 |  | Surface pressure | 38.53 |
| Longitude | 7.97 |  | CMAQ | 35.86 |
| Temperature at 2m | 7.86 |  | Latitude | 33.27 |
| Surface pressure | 6.40 |  | NLCD\_1km | 28.68 |
| State: California | 6.21 |  | Temperature at 2m | 21.47 |
| Region: Southwest | 5.47 |  | State: California | 19.57 |
| MAIAC AOD | 4.59 |  | Cosine of day of year | 15.11 |
| Cosine of day of year | 3.81 |  | Region: Southwest | 10.28 |
| Dew point temp. at 2m | 3.16 |  | Date | 9.58 |
| Date | 2.98 |  | Dew point temp. at 2m | 9.46 |
| Cosine of month | 2.84 |  | Cosine of month | 8.13 |
| Relative humidity at 2m | 2.78 |  | Region: Northwest | 7.95 |
| Planetary boundary layer height from surface | 2.07 |  | Region: Southwest + study period 2008-2016 | 7.71 |
| Sum of arterial and collector road lengths within 1000m | 1.84 |  | Relative humidity at 2m | 7.31 |
| NDVI | 1.62 |  | Sum of arterial and collector road lengths within 1000m | 6.52 |
| Sum of arterial and collector road lengths within 500m | 1.49 |  | Planetary boundary layer height from surface | 6.47 |
| Pressure reduced to mean sea level | 1.40 |  | MAIAC AOD | 6.03 |
| Region: Southwest + study period 2008-2016 | 1.34 |  | NDVI | 5.80 |
| Year | 1.30 |  | Sum of arterial and collector road lengths within 500m | 5.52 |

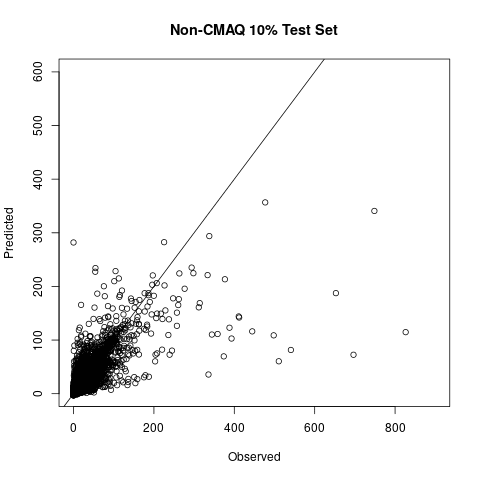
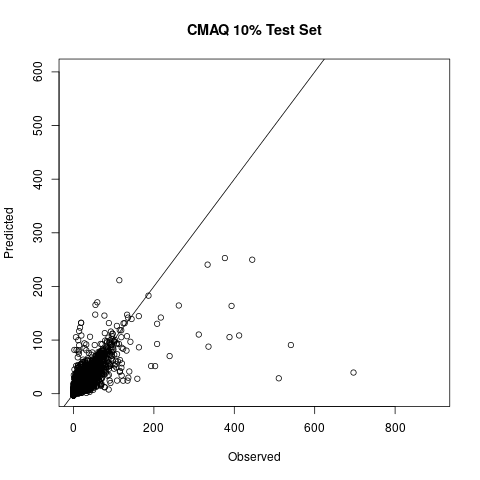
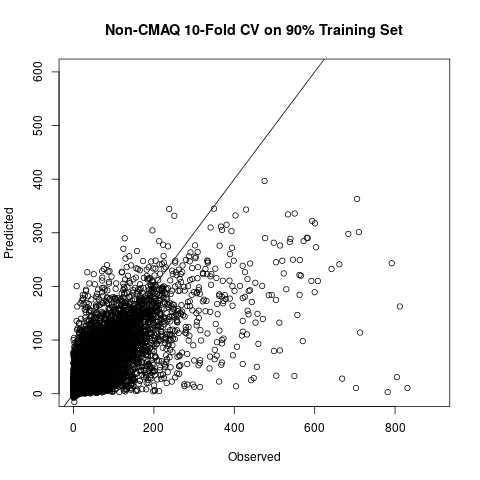
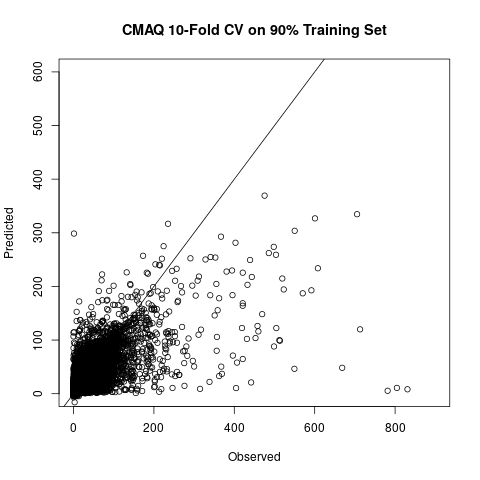
Spatial folds results:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **PM2.5 (µg/m3)** | **CMAQ Ensemble Training R2** | **CMAQ Ensemble Testing R2** | **Non-CMAQ Ensemble Training R2** | **Non-CMAQ Ensemble Testing R2** |
| Below 35 | 0.682 | 0.575 | 0.523 | 0.522 |
| Below 60 | 0.714 | 0.612 | 0.568 | 0.569 |
| Below 150 | 0.720 | 0.621 | 0.615 | 0.612 |
| Below 300 | 0.706 | 0.631 | 0.626 | 0.641 |
| Below 500 | 0.690 | 0.603 | 0.618 | 0.616 |
| Below 1000 | 0.659 | 0.589 | 0.598 | 0.593 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Year** | **CMAQ Ensemble Training R2** | **CMAQ Ensemble Testing R2** | **Non-CMAQ Ensemble Training R2** | **Non-CMAQ Ensemble Testing R2** |
| 2008 | 0.787 | 0.689 | 0.680 | 0.661 |
| 2009 | 0.744 | 0.637 | 0.614 | 0.602 |
| 2010 | 0.707 | 0.634 | 0.559 | 0.579 |
| 2011 | 0.723 | 0.642 | 0.598 | 0.615 |
| 2012 | 0.684 | 0.504 | 0.581 | 0.449 |
| 2013 | 0.683 | 0.588 | 0.562 | 0.576 |
| 2014 | 0.668 | 0.572 | 0.530 | 0.540 |
| 2015 | 0.667 | 0.687 | 0.571 | 0.645 |
| 2016 | 0.449 | 0.520 | 0.370 | 0.454 |
| 2017 | N/A | N/A | 0.684 | 0.669 |
| 2018 | N/A | N/A | 0.687 | 0.619 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **State** | **CMAQ Ensemble Training R2** | **CMAQ Ensemble Testing R2** | **Non-CMAQ Ensemble Training R2** | **Non-CMAQ Ensemble Testing R2** |
| Arizona | 0.614 | 0.374 | 0.469 | 0.418 |
| California | 0.719 | 0.756 | 0.632 | 0.724 |
| Colorado | 0.499 | 0.472 | 0.351 | 0.432 |
| Idaho | 0.663 | 0.546 | 0.605 | 0.607 |
| Montana | 0.648 | 0.593 | 0.636 | 0.595 |
| Nevada | 0.679 | 0.634 | 0.497 | 0.568 |
| New Mexico | 0.536 | 0.028 | 0.388 | 0.006 |
| Oregon | 0.597 | 0.362 | 0.612 | 0.436 |
| Utah | 0.638 | 0.647 | 0.495 | 0.665 |
| Washington | 0.583 | 0.586 | 0.529 | 0.515 |
| Wyoming | 0.516 | 0.531 | 0.418 | 0.465 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Season** | **CMAQ Ensemble Training R2** | **CMAQ Ensemble Testing R2** | **Non-CMAQ Ensemble Training R2** | **Non-CMAQ Ensemble Testing R2** |
| Fall | 0.598 | 0.520 | 0.599 | 0.561 |
| Spring | 0.611 | 0.561 | 0.444 | 0.543 |
| Summer | 0.635 | 0.571 | 0.605 | 0.591 |
| Winter | 0.767 | 0.658 | 0.639 | 0.650 |



Random folds results:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PM2.5 (µg/m3)** | **CMAQ Ensemble Training RMSE (µg/m3)** | **CMAQ Ensemble Training Spatial RMSE (µg/m3)** | **CMAQ Ensemble Training Temporal RMSE (µg/m3)** | **Non-CMAQ Ensemble Training RMSE (µg/m3)** | **Non-CMAQ Ensemble Training Spatial RMSE (µg/m3)** | **Non-CMAQ Ensemble Training Temporal RMSE (µg/m3)** |
| Below 35 | 2.792 | 3.046 | 2.601 | 3.165 | 3.870 | 2.918 |
| Below 60 | 3.071 | 3.527 | 2.851 | 3.548 | 4.758 | 3.238 |
| Below 150 | 3.407 | 4.947 | 3.076 | 4.043 | 6.659 | 3.569 |
| Below 300 | 3.725 | 6.888 | 3.208 | 4.509 | 8.820 | 3.803 |
| Below 500 | 4.058 | 8.752 | 3.313 | 4.945 | 10.913 | 3.970 |
| Below 1000 | 4.482 | 10.641 | 3.459 | 5.482 | 13.637 | 4.122 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Year** | **CMAQ Ensemble Training RMSE (µg/m3)** | **CMAQ Ensemble Training Spatial RMSE (µg/m3)** | **CMAQ Ensemble Training Temporal RMSE (µg/m3)** | **Non-CMAQ Ensemble Training RMSE (µg/m3)** | **Non-CMAQ Ensemble Training Spatial RMSE (µg/m3)** | **Non-CMAQ Ensemble Training Temporal RMSE (µg/m3)** |
| 2008 | 3.369 | 1.115 | 3.238 | 3.891 | 1.222 | 3.754 |
| 2009 | 3.206 | 0.983 | 3.094 | 3.678 | 1.180 | 3.553 |
| 2010 | 2.901 | 1.255 | 2.778 | 3.195 | 1.146 | 3.072 |
| 2011 | 3.377 | 2.838 | 3.171 | 3.665 | 2.640 | 3.442 |
| 2012 | 4.995 | 13.392 | 3.636 | 5.129 | 12.039 | 4.001 |
| 2013 | 4.398 | 8.985 | 3.626 | 4.606 | 9.316 | 3.823 |
| 2014 | 4.229 | 9.580 | 3.346 | 4.693 | 10.408 | 3.724 |
| 2015 | 5.275 | 12.185 | 3.883 | 5.643 | 11.660 | 4.320 |
| 2016 | 6.493 | 18.574 | 3.774 | 6.623 | 19.228 | 3.832 |
| 2017 | N/A | N/A | N/A | 8.243 | 21.555 | 5.618 |
| 2018 | N/A | N/A | N/A | 7.382 | 19.267 | 5.067 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **State** | **CMAQ Ensemble Training RMSE (µg/m3)** | **CMAQ Ensemble Training Spatial RMSE (µg/m3)** | **CMAQ Ensemble Training Temporal RMSE (µg/m3)** | **Non-CMAQ Ensemble Training RMSE (µg/m3)** | **Non-CMAQ Ensemble Training Spatial RMSE (µg/m3)** | **Non-CMAQ Ensemble Training Temporal RMSE (µg/m3)** |
| Arizona | 2.606 | 1.310 | 2.476 | 3.037 | 2.822 | 2.813 |
| California | 4.186 | 8.132 | 3.569 | 5.357 | 12.321 | 4.266 |
| Colorado | 5.003 | 10.662 | 2.696 | 5.393 | 11.899 | 2.876 |
| Idaho | 5.889 | 15.462 | 3.664 | 5.718 | 12.364 | 4.265 |
| Montana | 4.634 | 9.150 | 3.600 | 7.002 | 14.979 | 5.251 |
| Nevada | 3.154 | 1.209 | 3.043 | 3.212 | 1.192 | 3.099 |
| New Mexico | 3.089 | 5.966 | 2.425 | 3.109 | 5.267 | 2.582 |
| Oregon | 4.900 | 10.116 | 4.014 | 7.171 | 20.571 | 4.952 |
| Utah | 3.930 | 7.291 | 3.324 | 4.204 | 6.530 | 3.657 |
| Washington | 4.748 | 13.845 | 3.324 | 5.229 | 14.303 | 3.940 |
| Wyoming | 5.238 | 14.989 | 2.280 | 4.811 | 13.611 | 2.369 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Season** | **CMAQ Ensemble Training RMSE (µg/m3)** | **CMAQ Ensemble Training Spatial RMSE (µg/m3)** | **CMAQ Ensemble Training Temporal RMSE (µg/m3)** | **Non-CMAQ Ensemble Training RMSE (µg/m3)** | **Non-CMAQ Ensemble Training Spatial RMSE (µg/m3)** | **Non-CMAQ Ensemble Training Temporal RMSE (µg/m3)** |
| Fall | 5.478 | 8.237 | 3.507 | 6.917 | 10.921 | 4.395 |
| Spring | 2.577 | 1.726 | 2.258 | 2.770 | 2.060 | 2.401 |
| Summer | 5.191 | 9.095 | 3.032 | 6.635 | 10.898 | 4.102 |
| Winter | 3.994 | 2.618 | 3.487 | 4.436 | 2.649 | 3.911 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PM2.5 (µg/m3)** | **CMAQ Ensemble Training R2** | **CMAQ Ensemble Training Spatial R2** | **CMAQ Ensemble Training Temporal R2** | **Non-CMAQ Ensemble Training R2** | **Non-CMAQ Ensemble Training Spatial R2** | **Non-CMAQ Ensemble Training Temporal R2** |
| Below 35 | 0.779 | 0.648 | 0.550 | 0.729 | 0.589 | 0.663 |
| Below 60 | 0.800 | 0.679 | 0.601 | 0.753 | 0.640 | 0.703 |
| Below 150 | 0.798 | 0.659 | 0.615 | 0.767 | 0.683 | 0.732 |
| Below 300 | 0.781 | 0.631 | 0.559 | 0.762 | 0.684 | 0.735 |
| Below 500 | 0.762 | 0.617 | 0.550 | 0.745 | 0.678 | 0.724 |
| Below 1000 | 0.732 | 0.588 | 0.388 | 0.719 | 0.640 | 0.712 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Year** | **CMAQ Ensemble Training R2** | **CMAQ Ensemble Training Spatial R2** | **CMAQ Ensemble Training Temporal R2** | **Non-CMAQ Ensemble Training R2** | **Non-CMAQ Ensemble Training Spatial R2** | **Non-CMAQ Ensemble Training Temporal R2** |
| 2008 | 0.853 | 0.950 | 0.807 | 0.809 | 0.941 | 0.752 |
| 2009 | 0.816 | 0.954 | 0.749 | 0.767 | 0.932 | 0.689 |
| 2010 | 0.800 | 0.907 | 0.725 | 0.764 | 0.922 | 0.679 |
| 2011 | 0.802 | 0.734 | 0.745 | 0.772 | 0.757 | 0.712 |
| 2012 | 0.746 | 0.722 | 0.733 | 0.732 | 0.789 | 0.687 |
| 2013 | 0.754 | 0.551 | 0.751 | 0.733 | 0.554 | 0.727 |
| 2014 | 0.741 | 0.630 | 0.723 | 0.686 | 0.544 | 0.671 |
| 2015 | 0.733 | 0.693 | 0.726 | 0.697 | 0.674 | 0.677 |
| 2016 | 0.520 | 0.419 | 0.587 | 0.499 | 0.362 | 0.574 |
| 2017 | N/A | N/A | N/A | 0.750 | 0.714 | 0.759 |
| 2018 | N/A | N/A | N/A | 0.768 | 0.715 | 0.789 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **State** | **CMAQ Ensemble Training R2** | **CMAQ Ensemble Training Spatial R2** | **CMAQ Ensemble Training Temporal R2** | **Non-CMAQ Ensemble Training R2** | **Non-CMAQ Ensemble Training Spatial R2** | **Non-CMAQ Ensemble Training Temporal R2** |
| Arizona | 0.742 | 0.872 | 0.651 | 0.691 | 0.669 | 0.616 |
| California | 0.780 | 0.673 | 0.756 | 0.747 | 0.716 | 0.723 |
| Colorado | 0.572 | 0.530 | 0.591 | 0.513 | 0.467 | 0.562 |
| Idaho | 0.748 | 0.638 | 0.794 | 0.771 | 0.708 | 0.790 |
| Montana | 0.727 | 0.623 | 0.727 | 0.722 | 0.674 | 0.712 |
| Nevada | 0.794 | 0.935 | 0.753 | 0.780 | 0.933 | 0.743 |
| New Mexico | 0.643 | 0.557 | 0.652 | 0.629 | 0.491 | 0.617 |
| Oregon | 0.681 | 0.636 | 0.695 | 0.701 | 0.669 | 0.722 |
| Utah | 0.753 | 0.654 | 0.761 | 0.702 | 0.555 | 0.710 |
| Washington | 0.658 | 0.649 | 0.665 | 0.700 | 0.628 | 0.724 |
| Wyoming | 0.529 | 0.622 | 0.662 | 0.543 | 0.592 | 0.669 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Season** | **CMAQ Ensemble Training R2** | **CMAQ Ensemble Training Spatial R2** | **CMAQ Ensemble Training Temporal R2** | **Non-CMAQ Ensemble Training R2** | **Non-CMAQ Ensemble Training Spatial R2** | **Non-CMAQ Ensemble Training Temporal R2** |
| Fall | 0.658 | 0.584 | 0.647 | 0.690 | 0.635 | 0.696 |
| Spring | 0.737 | 0.803 | 0.604 | 0.697 | 0.733 | 0.566 |
| Summer | 0.700 | 0.624 | 0.715 | 0.708 | 0.680 | 0.687 |
| Winter | 0.832 | 0.879 | 0.735 | 0.794 | 0.872 | 0.689 |

Description and results of the high-low models:

We developed a preliminary classification model to split the data into “high” versus “low” values, with 15µg/m3 being the most plausible and accurate cut-point. Here, plausibility refers to the fact that with a split of 15 µg/m3or higher, variables associated with wildfires began ranking in the top 20 most important variables (calculated with the “permutation” importance algorithm). After developing this preliminary classification model, we pursued two approaches to develop regression models for the data classified as “high” or “low”. The first was to train the “high” and “low” regression models on data classified as “high” and “low”, respectively. The second was to train the “high” and “low” regression models on data that we knew to be greater than or less than 15 µg/m3respectively. Ultimately, we found that these splitting approaches did not improve our predictions at higher levels of PM2.5. In fact, they only slightly improved predictions at lower levels of PM2.5. Thus, we decided to use the overall models (not splitting into “high” or “low”) for the final analysis.