Supplemental Material

for

Testing the Performance of Field Calibration Techniques for Low-Cost Gas Sensors in New Deployment Locations: Across A County Line and Across Colorado

Joanna Gordon Casey and Michael Hannigan

These supplemental materials consist of a number of plots that support and have relevance to the methods and findings presented in the main article. Figure S1 includes six panels plotting RMSE against r^2 for training at BAO (left side panels) and test data (right side panels) at BAO and SJ Basin sites for ozone from the summer of 2015. The top panels in Fig. S1 are discussed in the main article, showing that while there are some groupings by U-Pod during training these patterns in performance are not replicated in the test data, and in general are much smaller in magnitude than the differences in performance for U-Pods deployed at different locations during the test period. The middle panels show the same comparison of training vs. test deployment ANN performance, this time including all sets of inputs instead of just the best-performing set of 7 inputs. The bottom two plots in Fig. S1 show model performance during training and testing during the summer of 2015, colored by the number of inputs included in each ANN. The paired plots in this figure show that there is a very strong correlation between model performance and the number of inputs included during training. The plots in the figure also show that this correlation breaks down when the models are applied to data collected in a different location than where the model training took place. This group of plots show that individual U-Pod deployment locations during testing had a greater impact on model performance than the number of inputs alone or the specific sensors included in each U-Pod.

The following Fig. S2 – Fig. S7 are referenced in the main article. Each of these figures shows the relative performance of LMs and ANNs with a number of inputs in the quantification of ozone (O_3) and carbon dioxide (CO_2), with the best performing set of inputs highlighted in purple. The best-performing model type and set of inputs are carried forward into the comparative analysis presented and discussed in the main article. Figures S8 – S22 show the relative ranges of O_3, CO_2, and environmental parameters during each training and test data pair, in order to access coverage and to identify extrapolation of field calibration models during testing period. Figure S23 is very similar to Fig. 7.8, but shows target plot performance of CO_2 and O_3 field calibration models in terms of MBE and CRMSE, normalized on an individual basis instead of uniformly. Table S1 shows when bad humidity signals in U-Pods were replaced and which signals they were replaced with.

Table S1 Relative humidity signal replacement in U-Pods with malfunctioning sensors

Deployment	U-Pod With Bad Sensor	Replacement Signal
Spring 2015	BF (Bloomfield)	BJ (Navajo Dam)
Summer 2015	BF (Sub Station)	BI (Ignacio)
Summer 2015	BA (Navajo Dam)	BI (Ignacio)
Summer 2015	BF (BAO)	BI (BAO)
Summer 2015	BA (BAO)	BI (BAO)
Fall 2017	BH (GRET)	Picarro (GRET)
Fall 2017	BI (GRET)	Picarro (GRET)
Spring 2017	BF (GRET)	BI (GRET)

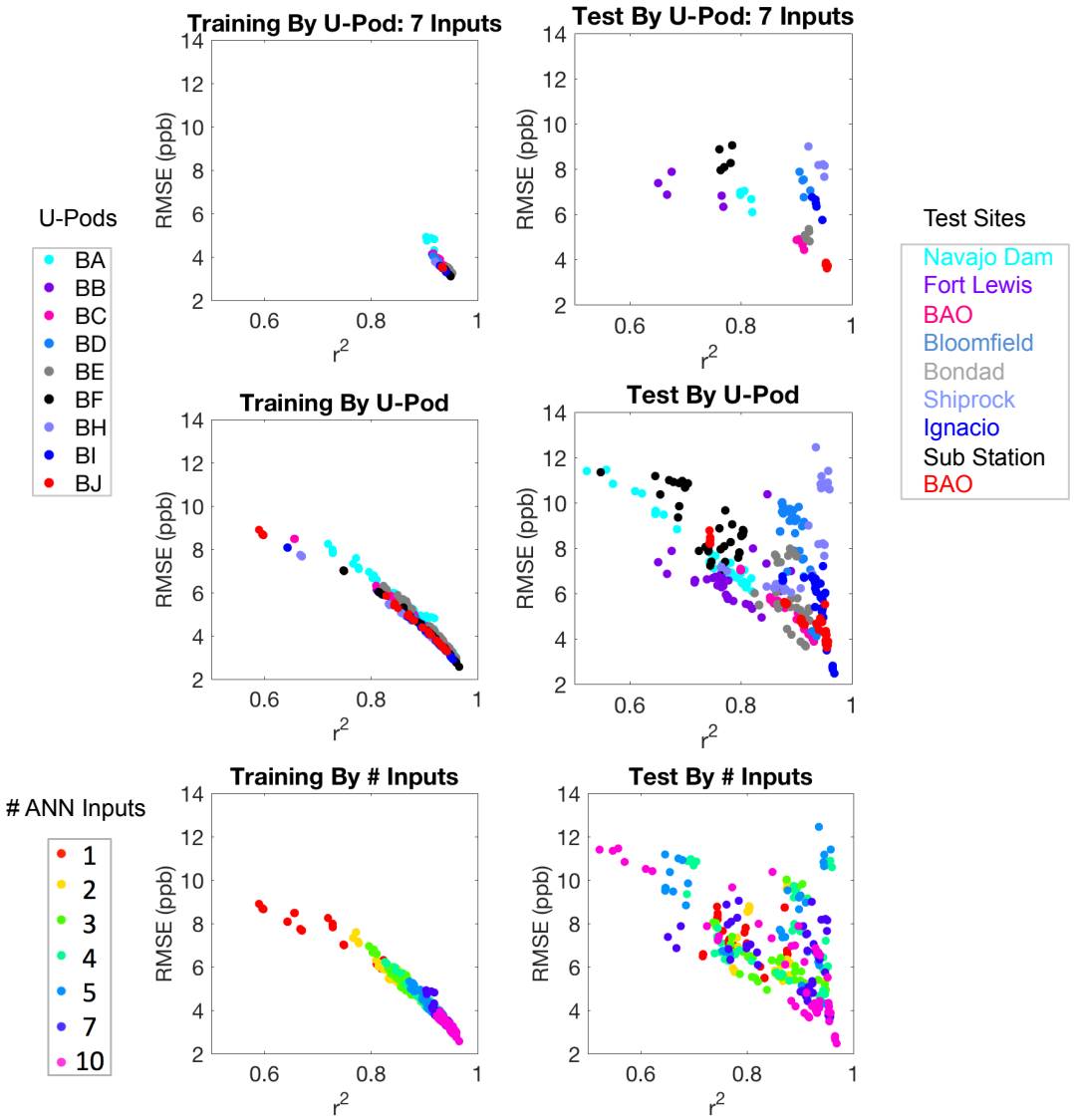

Figure S1 Comparisons of ANNs with various inputs and each U-Pod during the 2015 summer training period at BAO and the summer 2015 BAO and SJ Basin test data period. The plots on the left show the RMSE and r^2 metrics for models during the training period. The plots on the right show these metrics during the test period. In the bottom row, each of the plot markers are colored by the number of inputs included in each ANN. In the top and middle rows, each of the plot markers are colored by U-Pod code (unique sensor enclosures). In the top plots, only metrics for ANNs with 7 inputs are shown, while in the middle plots, ANNs with a range of inputs are included.

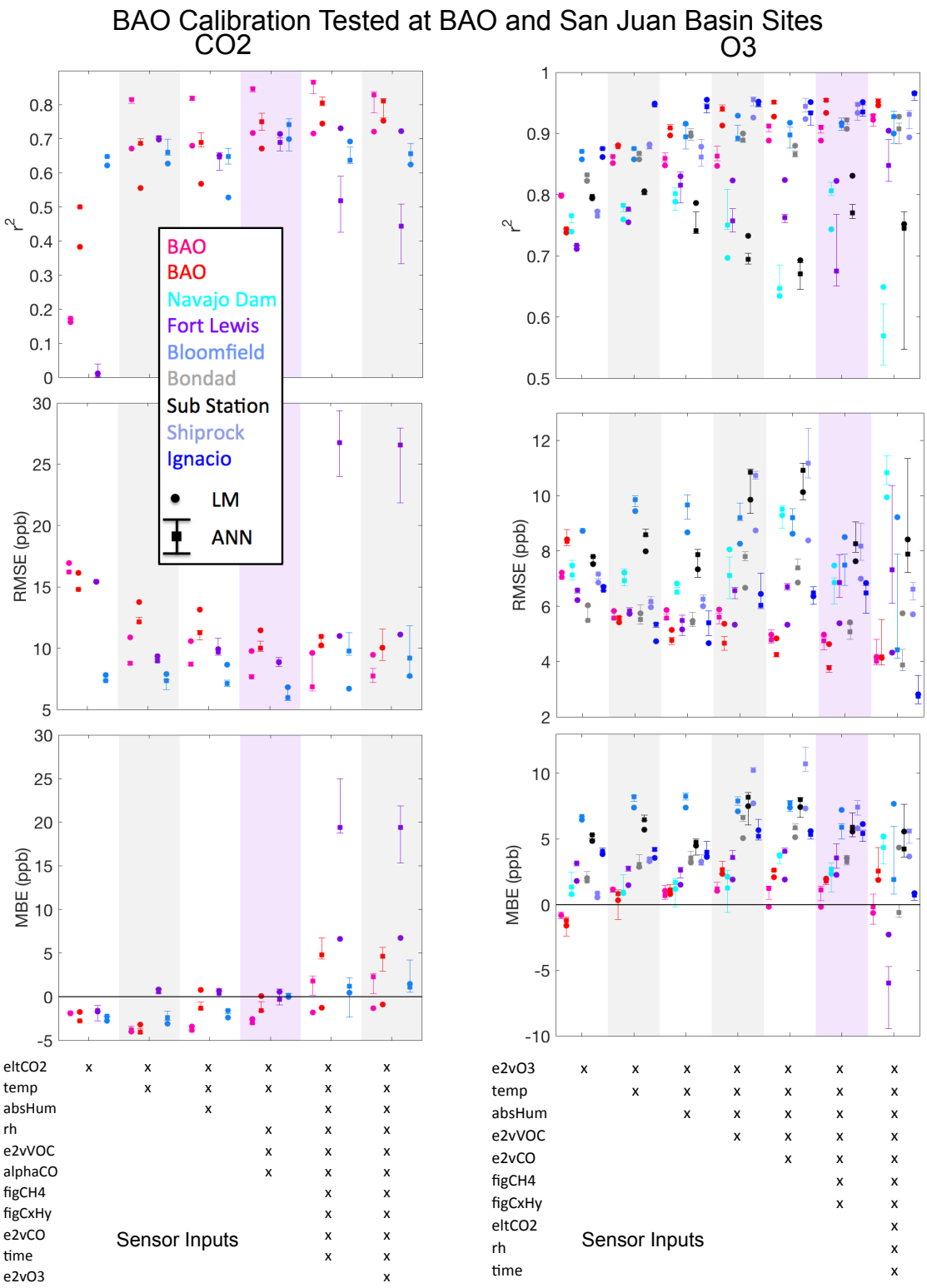

Figure S2 Comparisons of ANNs and LMs with various inputs during the summer 2015 BAO and SJ Basin test data period. ANN plot-markers and error bars show the median and range of 5 separate training iterations. ANNs performed better than LMs. The light purple band indicates the best performing set of model inputs.

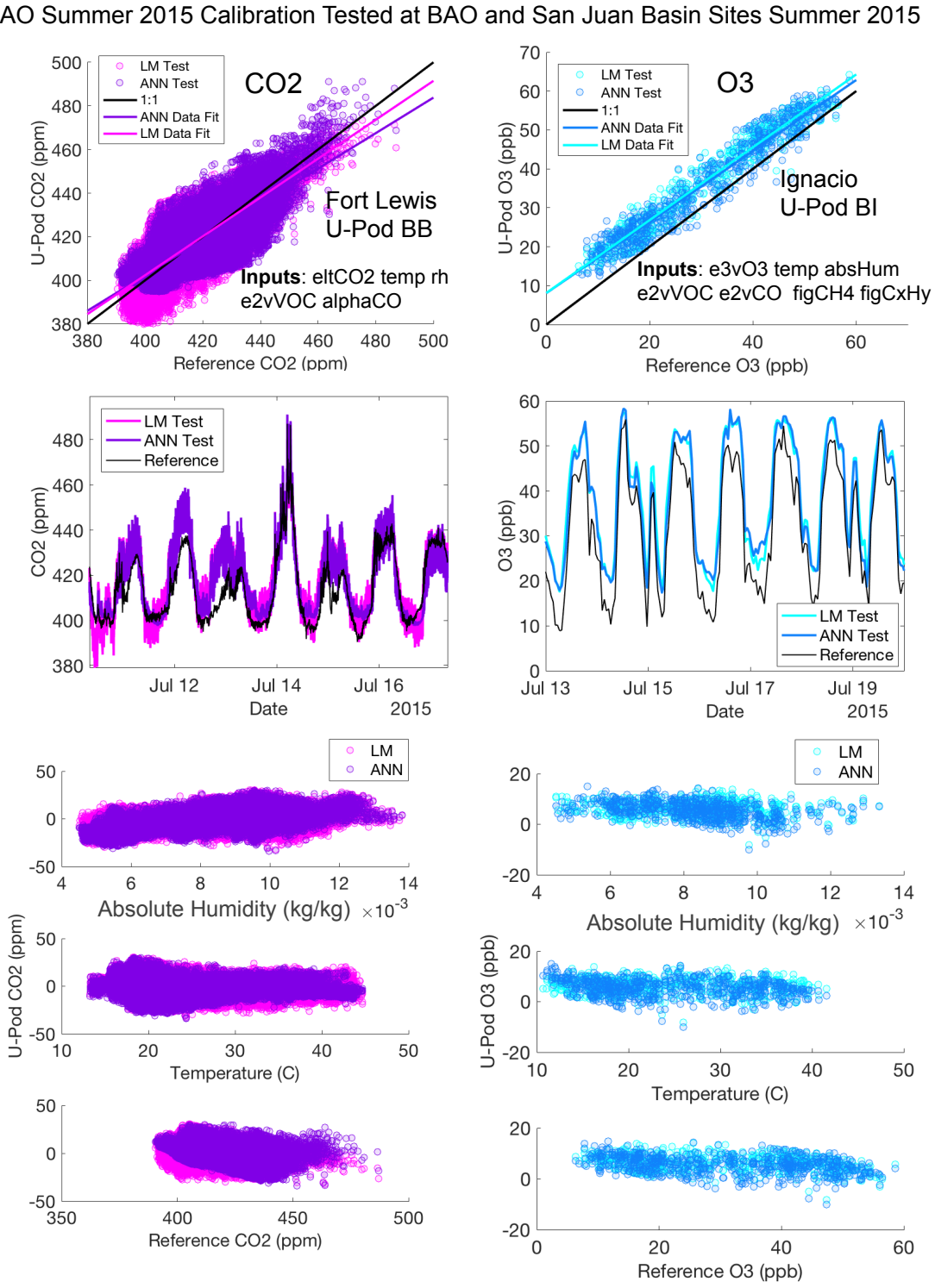

Figure S3 Representative performance of ANNs relative to LMs during the summer 2015 SJ Basin test data period. Other model inputs were optimal for other training and test data pairs, corresponding to different sampling sites and timing.

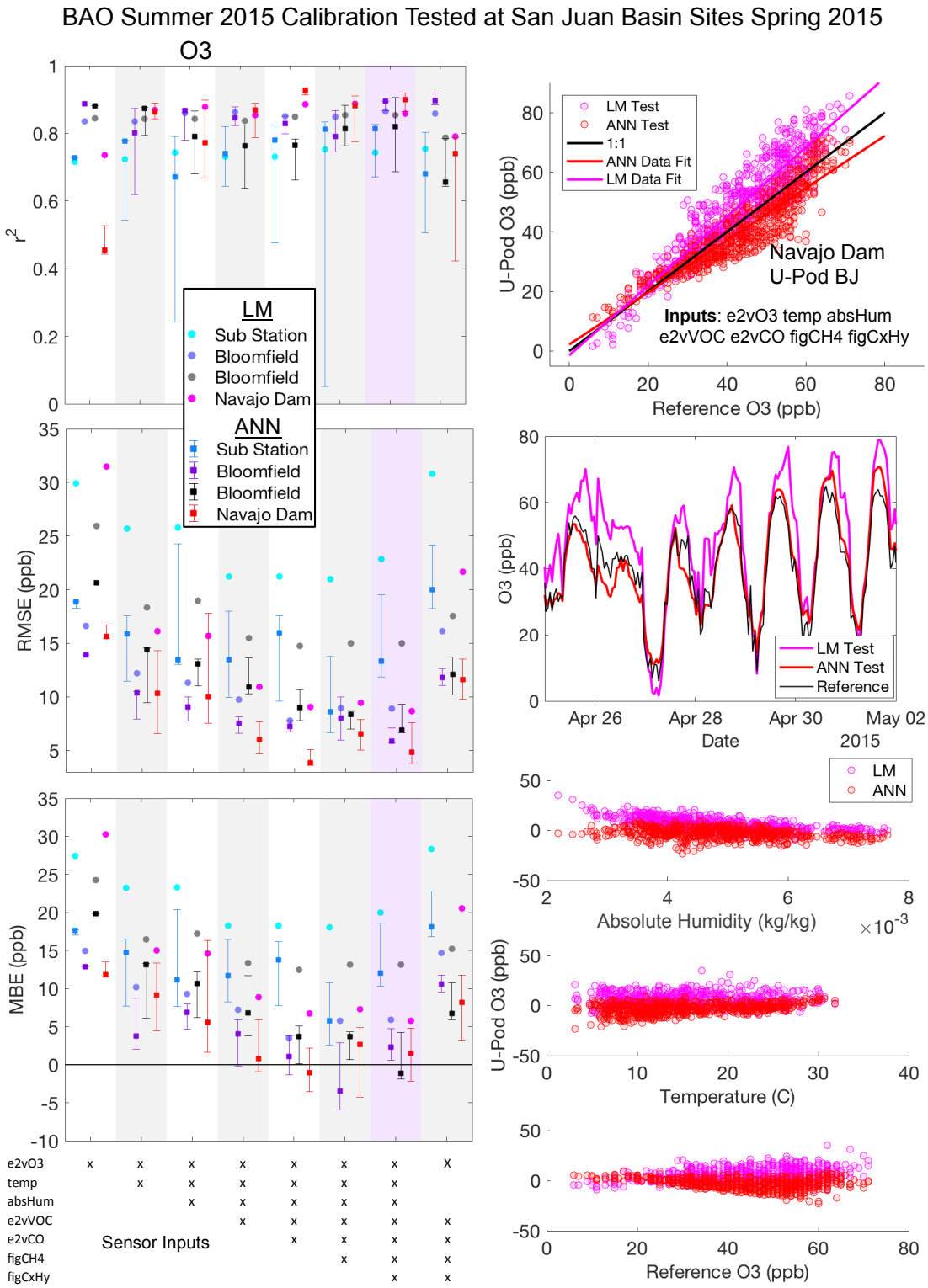

Figure S4 Comparisons of ANNs and LMs with various inputs during the spring 2015 SJ Basin test data period. ANN plot-markers and error bars show the median and range of 5 separate training iterations. ANNs performed better than LMs. The light purple band indicates the best performing set of model inputs.

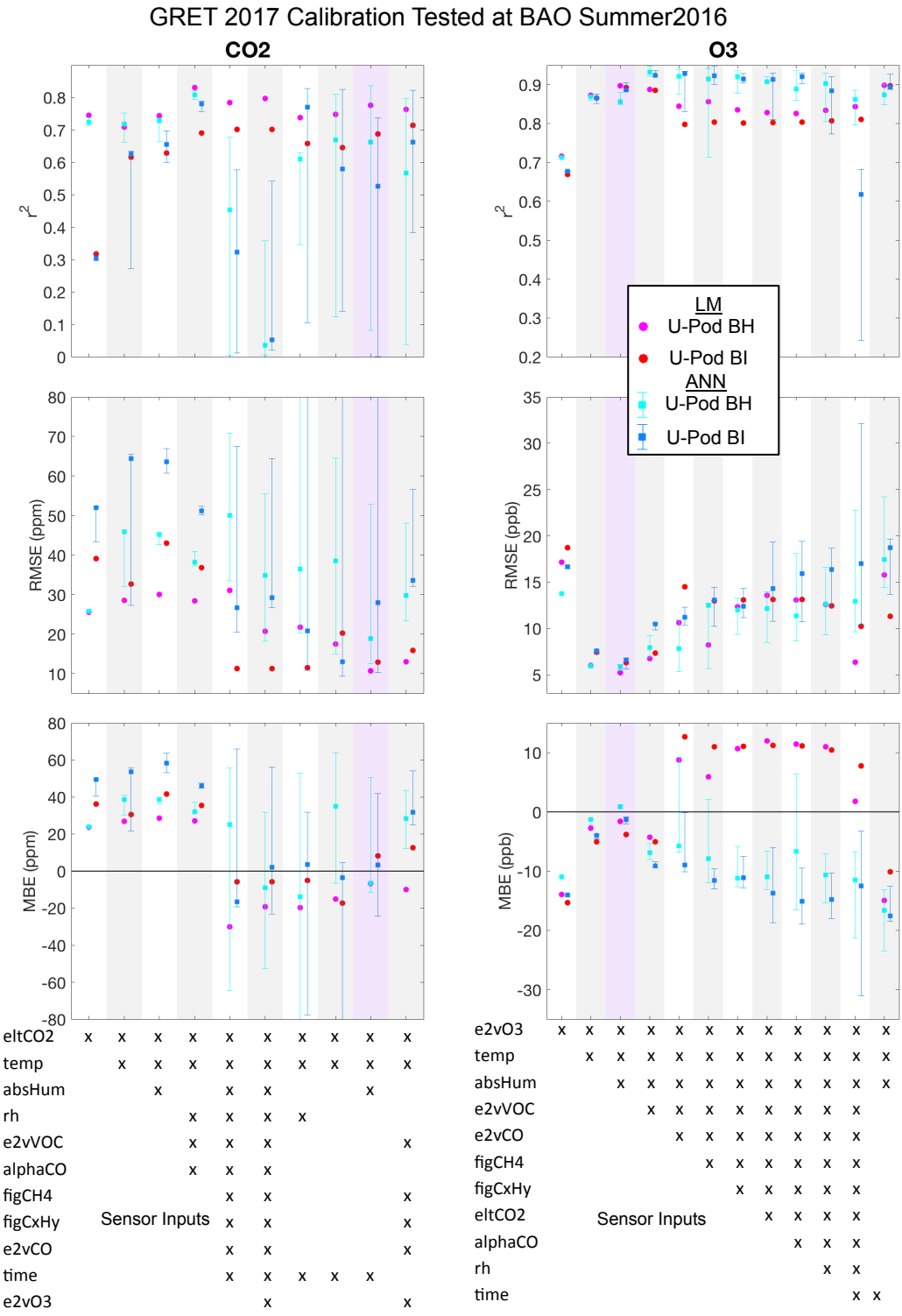

Figure S5 Comparisons of ANNs and LMs with various inputs during the summer 2016 BAO test data period. ANN plot-markers and error bars show the median and range of 5 separate training iterations. LMs performed better than ANNs. The light purple band indicates the best performing set of model inputs.

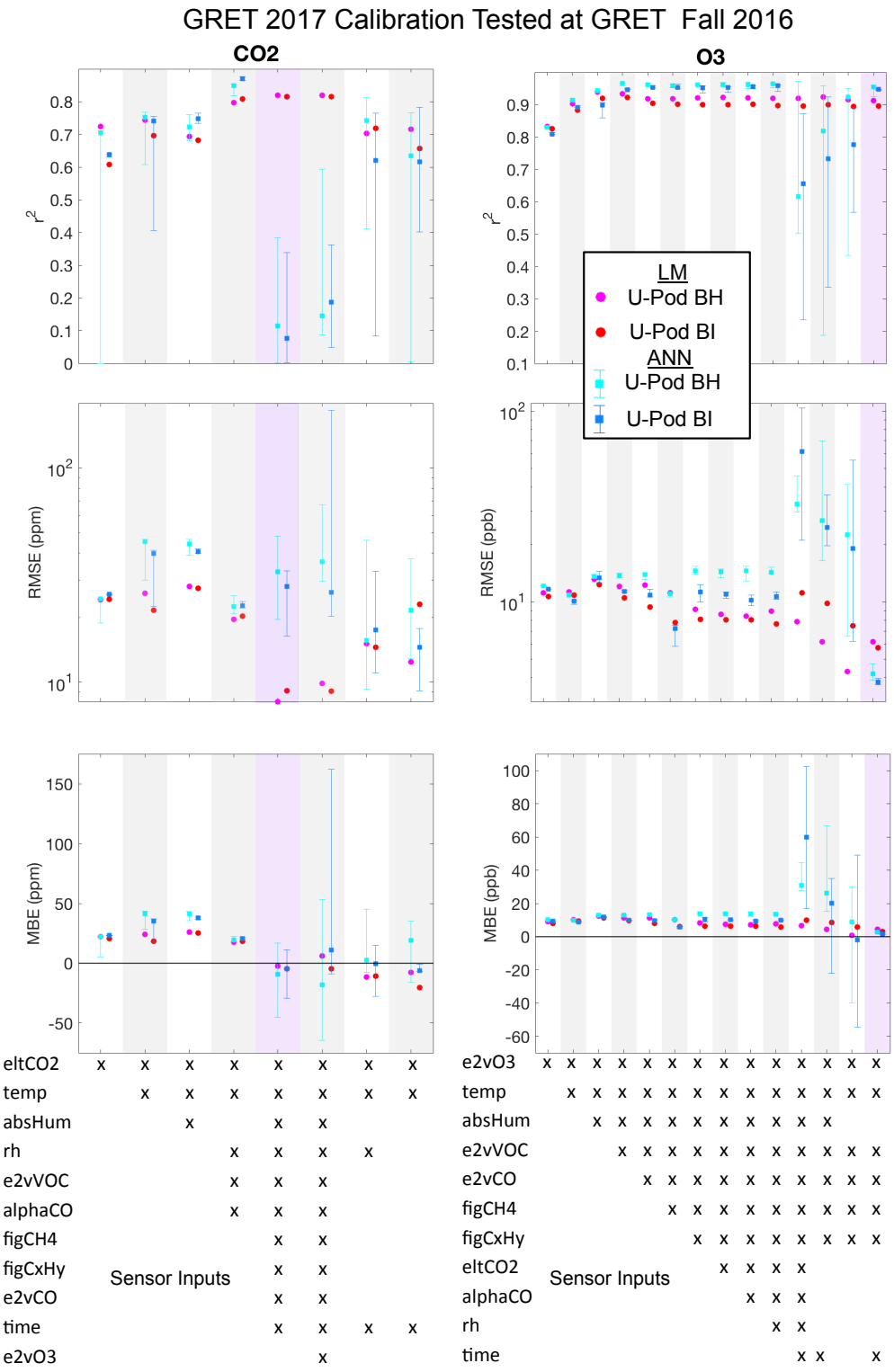

Figure S6 Comparisons of ANNs and LMs with various inputs during the fall 2016 GRET test data period. ANN plot-markers and error bars show the median and range of 5 separate training iterations. Note the log scale on the RMSE plots for CO_2 and O_3. For CO_2 LMs performed better but for O_3 ANNs performed better. The light purple band indicates the best performing set of model inputs.

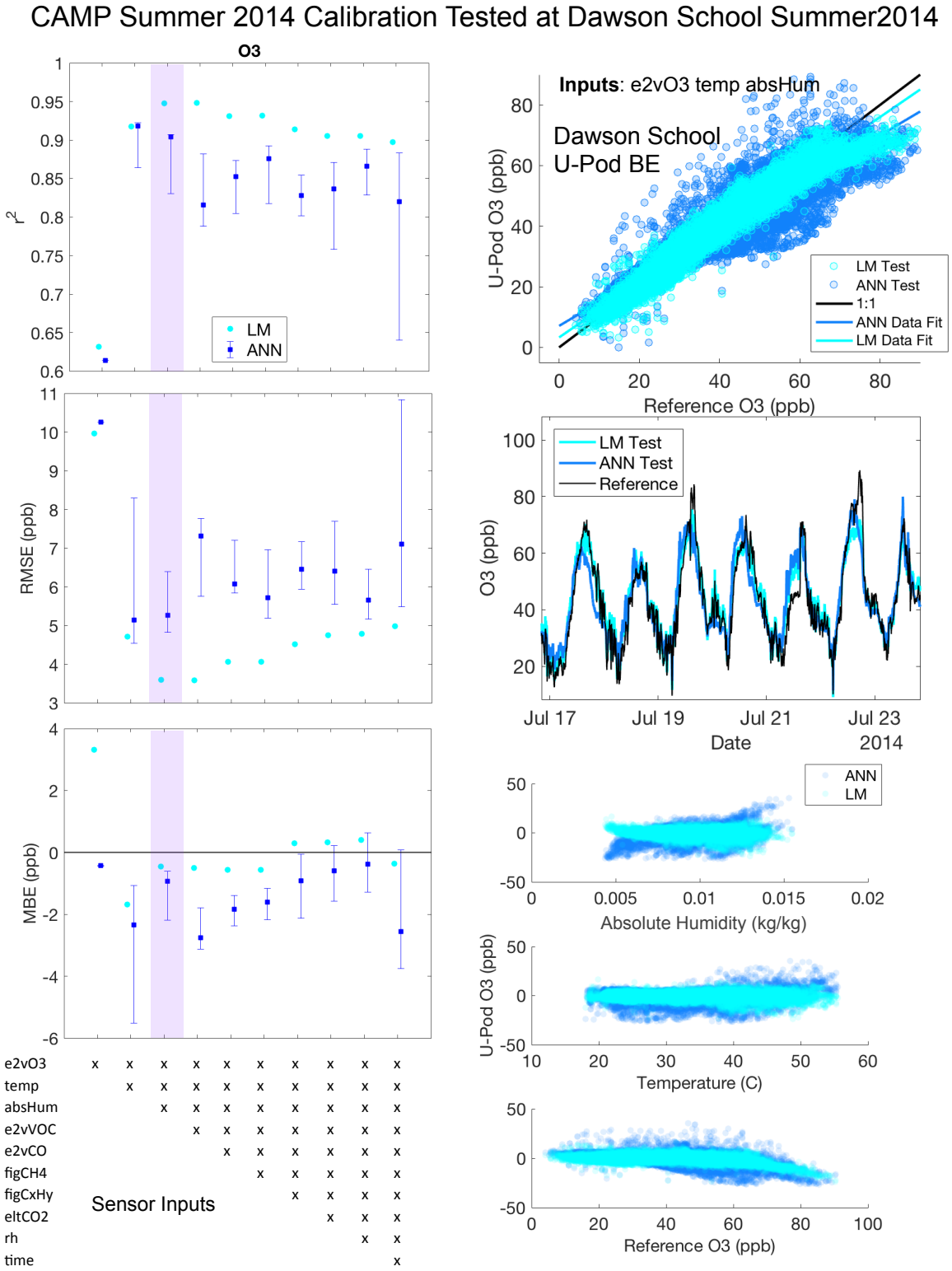

Figure S7 Comparisons of ANNs and LMs with various inputs during the summer 2014 Dawson test data period using models trained with data from CAMP. ANN plot-markers and error bars show the median and range of 5 separate training iterations. LMs performed better than ANNs. The light purple band indicates the best performing set of model inputs.

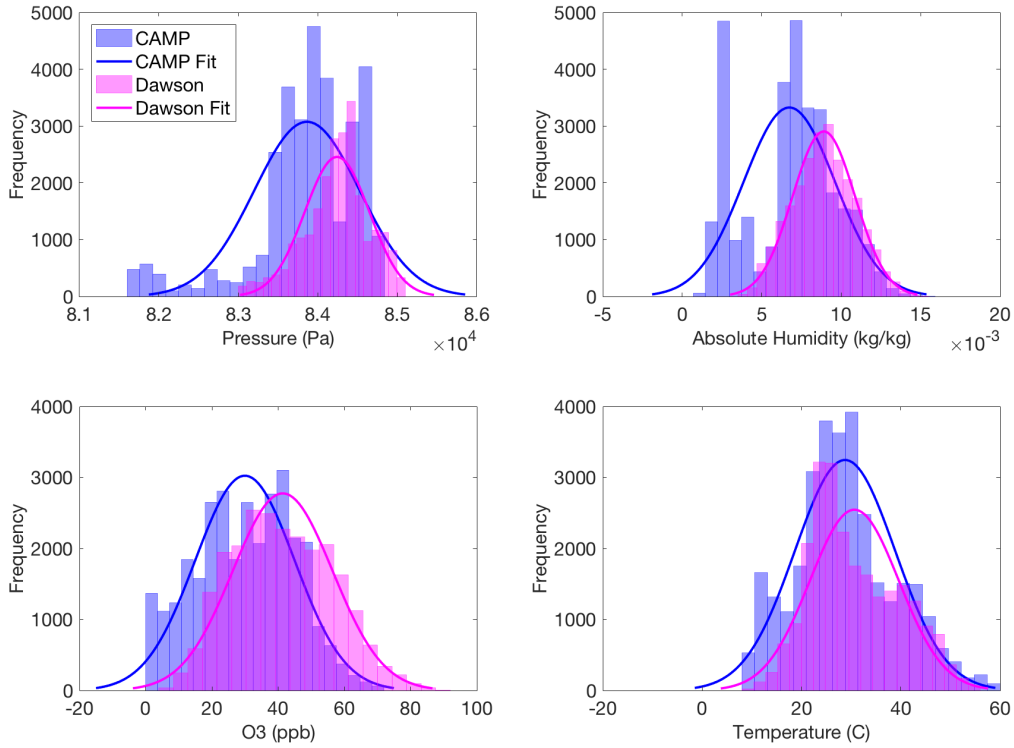

Figure S8 Variable Space during CAMP training and Dawson School test

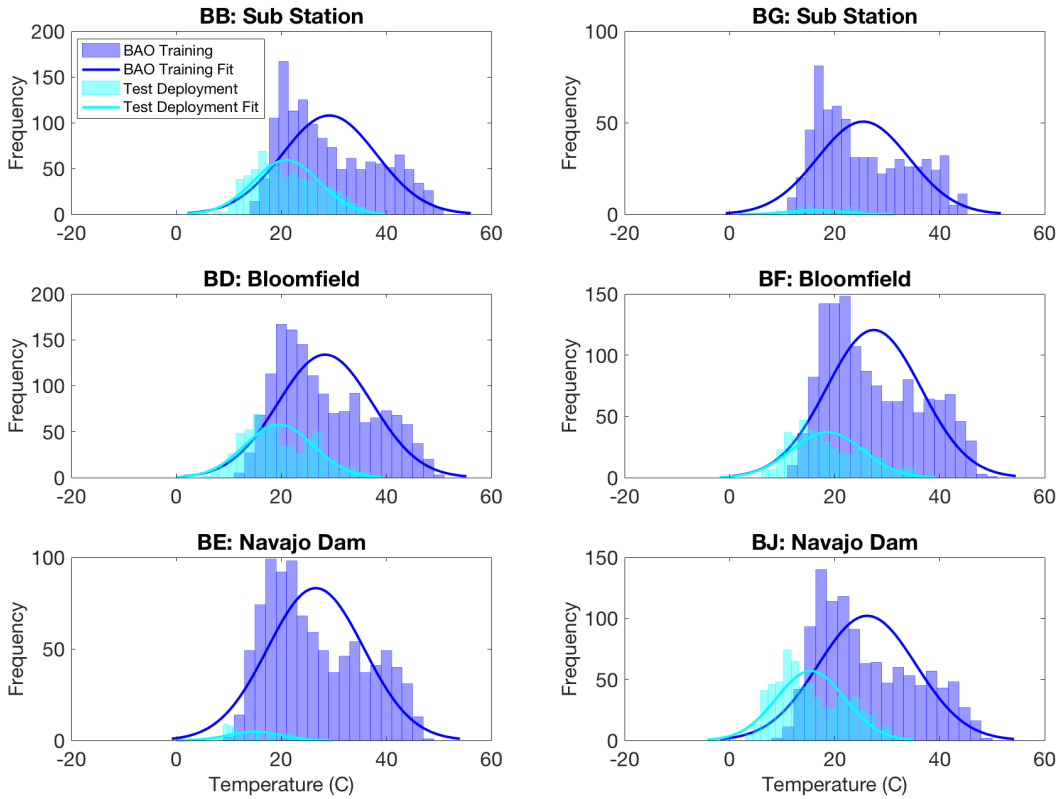

Figure S9 Temperatures during BAO training and spring SJ Basin test. Note that due to significant amounts of missing data, U-Pods BG and BE were not included in the analysis.

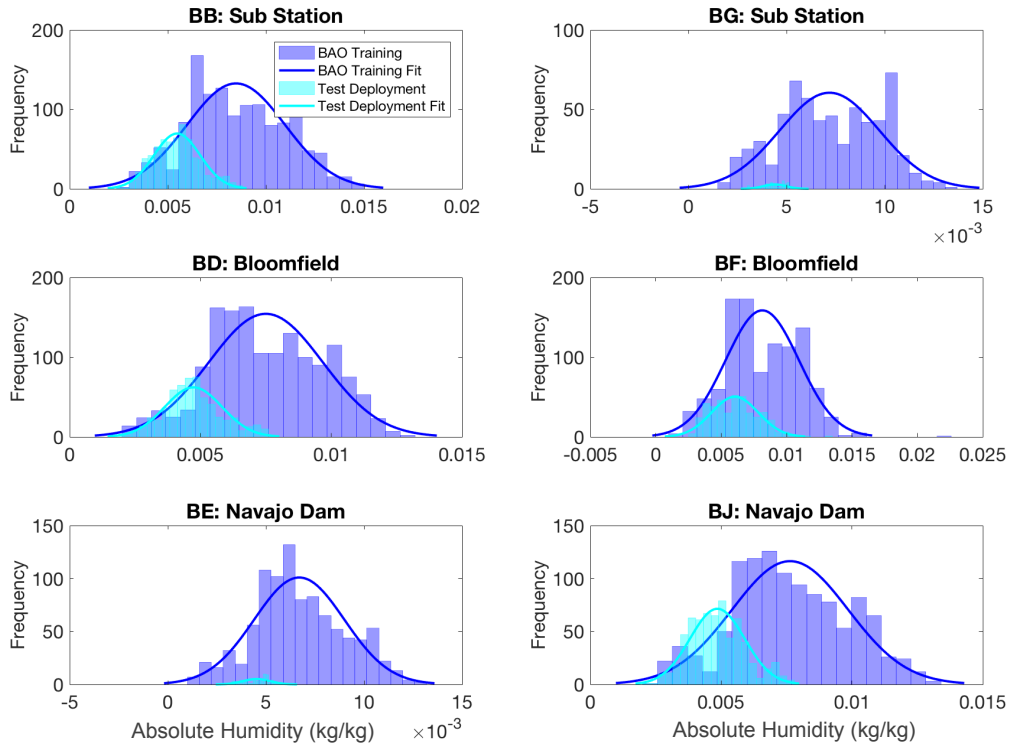

Figure S10 Humidity during BAO training and spring SJ Basin test. Note that due to significant amounts of missing data, U-Pods BG and BE were not included in the analysis.

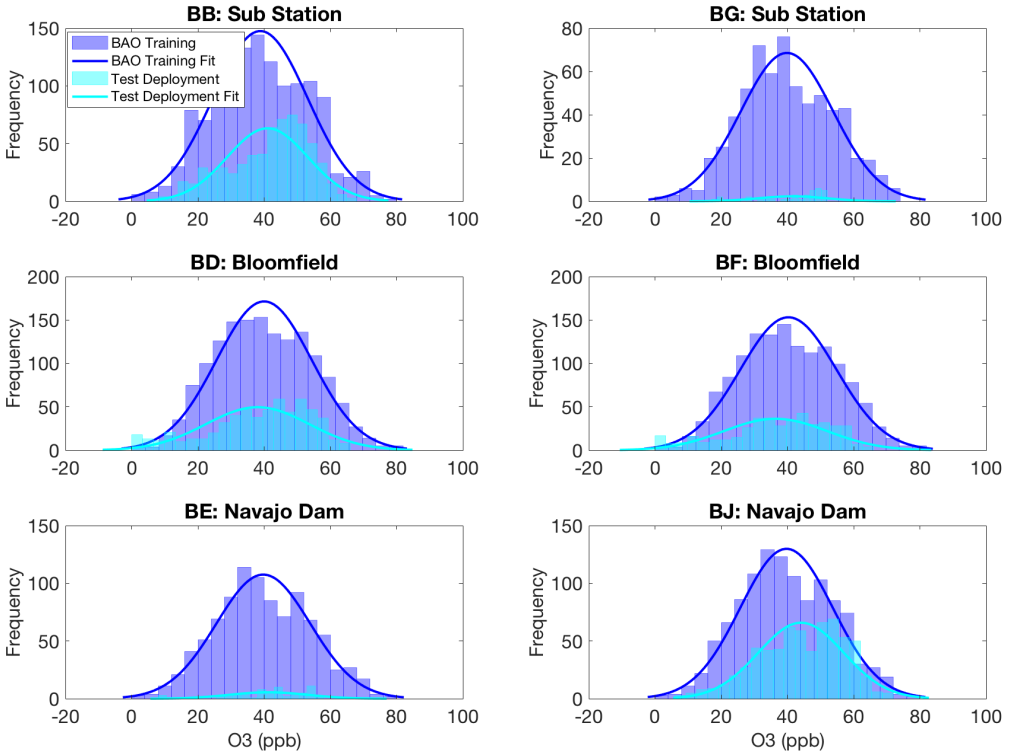

Figure S11 O_3 during BAO training and spring SJ Basin testing. Note that due to significant amounts of missing data, U-Pods BG and BE were not included in the analysis.

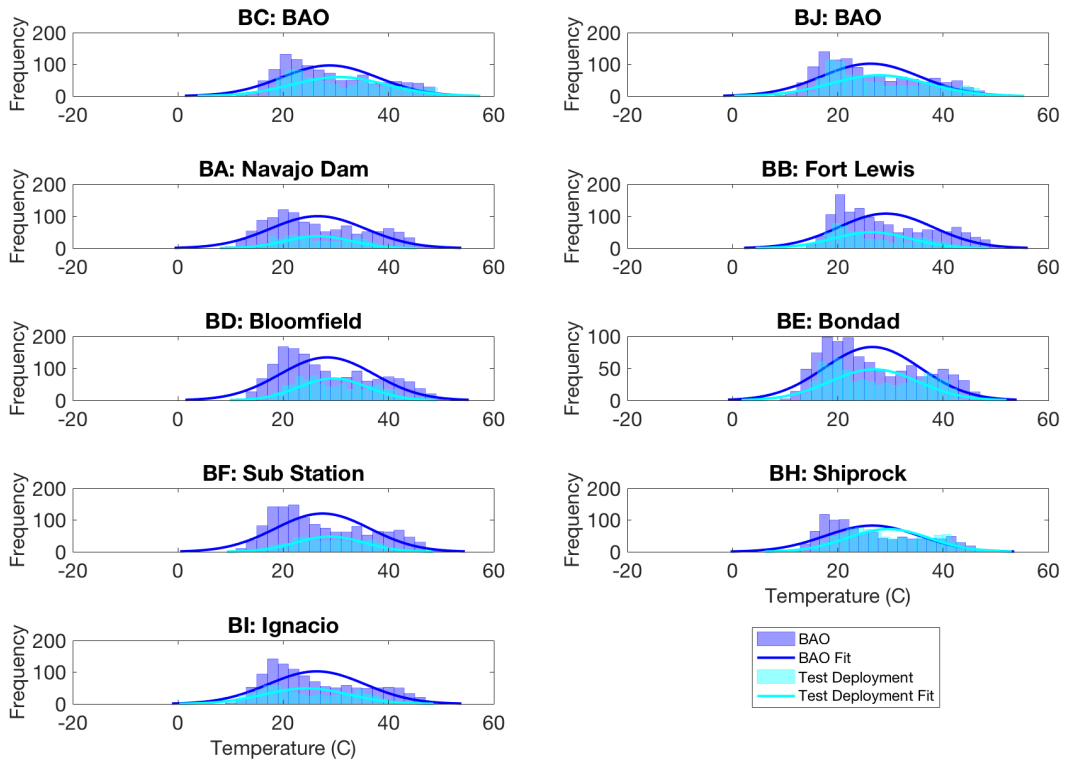

Figure S12 Temperatures during BAO training and summer BAO and SJ Basin testing

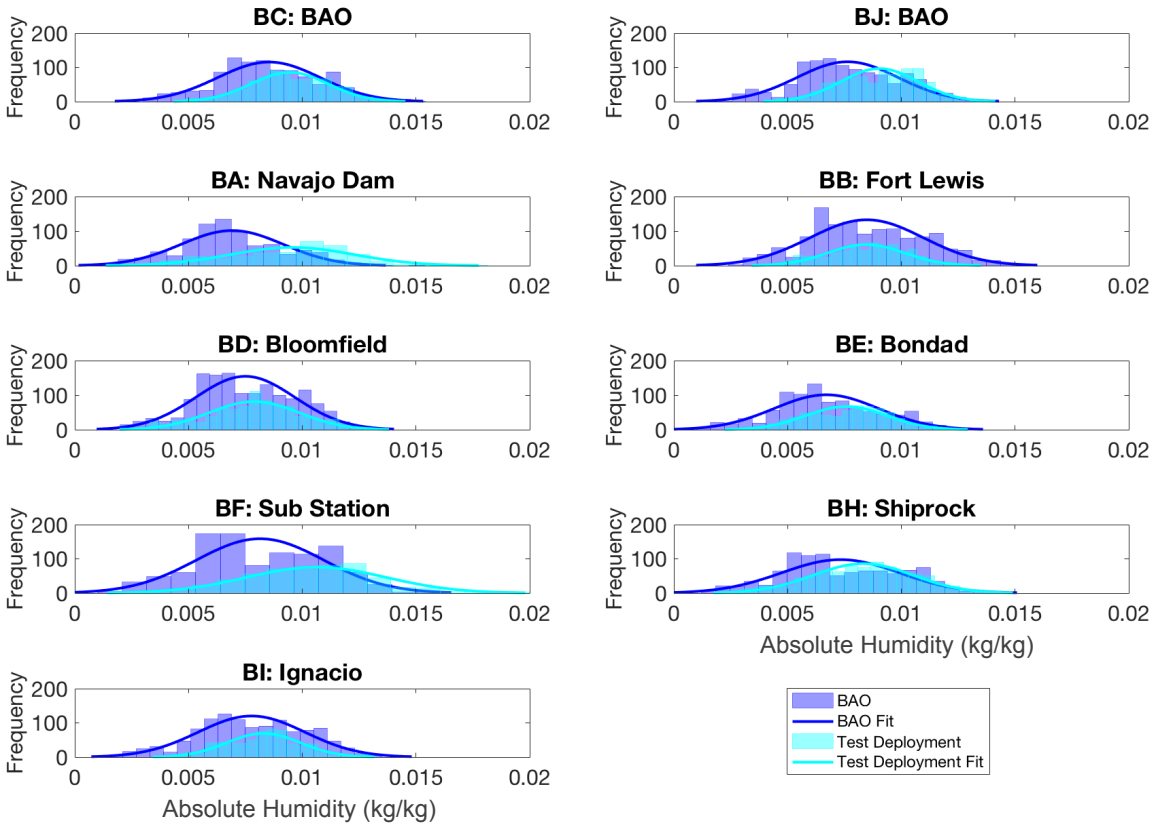

Figure S13 Humidity during BAO training and summer BAO and SJ Basin testing

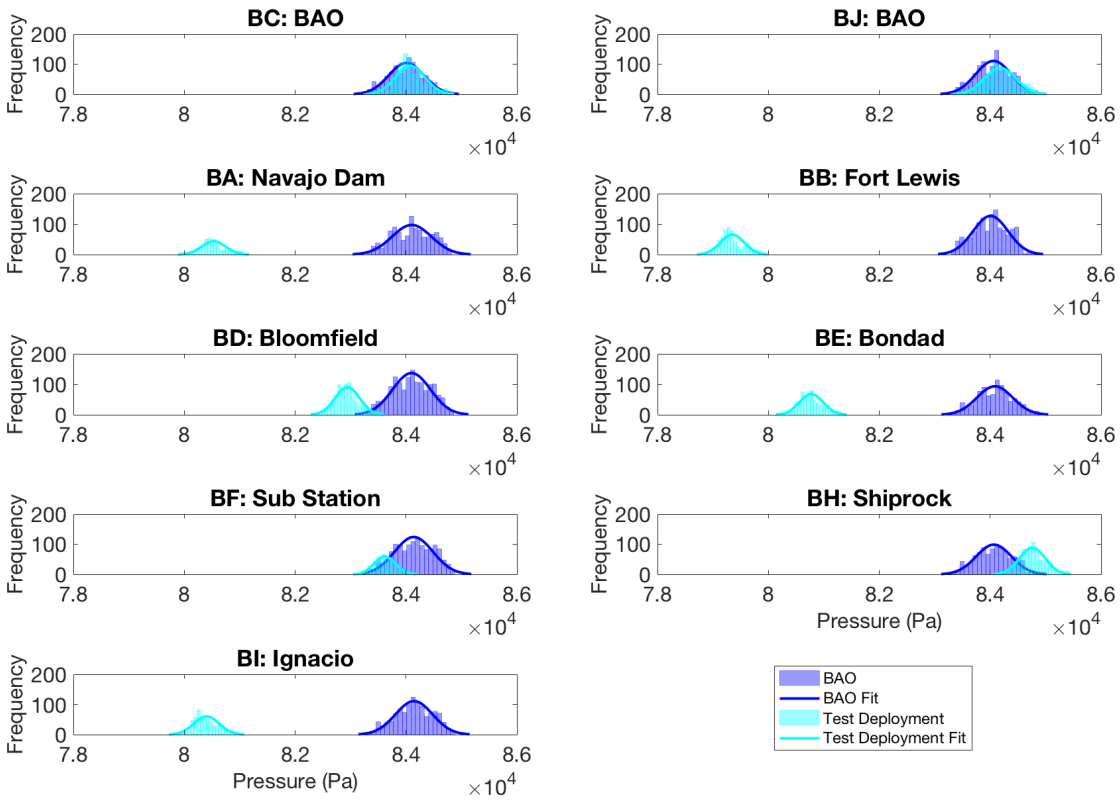

Figure S14 Pressure during BAO training and summer BAO and SJ Basin testing

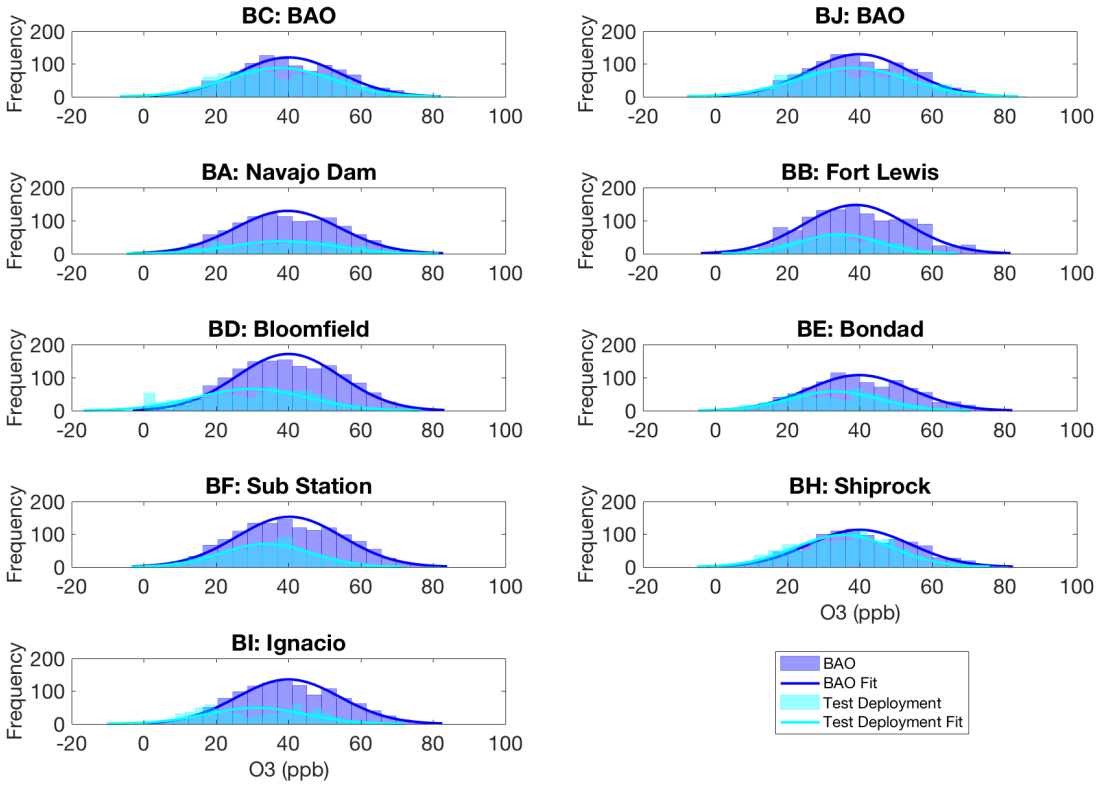

Figure S15 O_3 during BAO training and summer BAO and SJ Basin testing

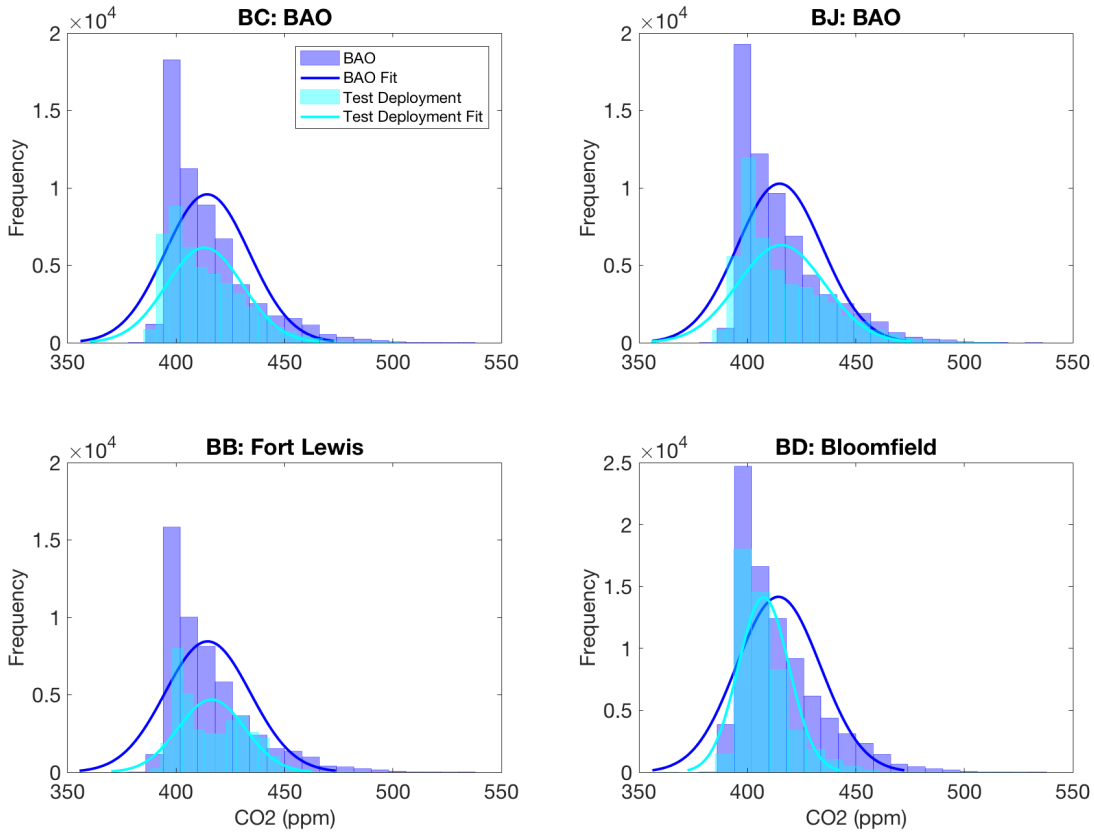

Figure S16 CO_2 during BAO training and summer BAO and SJ Basin testing

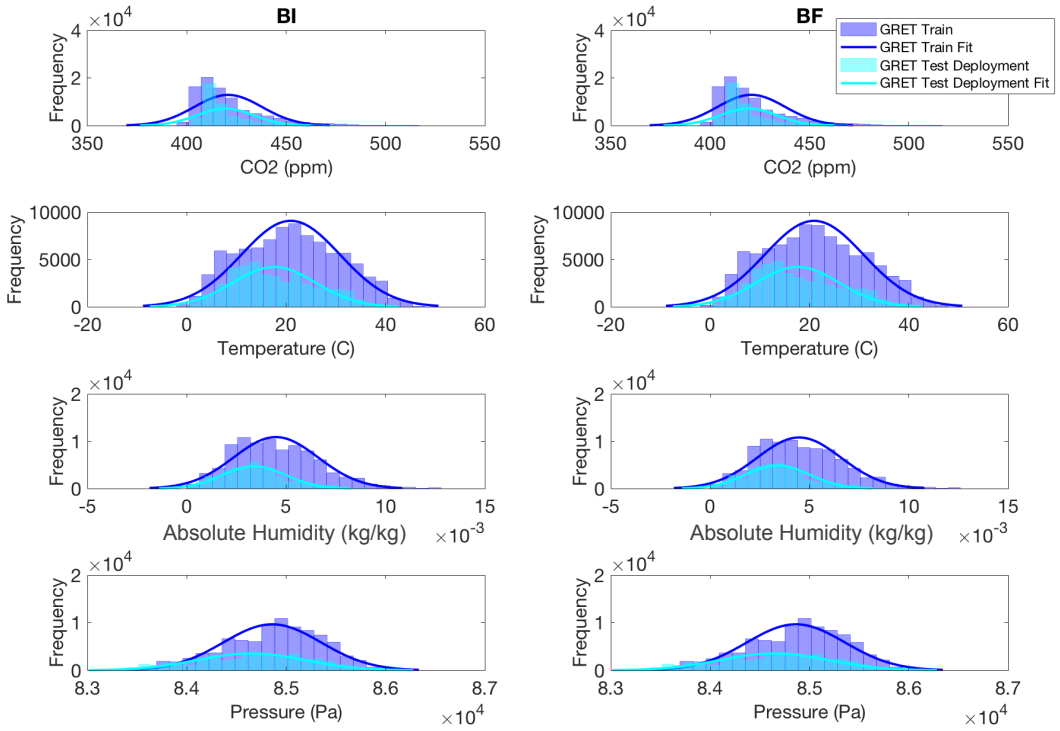

Figure S17 Variable space during GRET spring CO_2 training and GRET spring testing.

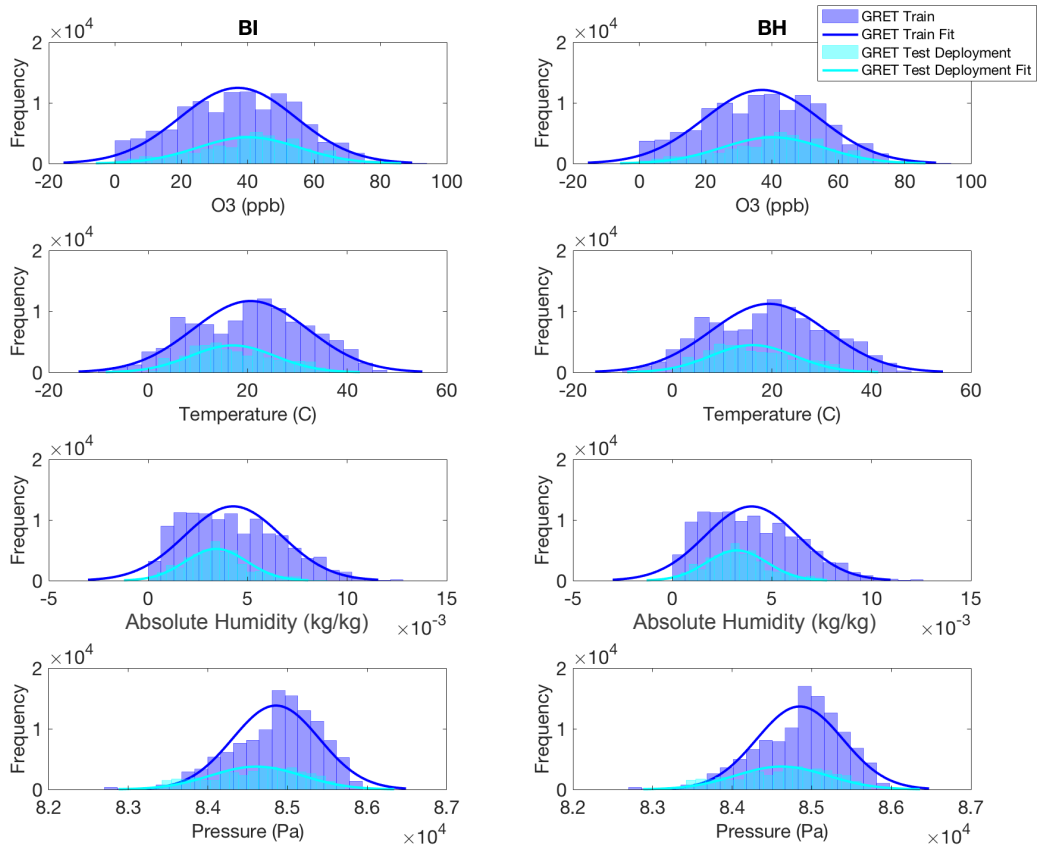

Figure S18 Variable space during GRET spring O_3 training and GRET spring testing.

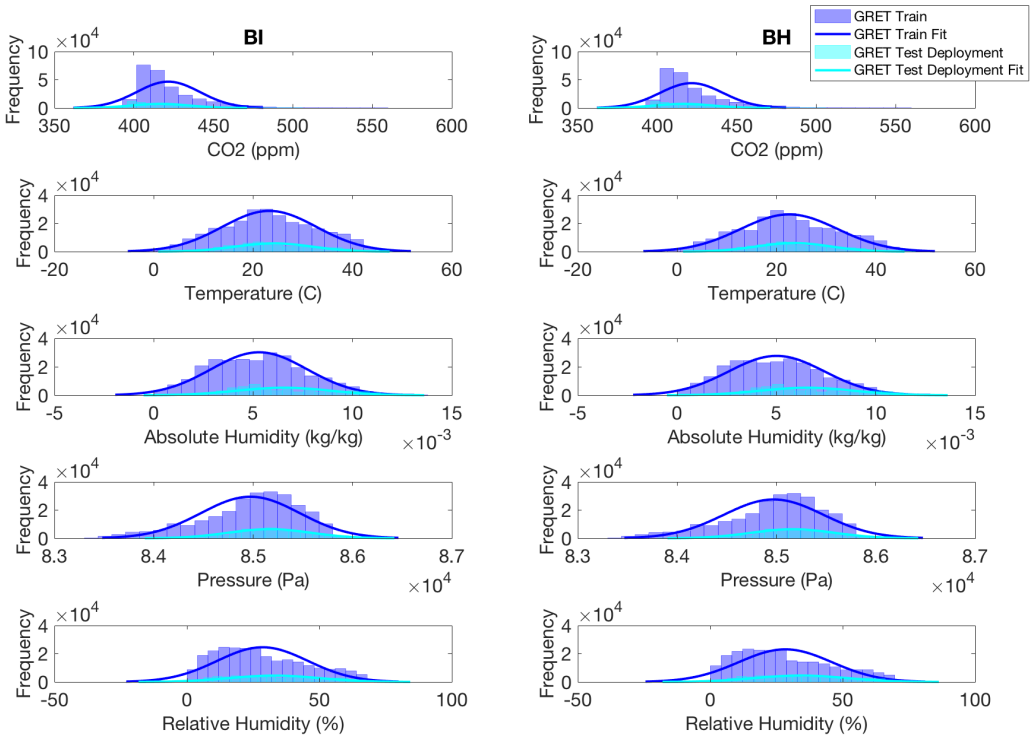

Figure S19 Variable space during 2017 GRET CO_2 training and GRET fall 2016 testing.

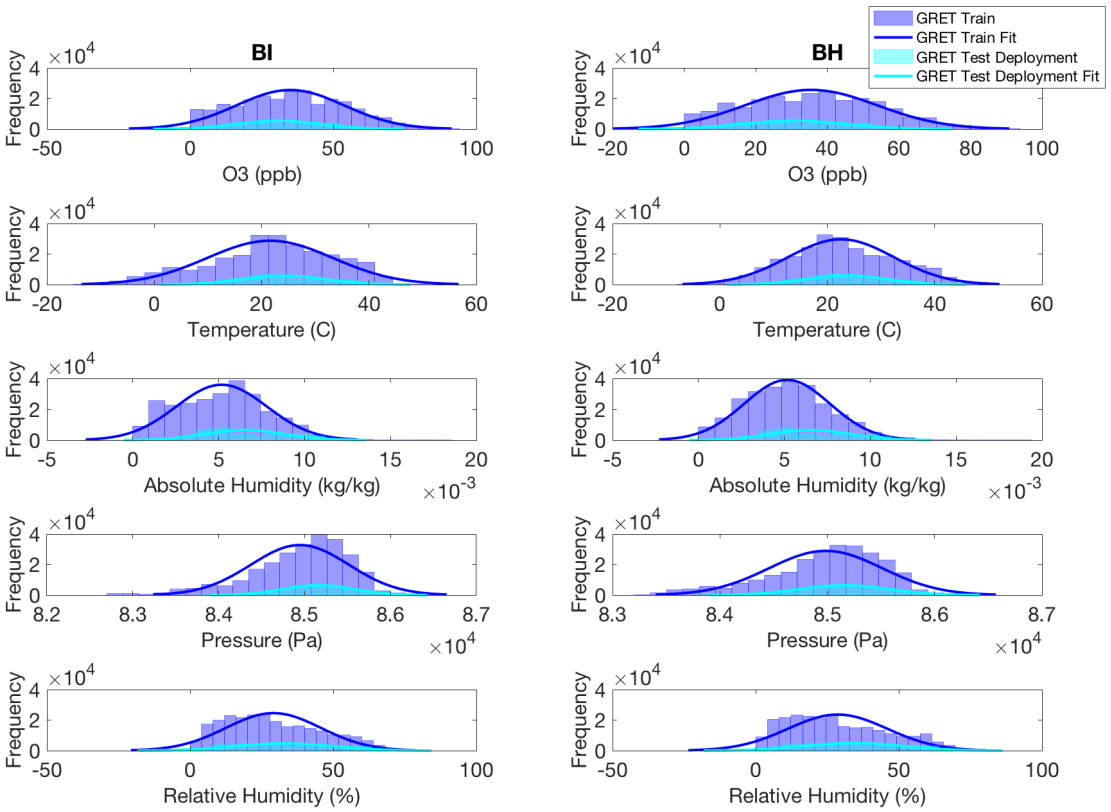

Figure S20 Variable space during 2016-2017 GRET O_3 training and GRET fall 2016 testing.

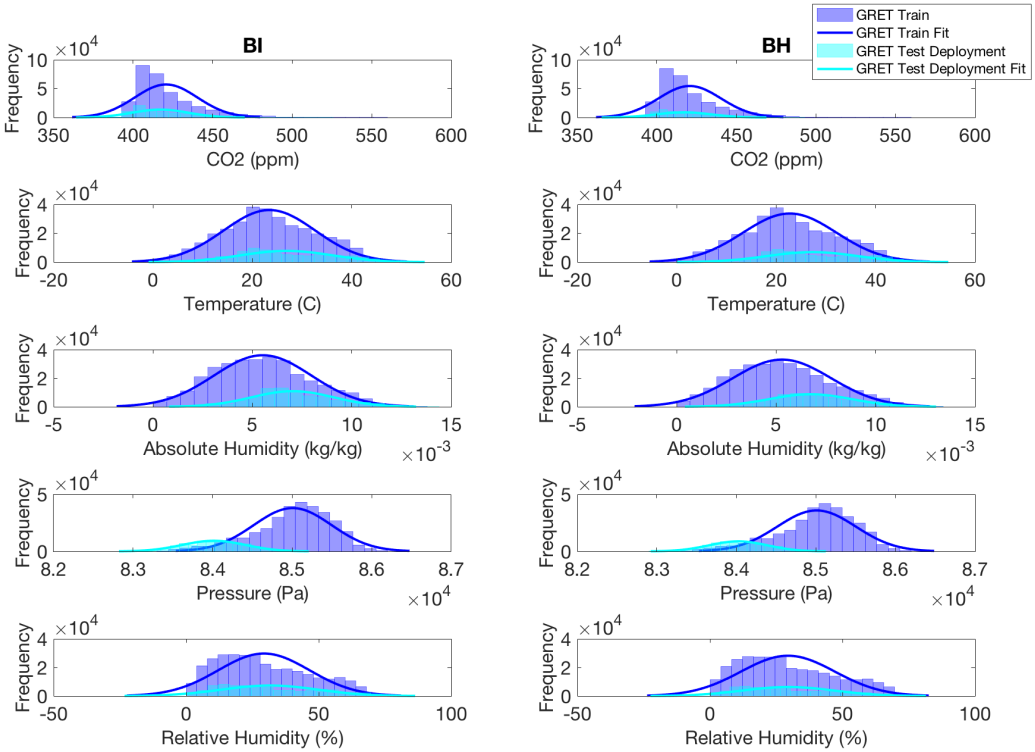

Figure S21 Variable space during 2017 GRET CO_2 training and BAO summer 2016 testing.

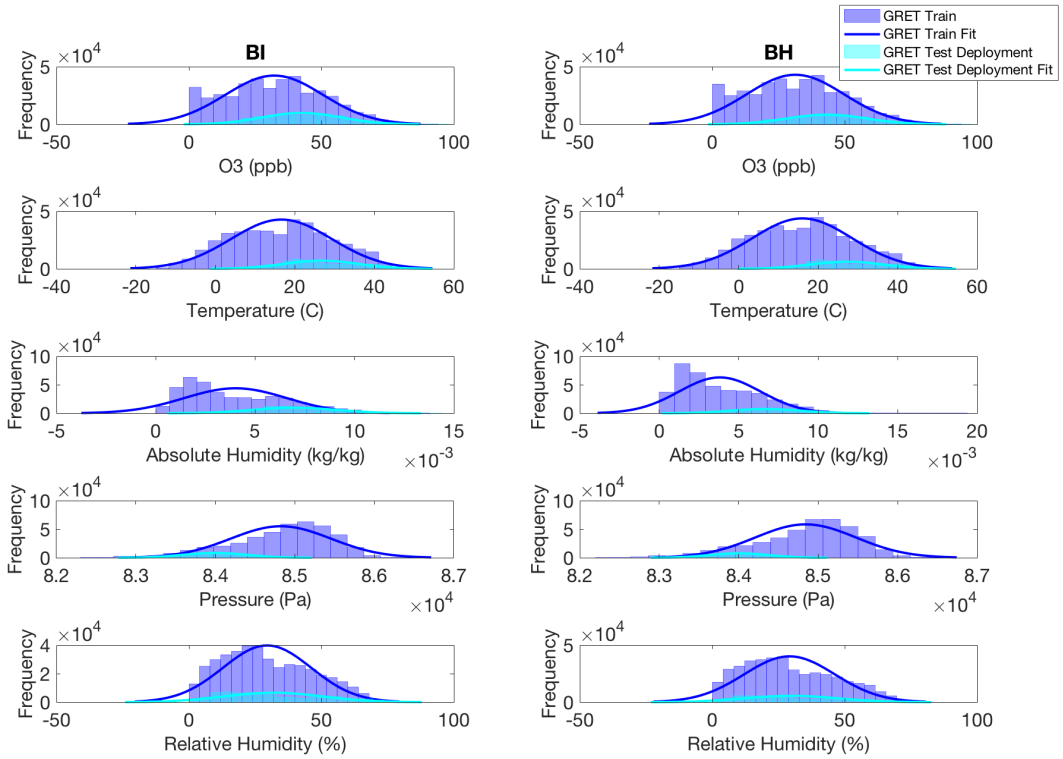

Figure S22 Variable space during 2016-2017 GRET O_3 training and BAO summer 2016 testing.

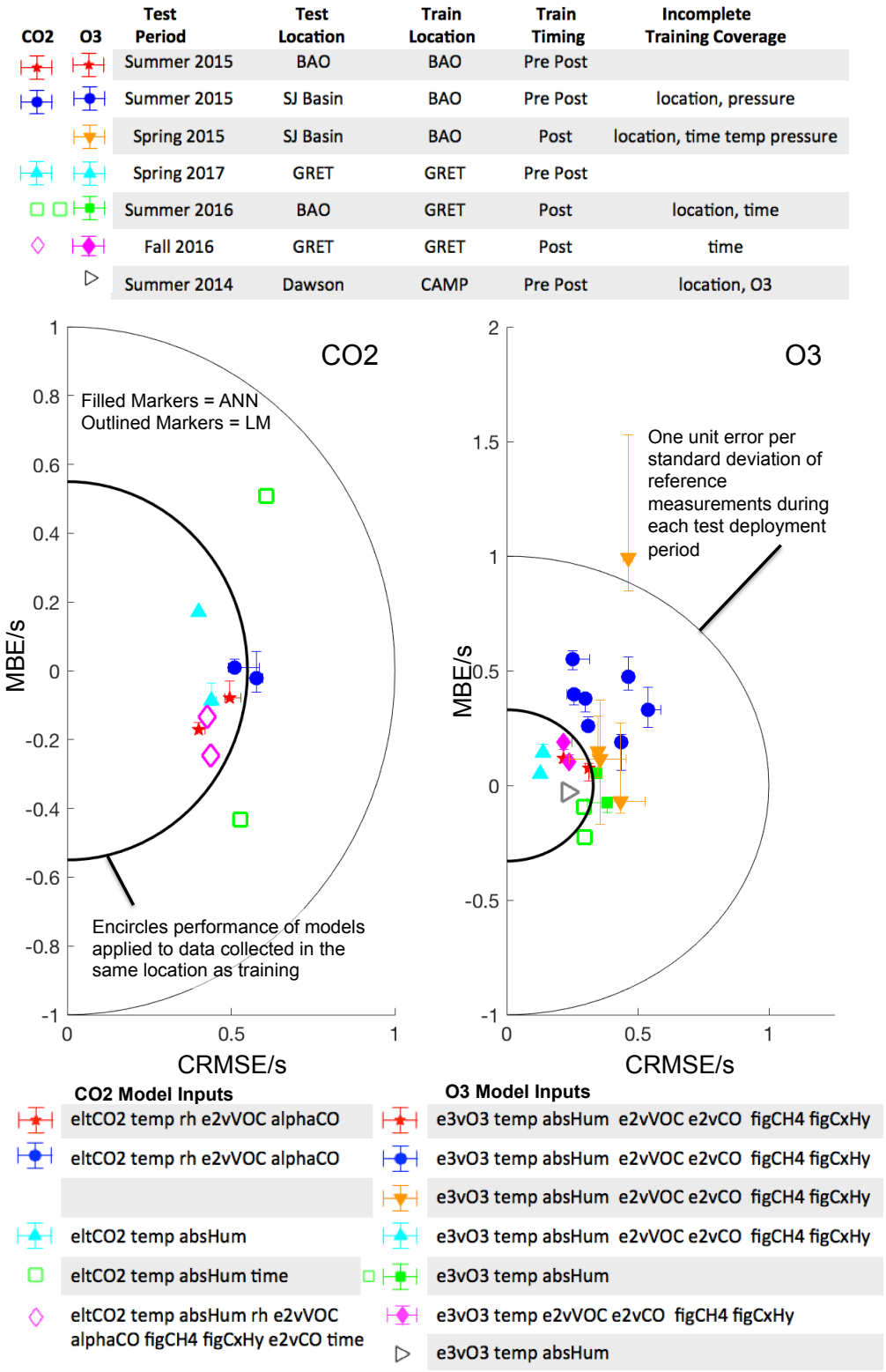

Figure S23 Target diagrams for CO_2 and O_3 calibration model performance when tested on data from a number of field deployments. MBE and CRMSE on the left and bottom axis (black) are normalized by s, the standard deviation of reference measurements during individual test periods.

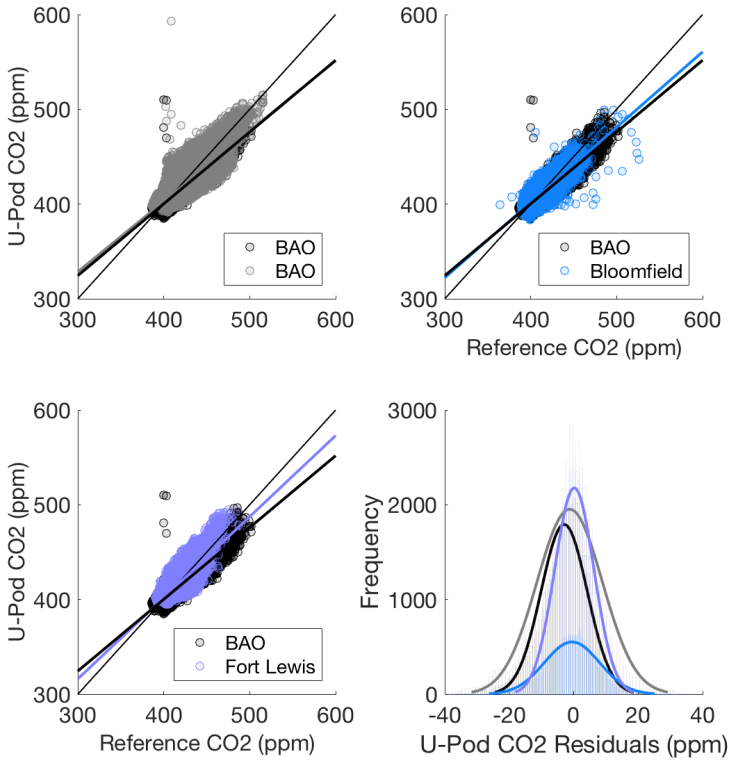

Figure S24 Scatter plots of U-Pod CO_2 vs. reference CO_2 and overlaid histograms of U-Pod CO_2 residuals. A 1:1 single-weight reference line is included in each scatter plot along with double-weight lines of best fit for U-Pods at each sampling location. Data from U-Pod BC at BAO is plotted in black along with U-Pods BJ, BB, and BD at BAO, Fort Lewis, and Bloomfield, respectively. (sensor signal inputs: eltCO2, temp, rh, e2vVOC, and alphaCO)

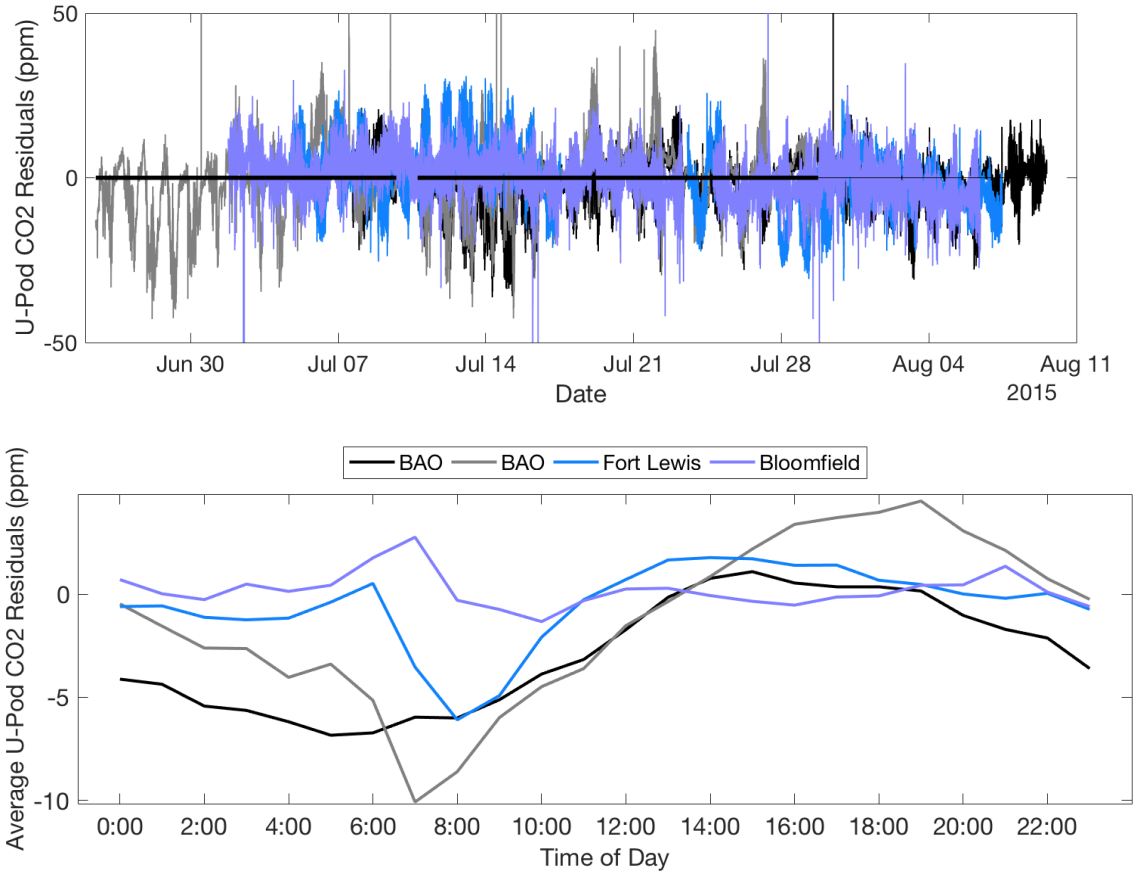

Figure S25 U-Pod CO_2 residuals by of time of day and throughout the duration of the deployment period (sensor signal inputs: eltCO2, temp, rh, e2vVOC, and alphaCO)