Supplementary Material

Erick E. Mollinedo

2024-04-19

# 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 |

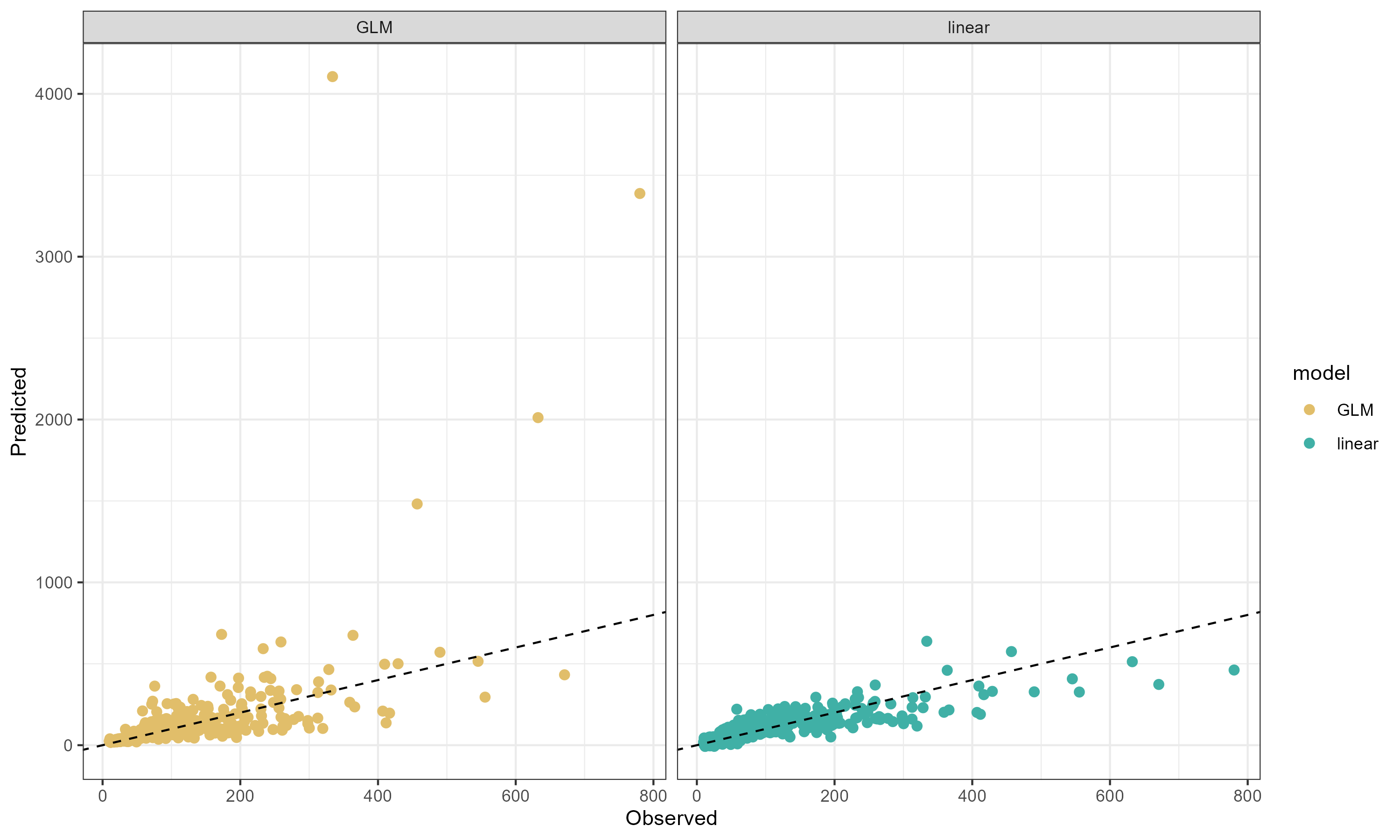
|  |
| --- |
| Figure 4: PM2.5 concentrations by type of fuel |

The following tables show the model summaries for the final models selected. In this case the linear regression model for the PM2.5 mass concentrations and the gamma GLM model for the BC mass concentrations.

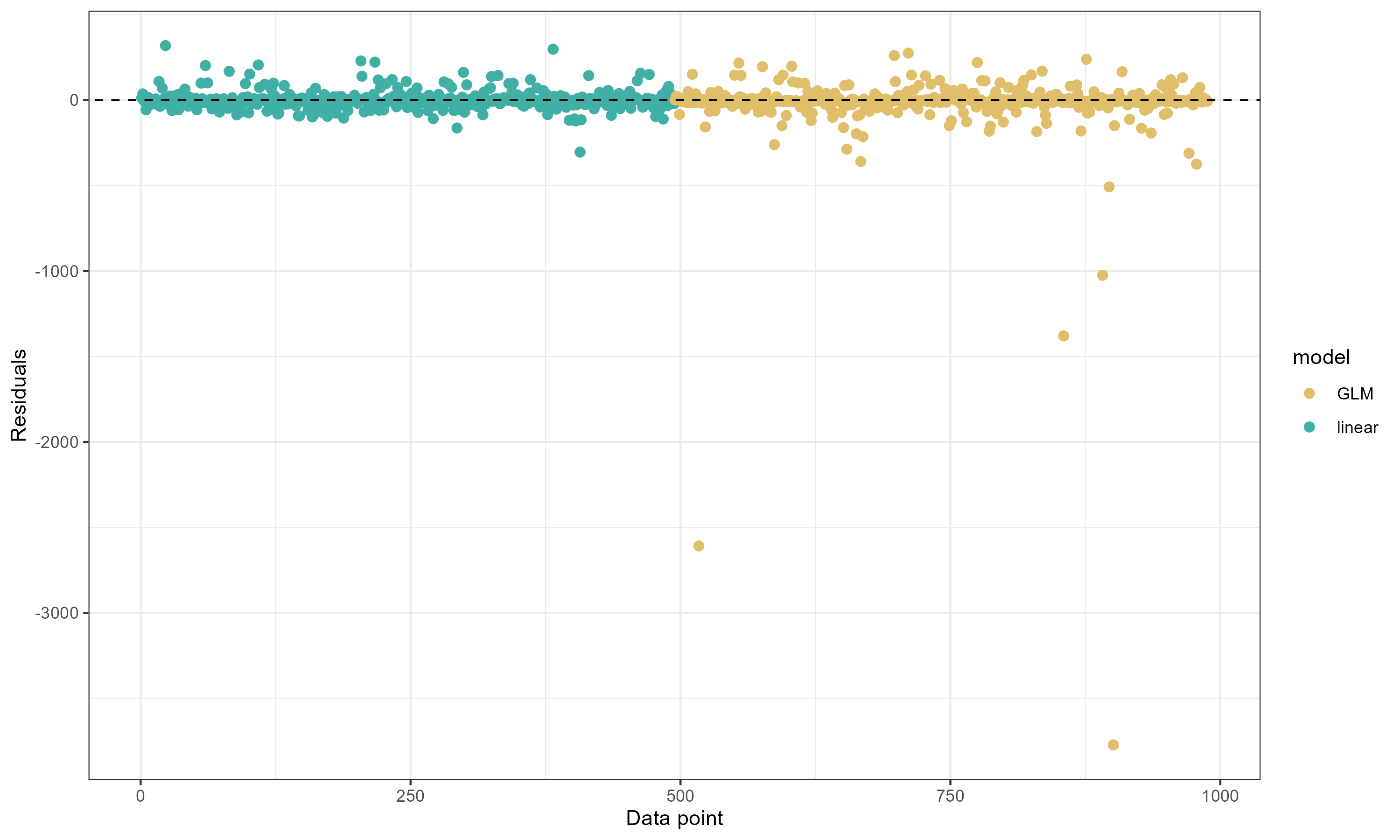
# A tibble: 12 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) -10.7 7.17 -1.49 1.37e- 1  
 2 Mg -177. 48.5 -3.66 2.85e- 4  
 3 Al 13.2 13.2 1.00 3.16e- 1  
 4 Si 4.29 7.95 0.540 5.90e- 1  
 5 S -8.96 9.88 -0.907 3.65e- 1  
 6 K 59.1 2.81 21.1 2.86e-70  
 7 Ca 12.3 14.7 0.831 4.06e- 1  
 8 Ti -71.3 160. -0.445 6.57e- 1  
 9 Mn 302. 176. 1.71 8.80e- 2  
10 Fe -6.15 28.2 -0.218 8.27e- 1  
11 Zn 184. 199. 0.924 3.56e- 1  
12 BC 5.32 0.532 9.99 1.68e-21

# A tibble: 9 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
1 (Intercept) 2.44 0.0359 67.8 1.60e-249  
2 smokeYes -0.118 0.232 -0.509 6.11e- 1  
3 coilYes 0.194 0.172 1.13 2.59e- 1  
4 trashYes -0.0312 0.0940 -0.332 7.40e- 1  
5 keroseneYes 0.392 0.206 1.91 5.73e- 2  
6 smokyYes -0.0842 0.217 -0.388 6.98e- 1  
7 cropYes 0.0352 0.237 0.148 8.82e- 1  
8 stove\_otherYes 0.175 0.0679 2.57 1.04e- 2  
9 fueltypeLPG -0.974 0.0536 -18.2 1.08e- 56

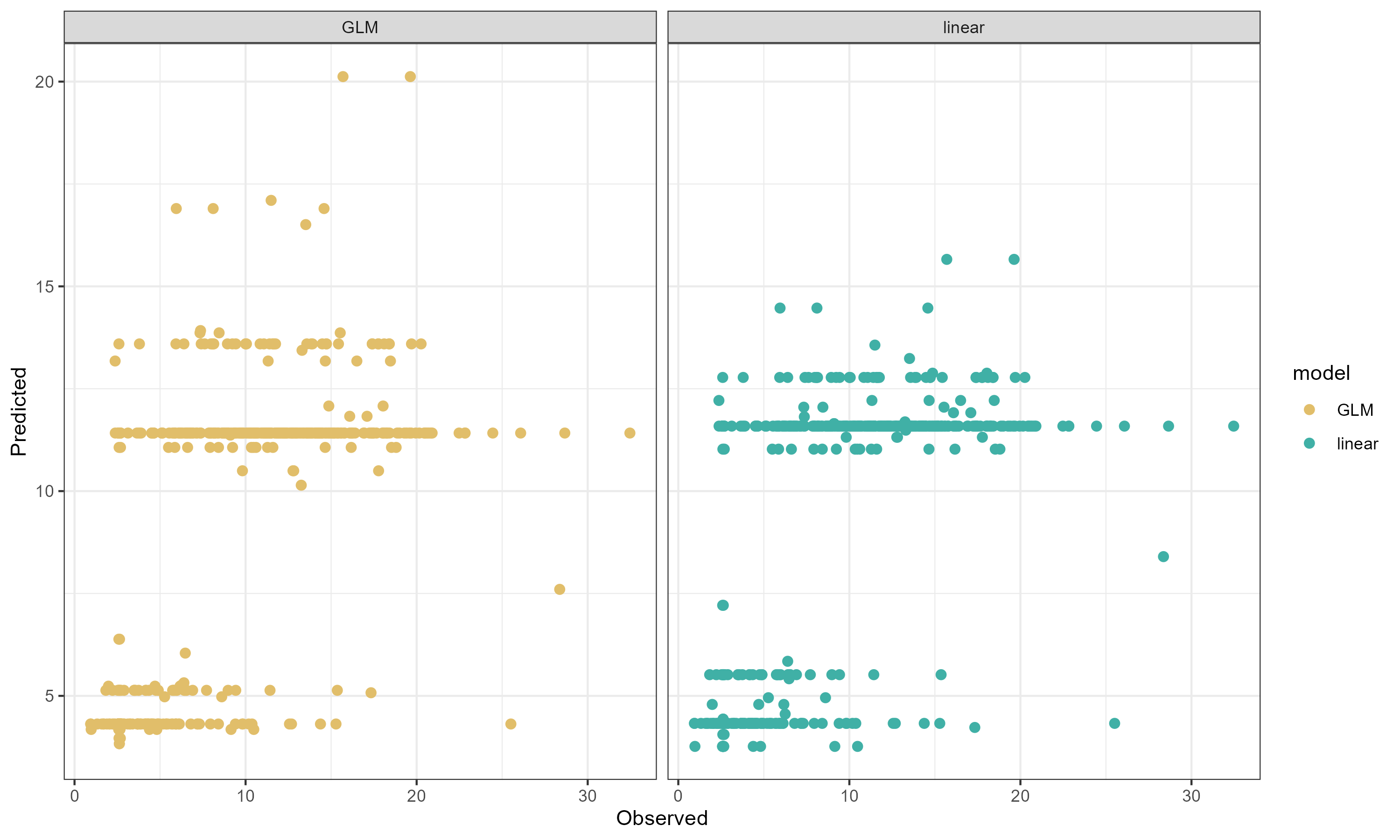
The following figures illustrate the observed vs predicted values plots for both PM2.5 and BC models, and the residual plots.



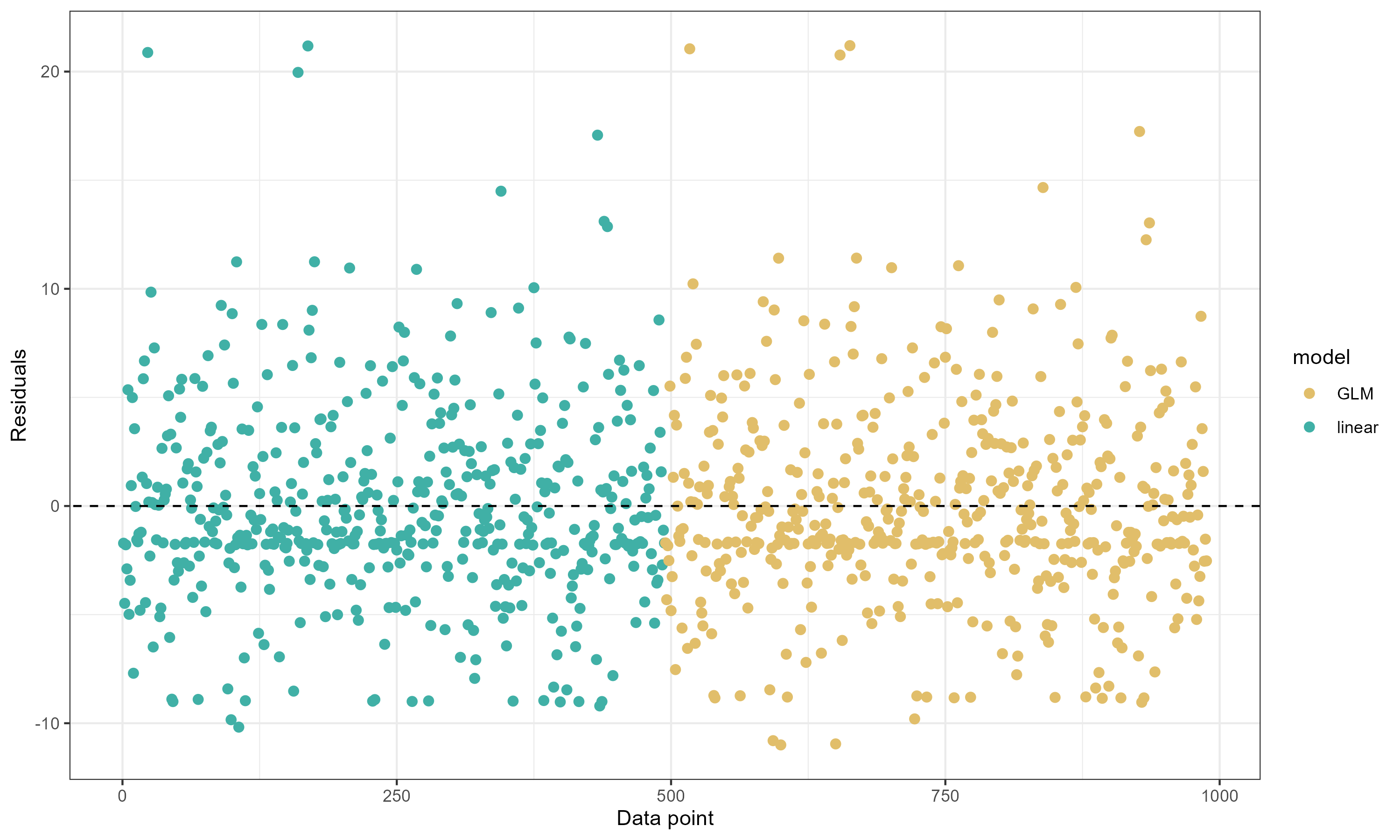
Observed vs Predicted values of the PM2.5 models



PM2.5 Models residuals



Observed vs Predicted values of the BC models



BC models residuals