Supplementary Material

Project: Part 4

Erick E. Mollinedo

2024-04-05

# 1. Supplementary Material

The following document includes some of the results from the exploratory data analysis.

## 1.1 Supplementary Results

[Figure 1](#fig-conc) shows the summary of concentrations for each chemical species categorized by study arm. It is observed a similar pattern observed from the categorization by type of fuel, where Mg, Mn, Ca, K, BC, Ti, Si and S have statistically significant differences in concentrations (shown as \*\*\*\*).

|  |
| --- |
| Figure 1: Chemical species concentrations by study arm |

In terms of the categorical exposures [Figure 2](#fig-cat) summarizes the responses from all the exposures. It is observed that for the incence and generator variables there were just a few or none observations that recorded being exposed to these sources. Meanwhile, stove usage was the exposure most recorded.

|  |
| --- |
| Figure 2: Summary of categorical exposures |

The following figures show the distribution of the concentrations of Carbon monoxide ([Figure 3](#fig-co)) and PM2.5 ([Figure 4](#fig-pm)), which were considered as covariates (CO) or independent variables (PM2.5) to be included in the modelling process with the categorical exposures.

|  |
| --- |
| Figure 3: Carbon Monoxide concentrations by type of fuel |

|  |
| --- |
| Figure 4: PM2.5 concentrations by type of fuel |

The following tables show the results when conducting the generalized linear models using a single or multiple variables for the chemical species that can be interpreted by sources of exposure.

Warning: `pull\_workflow\_fit()` was deprecated in workflows 0.2.3.  
ℹ Please use `extract\_fit\_parsnip()` instead.

# A tibble: 10 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) 0.0509 0.239 0.213 0.831   
 2 stove\_Yes -0.446 0.237 -1.88 0.0606  
 3 smoke\_Yes -0.219 0.232 -0.945 0.345   
 4 coil\_Yes -0.243 0.185 -1.32 0.188   
 5 trash\_Yes -0.0716 0.0992 -0.722 0.471   
 6 kerosene\_Yes 0.270 0.203 1.33 0.183   
 7 smoky\_Yes -0.0760 0.220 -0.345 0.730   
 8 crop\_Yes 0.239 0.233 1.03 0.305   
 9 stove\_other\_Yes 0.0378 0.0671 0.564 0.573   
10 fueltype\_LPG -0.0567 0.0546 -1.04 0.300

# A tibble: 10 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) 2.45 0.237 10.4 2.86e-23  
 2 stove\_Yes 0.000402 0.235 0.00171 9.99e- 1  
 3 smoke\_Yes -0.199 0.230 -0.868 3.86e- 1  
 4 coil\_Yes 0.107 0.183 0.583 5.60e- 1  
 5 trash\_Yes 0.0241 0.0982 0.245 8.06e- 1  
 6 kerosene\_Yes 0.885 0.201 4.41 1.24e- 5  
 7 smoky\_Yes -0.184 0.218 -0.845 3.98e- 1  
 8 crop\_Yes 0.0817 0.231 0.354 7.24e- 1  
 9 stove\_other\_Yes 0.162 0.0665 2.44 1.51e- 2  
10 fueltype\_LPG -0.981 0.0541 -18.1 4.78e-59

# A tibble: 10 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) -0.630 0.295 -2.13 3.33e- 2  
 2 stove\_Yes -0.286 0.293 -0.977 3.29e- 1  
 3 smoke\_Yes -0.173 0.286 -0.603 5.47e- 1  
 4 coil\_Yes -0.0749 0.228 -0.329 7.42e- 1  
 5 trash\_Yes -0.0687 0.122 -0.561 5.75e- 1  
 6 kerosene\_Yes 0.199 0.250 0.796 4.26e- 1  
 7 smoky\_Yes 0.138 0.272 0.507 6.12e- 1  
 8 crop\_Yes -0.105 0.288 -0.366 7.15e- 1  
 9 stove\_other\_Yes 0.0871 0.0829 1.05 2.94e- 1  
10 fueltype\_LPG -0.628 0.0675 -9.31 2.33e-19

# A tibble: 10 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) -0.475 0.308 -1.54 0.124   
 2 stove\_Yes -0.664 0.306 -2.17 0.0304  
 3 smoke\_Yes -0.122 0.299 -0.408 0.683   
 4 coil\_Yes -0.0685 0.238 -0.288 0.774   
 5 trash\_Yes -0.0762 0.128 -0.596 0.552   
 6 kerosene\_Yes 0.164 0.261 0.627 0.531   
 7 smoky\_Yes -0.141 0.284 -0.498 0.618   
 8 crop\_Yes 0.0119 0.301 0.0397 0.968   
 9 stove\_other\_Yes 0.0905 0.0866 1.04 0.296   
10 fueltype\_LPG -0.129 0.0705 -1.83 0.0680

# A tibble: 10 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) 0.354 0.289 1.23 2.20e- 1  
 2 stove\_Yes 0.00839 0.286 0.0293 9.77e- 1  
 3 smoke\_Yes -0.548 0.280 -1.96 5.08e- 2  
 4 coil\_Yes -0.201 0.223 -0.902 3.67e- 1  
 5 trash\_Yes 0.101 0.120 0.847 3.97e- 1  
 6 kerosene\_Yes 0.0189 0.245 0.0771 9.39e- 1  
 7 smoky\_Yes 0.0755 0.265 0.284 7.76e- 1  
 8 crop\_Yes -0.249 0.282 -0.885 3.77e- 1  
 9 stove\_other\_Yes 0.235 0.0810 2.90 3.85e- 3  
10 fueltype\_LPG -1.32 0.0659 -20.0 1.32e-68

# A tibble: 10 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) -1.77 0.240 -7.39 4.97e-13  
 2 stove\_Yes -0.174 0.238 -0.730 4.66e- 1  
 3 smoke\_Yes -0.223 0.233 -0.960 3.37e- 1  
 4 coil\_Yes -0.217 0.185 -1.17 2.41e- 1  
 5 trash\_Yes 0.00157 0.0995 0.0158 9.87e- 1  
 6 kerosene\_Yes 0.395 0.203 1.94 5.28e- 2  
 7 smoky\_Yes 0.0288 0.221 0.131 8.96e- 1  
 8 crop\_Yes 0.0759 0.234 0.324 7.46e- 1  
 9 stove\_other\_Yes 0.0389 0.0673 0.578 5.63e- 1  
10 fueltype\_LPG -0.353 0.0548 -6.45 2.31e-10

# A tibble: 10 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) -3.30 0.347 -9.53 3.75e-20  
 2 stove\_Yes -0.520 0.344 -1.51 1.31e- 1  
 3 smoke\_Yes -0.0384 0.337 -0.114 9.09e- 1  
 4 coil\_Yes 0.173 0.268 0.645 5.19e- 1  
 5 trash\_Yes -0.0574 0.144 -0.399 6.90e- 1  
 6 kerosene\_Yes 0.0673 0.294 0.229 8.19e- 1  
 7 smoky\_Yes 0.0304 0.319 0.0951 9.24e- 1  
 8 crop\_Yes -0.293 0.338 -0.867 3.86e- 1  
 9 stove\_other\_Yes -0.0339 0.0974 -0.348 7.28e- 1  
10 fueltype\_LPG -0.325 0.0793 -4.09 4.80e- 5

# A tibble: 10 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) -0.567 0.257 -2.21 0.0275   
 2 stove\_Yes -0.260 0.255 -1.02 0.307   
 3 smoke\_Yes -0.0341 0.249 -0.137 0.891   
 4 coil\_Yes 0.164 0.198 0.828 0.408   
 5 trash\_Yes 0.0814 0.106 0.765 0.445   
 6 kerosene\_Yes -0.115 0.218 -0.530 0.597   
 7 smoky\_Yes -0.0391 0.236 -0.166 0.868   
 8 crop\_Yes -0.230 0.250 -0.917 0.359   
 9 stove\_other\_Yes 0.0531 0.0721 0.737 0.461   
10 fueltype\_LPG -0.341 0.0586 -5.81 0.0000000103

# A tibble: 10 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) 0.510 0.366 1.39 0.164   
 2 stove\_Yes -0.624 0.363 -1.72 0.0867   
 3 smoke\_Yes -0.303 0.356 -0.853 0.394   
 4 coil\_Yes -0.493 0.283 -1.74 0.0821   
 5 trash\_Yes -0.0425 0.152 -0.279 0.780   
 6 kerosene\_Yes 0.268 0.311 0.864 0.388   
 7 smoky\_Yes -0.184 0.337 -0.545 0.586   
 8 crop\_Yes 0.234 0.357 0.654 0.513   
 9 stove\_other\_Yes 0.156 0.103 1.51 0.131   
10 fueltype\_LPG -0.227 0.0837 -2.71 0.00692

# A tibble: 10 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) -2.66 0.295 -9.02 2.38e-18  
 2 stove\_Yes -0.498 0.293 -1.70 8.95e- 2  
 3 smoke\_Yes -0.0406 0.286 -0.142 8.87e- 1  
 4 coil\_Yes -0.0493 0.228 -0.216 8.29e- 1  
 5 trash\_Yes -0.0935 0.122 -0.764 4.45e- 1  
 6 kerosene\_Yes -0.00443 0.250 -0.0177 9.86e- 1  
 7 smoky\_Yes 0.0115 0.272 0.0425 9.66e- 1  
 8 crop\_Yes -0.0380 0.288 -0.132 8.95e- 1  
 9 stove\_other\_Yes 0.0625 0.0829 0.754 4.51e- 1  
10 fueltype\_LPG -0.164 0.0675 -2.42 1.57e- 2

# A tibble: 10 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) -4.05 0.416 -9.73 6.58e-21  
 2 stove\_Yes -0.292 0.413 -0.706 4.81e- 1  
 3 smoke\_Yes 0.187 0.404 0.463 6.43e- 1  
 4 coil\_Yes 0.254 0.322 0.790 4.30e- 1  
 5 trash\_Yes -0.155 0.173 -0.897 3.70e- 1  
 6 kerosene\_Yes 0.176 0.353 0.498 6.19e- 1  
 7 smoky\_Yes -0.505 0.383 -1.32 1.88e- 1  
 8 crop\_Yes 0.0151 0.406 0.0371 9.70e- 1  
 9 stove\_other\_Yes -0.0719 0.117 -0.615 5.39e- 1  
10 fueltype\_LPG -0.137 0.0952 -1.43 1.52e- 1

The following table shows the values used to judge models (RMSE and R-square) for all chemical species between Model 1 and Model 2.

| Species | Metric | Model1 | Model2 |
| --- | --- | --- | --- |
| Al | rmse | 0.440169511 | 0.439470704 |
| Al | rsq | 0.040423126 | 0.022147345 |
| BC | rmse | 5.784293838 | 5.735175716 |
| BC | rsq | 0.292695691 | 0.286040808 |
| Ca | rmse | 0.274855750 | 0.273633522 |
| Ca | rsq | 0.091593067 | 0.099423672 |
| Fe | rmse | 0.266682770 | 0.265707118 |
| Fe | rsq | 0.032549451 | 0.011517699 |
| K | rmse | 1.011600802 | 1.009296705 |
| K | rsq | 0.210173790 | 0.212270684 |
| Mg | rmse | 0.084923122 | 0.084723717 |
| Mg | rsq | 0.052590391 | 0.047682072 |
| Mn | rmse | 0.018369471 | 0.018295808 |
| Mn | rsq | 0.024020094 | 0.025648627 |
| Si | rmse | 0.861906148 | 0.858238387 |
| Si | rsq | 0.026369796 | 0.011574612 |
| S | rmse | 0.277407523 | 0.274785778 |
| S | rsq | 0.028761304 | 0.038096777 |
| Ti | rmse | 0.033103748 | 0.032866467 |
| Ti | rsq | 0.019728205 | 0.009791549 |
| Zn | rmse | 0.013691216 | 0.013651982 |
| Zn | rsq | 0.008173018 | 0.009604463 |