Supplement of Atmos. Meas. Tech., 14, 5809–5821, 2021
https://doi.org/10.5194/amt-14-5809-2021-supplement

Atmospheric Measurement Techniques
Open Access

Supplement of

SIBaR: a new method for background quantification and removal from mobile air pollution measurements

Blake Actkinson et al.

Correspondence to: Robert J. Griffin (rob.griffin@rice.edu)

Census Tracts	Population Total	# Metal Recyclers	# Concrete Batch Plants	# Petrochemical Facilities
Northwest Domain	**34873**	**7**	**2**	**0**
North Spring Branch	5126	0	0	0
South Spring Branch	3604	0	0	0
Memorial Park	6908	0	0	0
Washington Corridor	5432	2	0	0
North River Oaks	1803	0	0	0
South River Oaks	2775	0	0	0
West Eastex	2753	5	2	0
North Heights	6472	1	0	0
Southwest Domain	**24927**	**0**	**1**	**0**
Westchase	5548	0	0	0
Sharpstown	5616	0	0	0
Sharpstown North	3484	0	1	0
Sharpstown South	5196	0	0	0
Bayland Park	5083	0	0	0
South Beltway Central Domain	**2530**	**3**	**8**	**0**
South Beltway Central	2530	3	8	0
Rice Domain	**8247**	**0**	**0**	**0**
North Rice	2892	0	0	0
South Rice	5355	0	0	0
Ship Channel Domain	**20177**	**4**	**1**	**4**
Clinton	2127	2	1	1
West Galena Park	5245	0	0	0
East Galena Park	3000	0	0	0
Manchester	1647	0	0	1
Harrisburg	1496	2	0	2
Milby Park	6662	0	0	0

Table S1. Neighborhood summary table which includes the total population and number of important point source emitters within each neighborhood. Data courtesy of the U.S. Census and Environmental Defense Fund (Census 2010; Environmental Defense Fund). Domain names refer to direction from downtown or local landmarks (such as Rice University).

Parameter	Instrument	Bias	Precision	Minimum Detection Limit	Measurement Technique
NO	T200 NO Analyzer	± 10 %	± 3 %	4 ppbv	Chemiluminescence
NO_2	T500U NO_2 Analyzer	± 10%	± 1 %	0.95 ppbv	Chemiluminescence
CO_2	Li-COR CO_2 Analyzer	± 1 %	< 0.05 %	N/A	Spectroscopy

Table S2. Instruments used in campaign to make measurements.

	1	2	3	4	5	6	7	8	9	10	11	12
1	0.000	0.082	0.073	0.039	0.027	0.000	0.066	0.037	0.067	0.064	0.062	0.078
2	0.000	0.000	0.042	0.071	0.083	0.082	0.053	0.074	0.053	0.052	0.053	0.037
3	0.000	0.000	0.000	0.063	0.072	0.073	0.038	0.063	0.036	0.039	0.038	0.054
4	0.000	0.000	0.000	0.000	0.034	0.039	0.051	0.049	0.056	0.055	0.056	0.062
5	0.000	0.000	0.000	0.000	0.000	0.027	0.063	0.036	0.068	0.070	0.068	0.075
6	0.000	0.000	0.000	0.000	0.000	0.000	0.066	0.037	0.067	0.064	0.062	0.078
7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.072	0.025	0.029	0.023	0.041
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.072	0.074	0.072	0.083
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.027	0.024	0.046
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.044
11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.047
12	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table S3. Pairwise RMSE values between different iterations of SIBaR for NO_x.

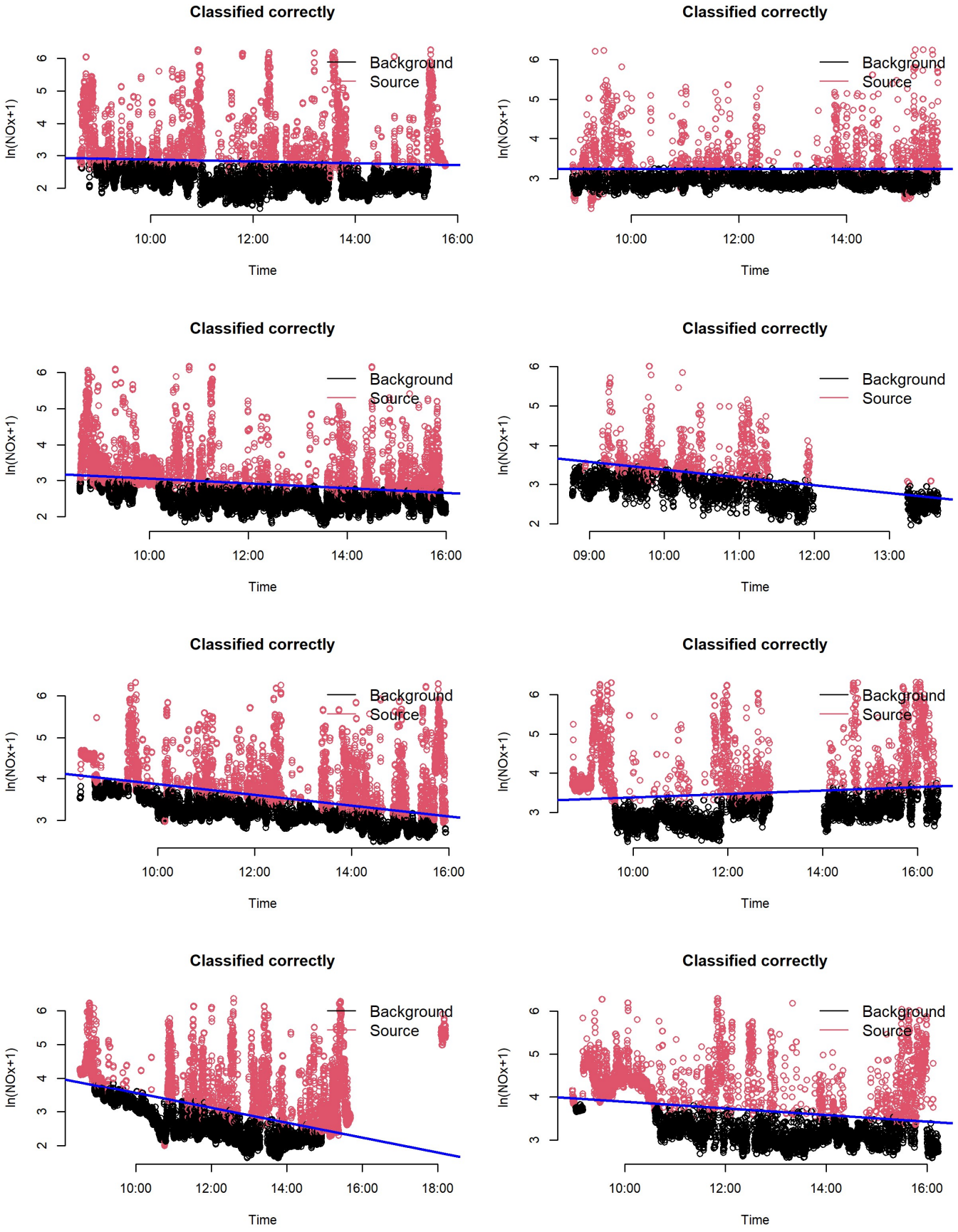

Figure S1. Results from the fitted line classifier. The blue line is the line of best fit for pollutant averages around transition and is assumed to be the line of separation.

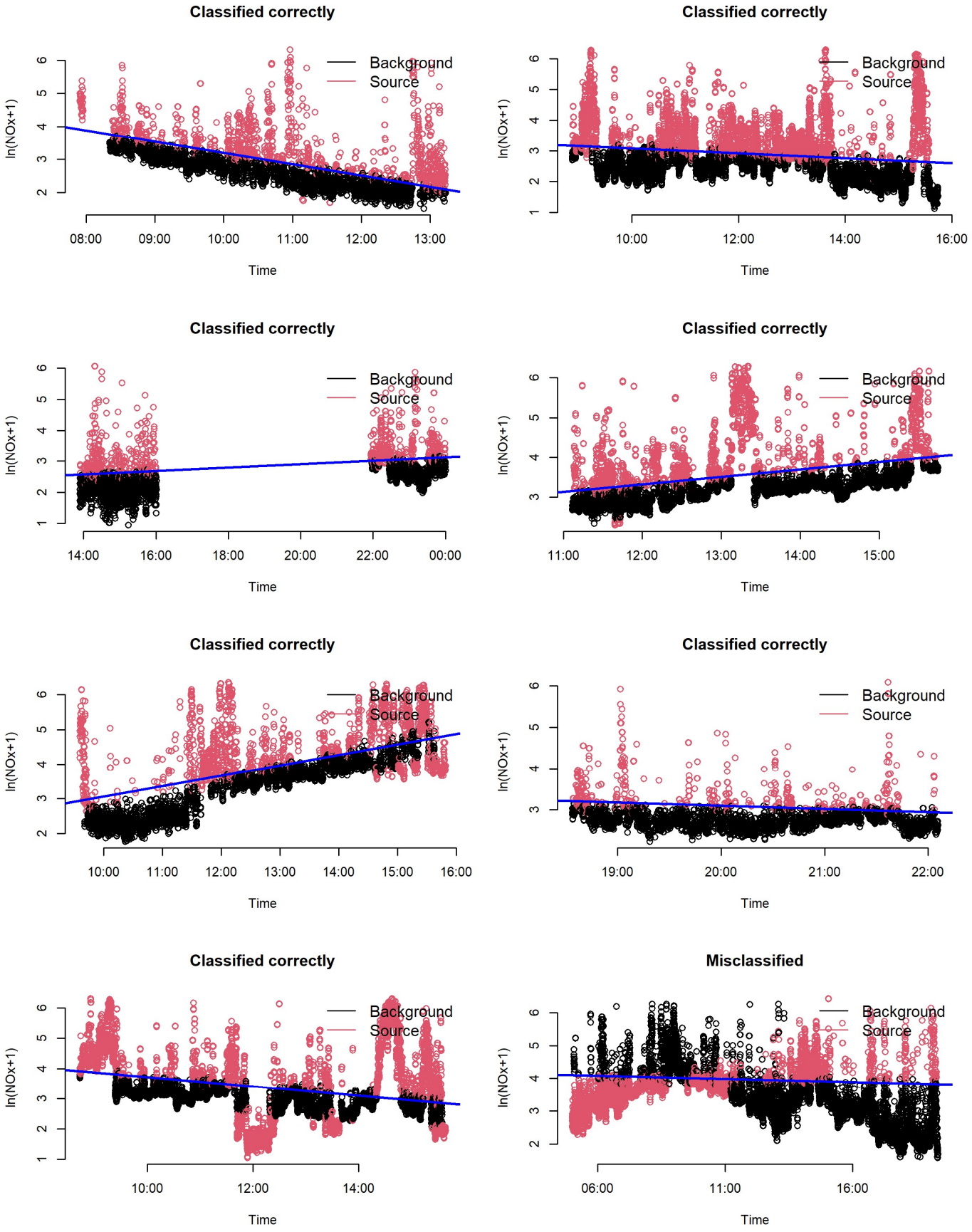

Figure S2. Results from the fitted line classifier. The blue line is the line of best fit for pollutant averages around transition and is assumed to be the line of separation.

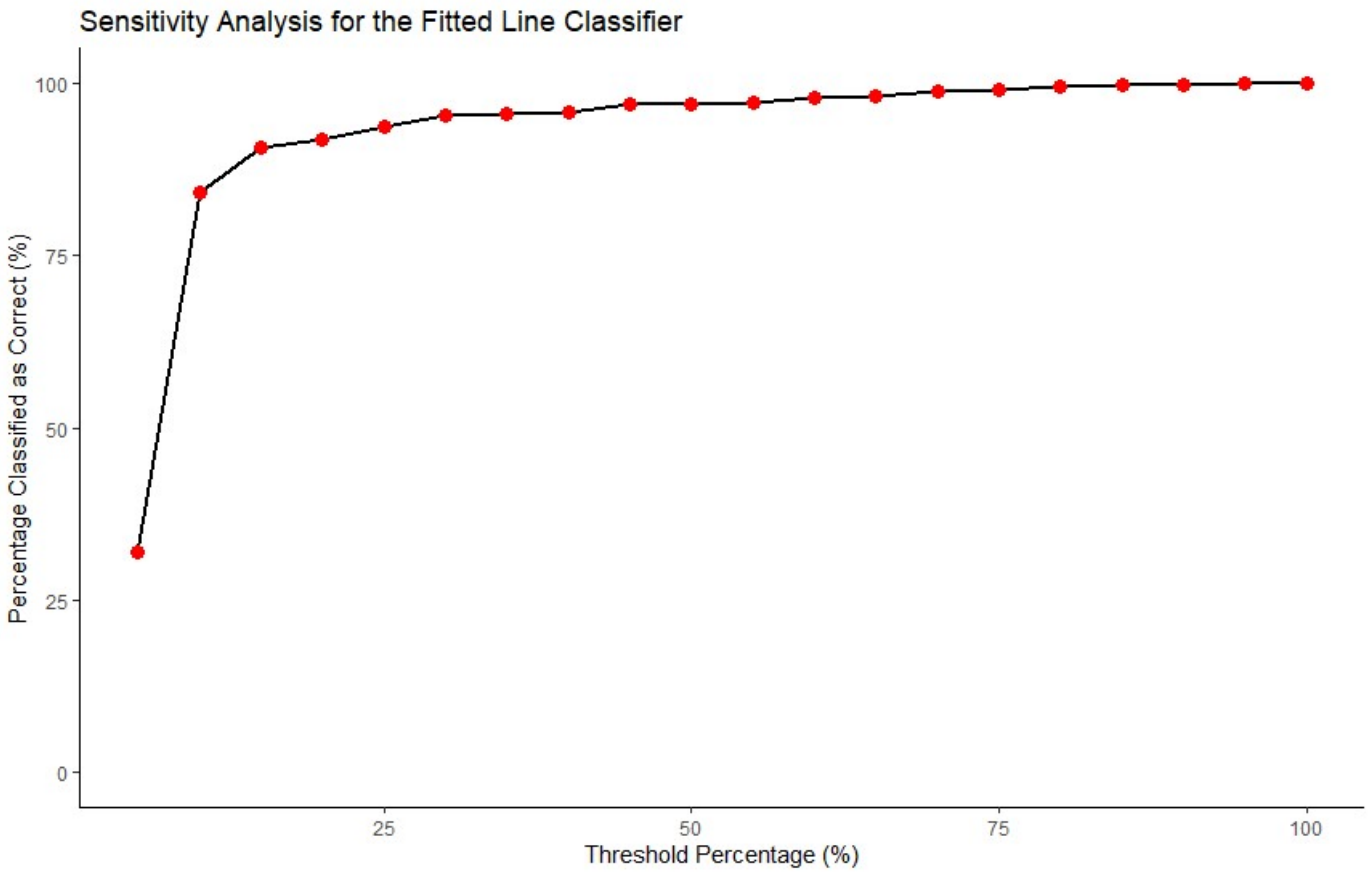

Figure S3. Sensitivity analysis on the threshold percentage of the fitted line classifier.

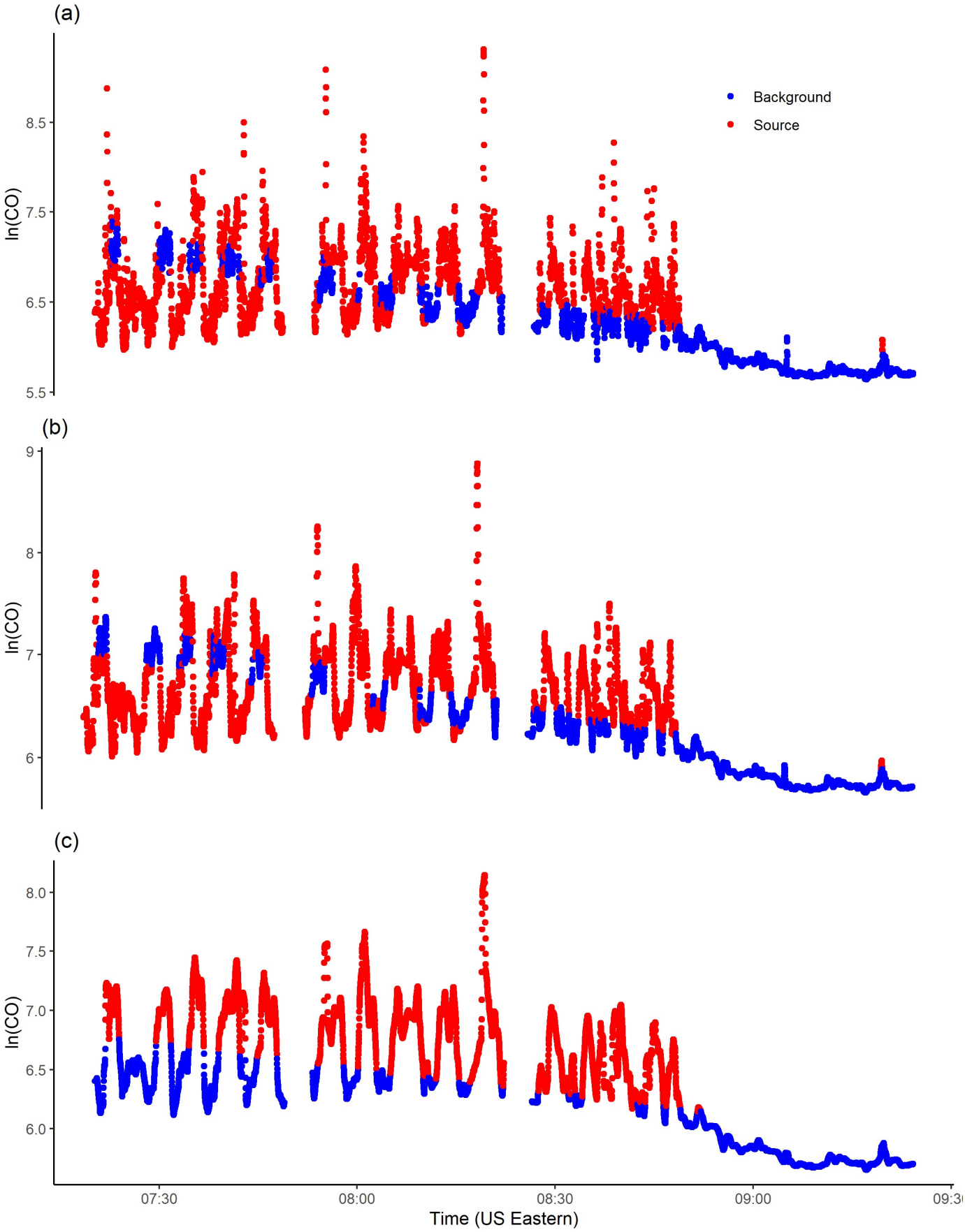

Figure S4. SIBaR state designation time window sensitivity. (a) No smoothing applied. (b) 10 s smoothing applied. (c) 30 s smoothing applied.

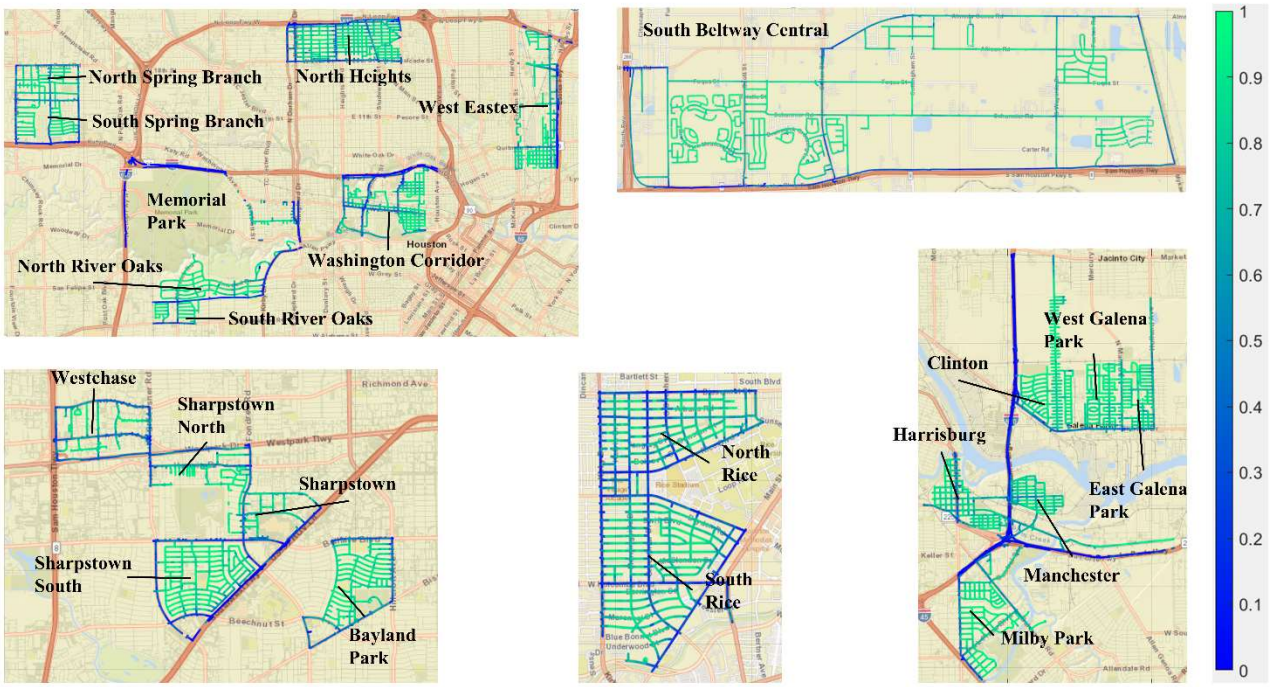

Figure S5. CO_2 mapped fractional contributions. A value of 1 implies all measurements reflect background, while a value of 0 implies all measurements reflect non-background. Basemap generated by Matlab geobasemap 'streets' and is hosted by ESRI (Sources: Esri, DeLorme, HERE, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, Tomtom).

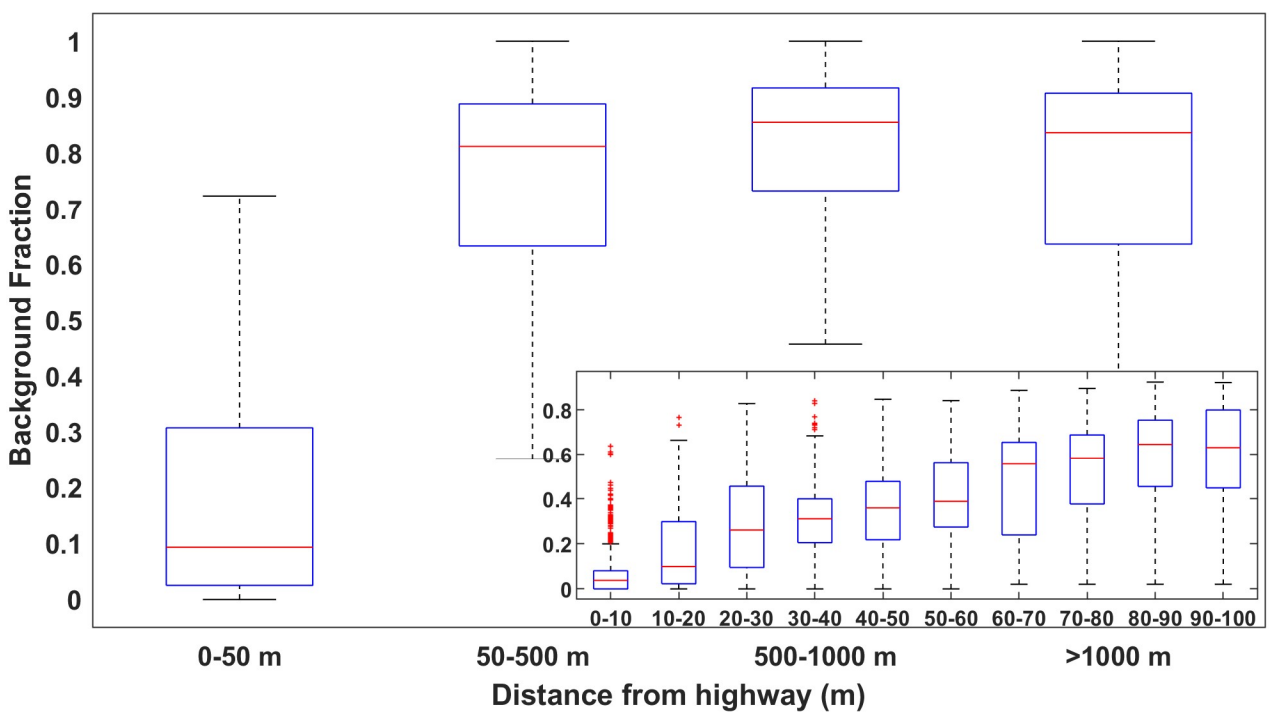

Figure S6. CO_2 background fraction boxplot binned by distance; format identical to Figure 4 in the main manuscript.

Figure S7. Mapped median NO_x source contributions. Northwest quadrant. Basemap generated by Matlab geobasemap 'streets' and is hosted by ESRI (Sources: Esri, DeLorme, HERE, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, Tomtom).

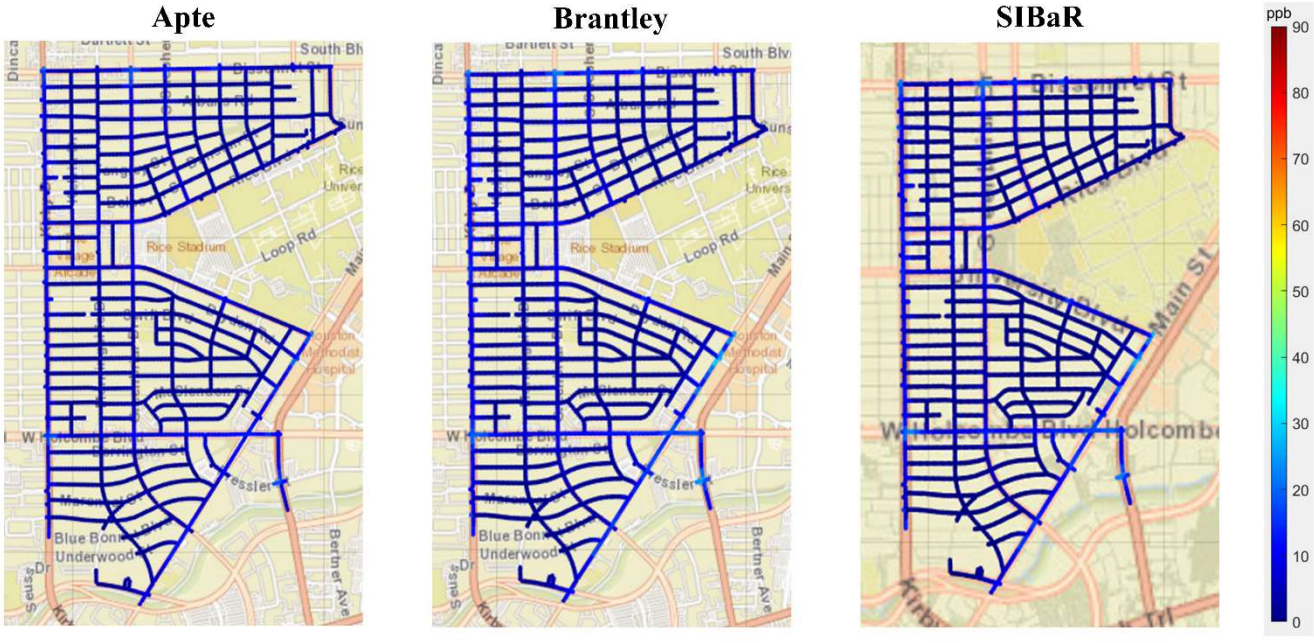

Figure S8. Mapped median NO_x source contributions. North and South Rice. Basemap generated by Matlab geobasemap 'streets' and is hosted by ESRI (Sources: Esri, DeLorme, HERE, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, Tomtom).

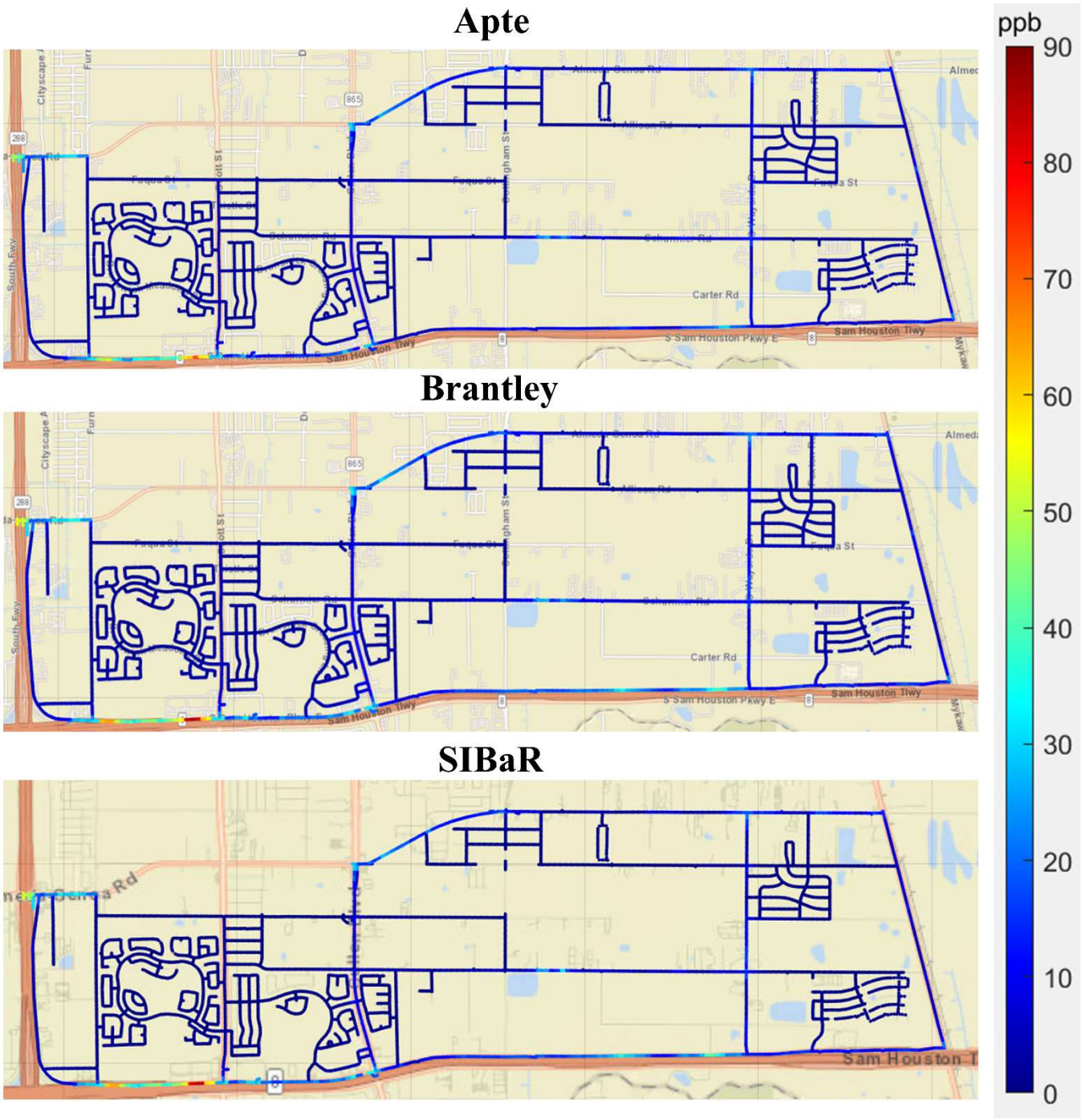

Figure S9. Mapped median NO_x source contributions. South Beltway Central. Basemap generated by Matlab geobasemap 'streets' and is hosted by ESRI (Sources: Esri, DeLorme, HERE, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, Tomtom).

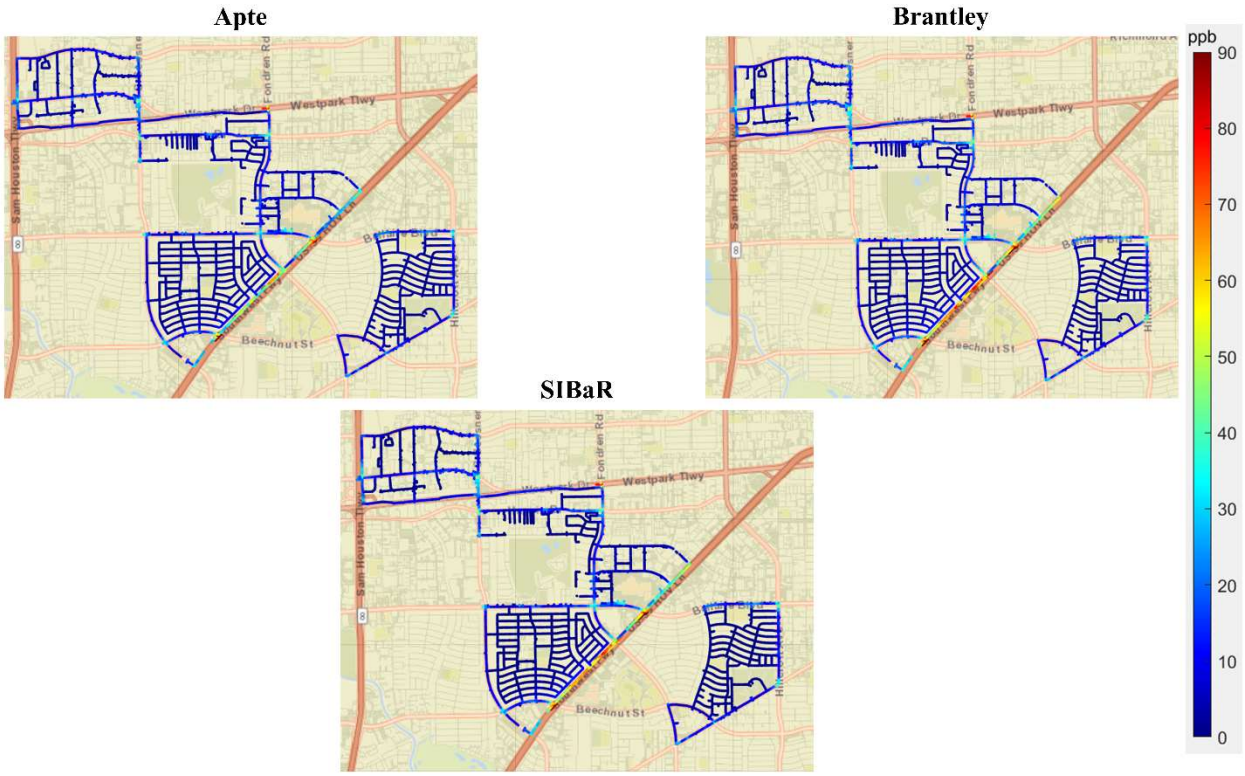

Figure S10. Mapped median NO_x source contributions. Southwest quadrant. Basemap generated by Matlab geobasemap 'streets' and is hosted by ESRI (Sources: Esri, DeLorme, HERE, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, Tomtom).

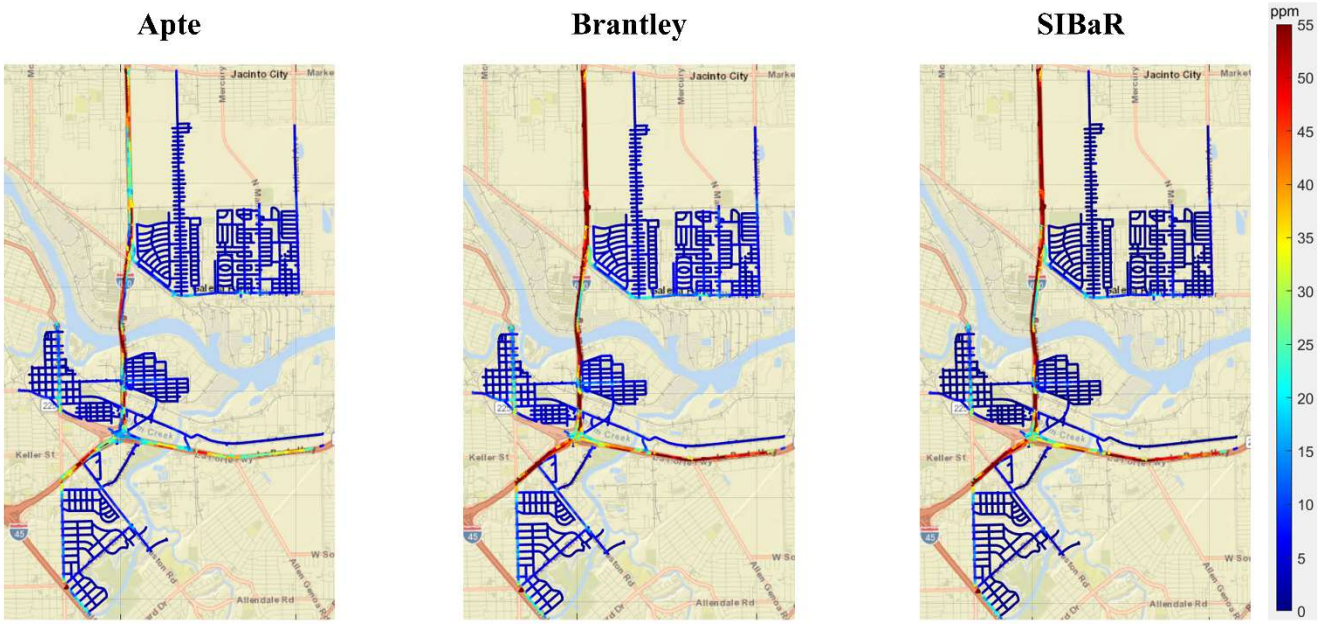

Figure S11. CO_2 mapped median source contributions. Ship Channel. Basemap generated by Matlab geobasemap 'streets' and is hosted by ESRI (Sources: Esri, DeLorme, HERE, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, Tomtom).

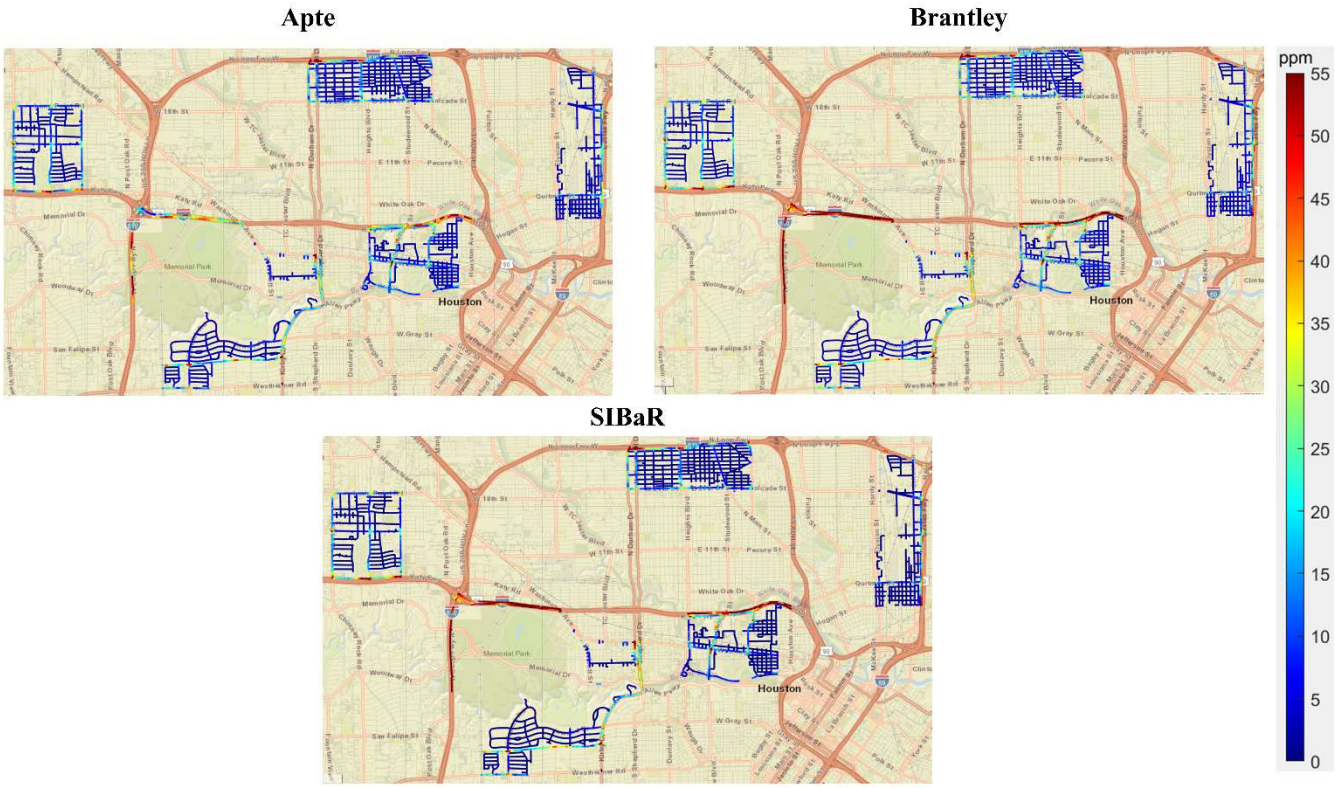

Figure S12. CO_2 mapped median source contributions. Northwest domain. Basemap generated by Matlab geobasemap 'streets' and is hosted by ESRI (Sources: Esri, DeLorme, HERE, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, Tomtom).

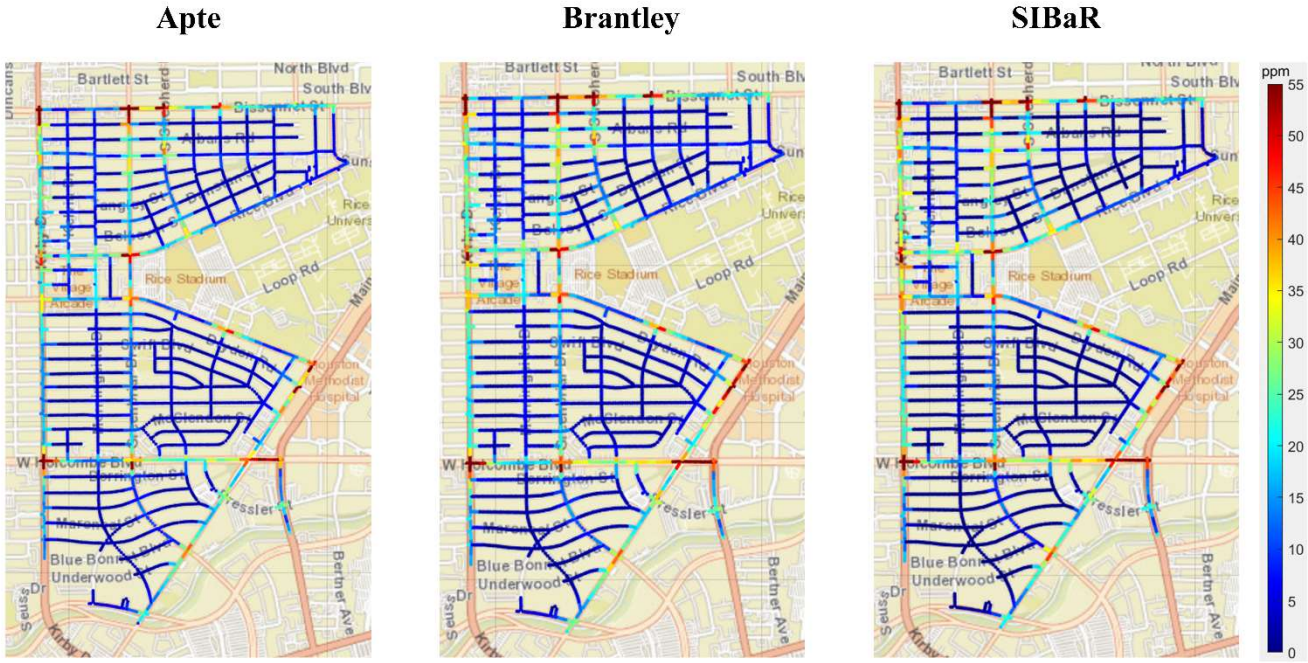

Figure S13. CO_2 mapped median source contributions. North and South Rice. Basemap generated by Matlab geobasemap 'streets' and is hosted by ESRI (Sources: Esri, DeLorme, HERE, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, Tomtom).

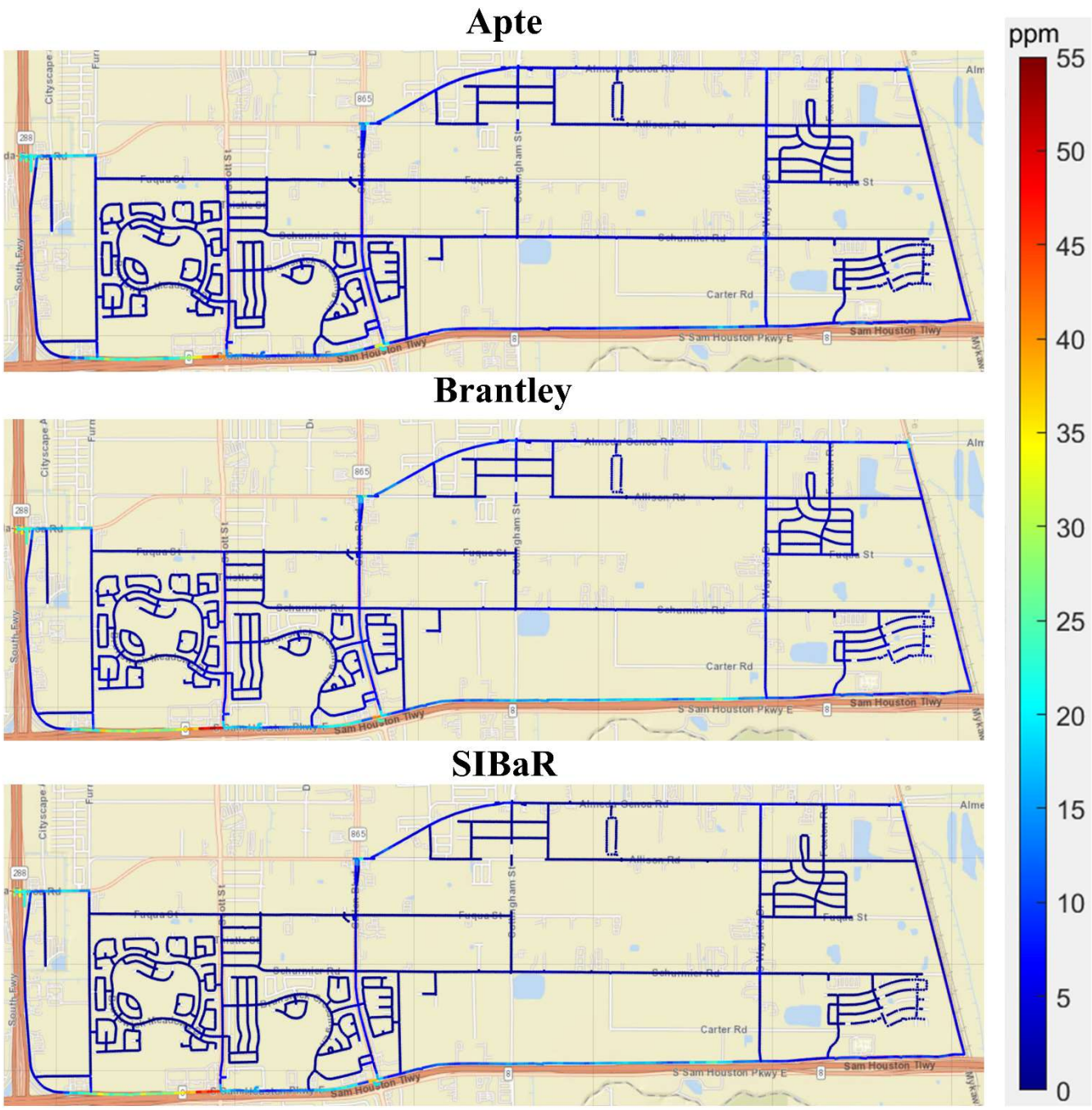

Figure S14. CO_2 mapped median source contributions. South Beltway Central. Basemap generated by Matlab geobasemap 'streets' and is hosted by ESRI (Sources: Esri, DeLorme, HERE, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, Tomtom).

Figure S15. CO_2 mapped median source contributions. Southwest quadrant. Basemap generated by Matlab geobasemap 'streets' and is hosted by ESRI (Sources: Esri, DeLorme, HERE, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, Tomtom).

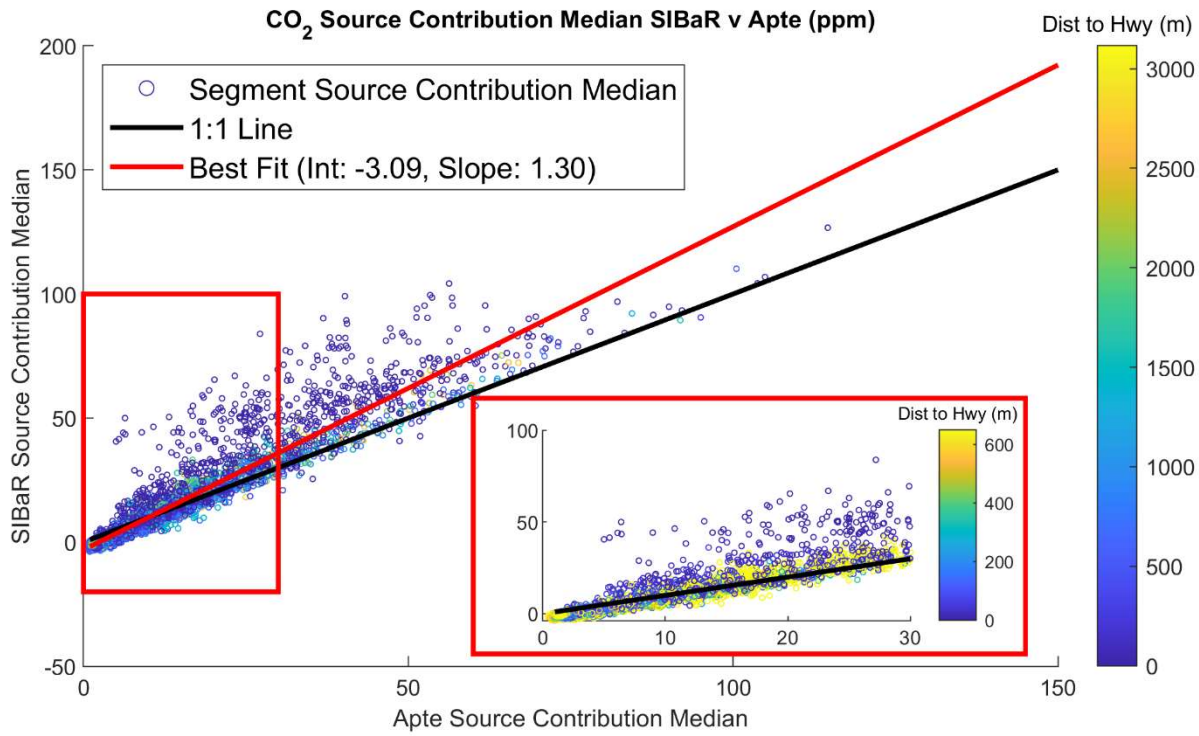

Figure S16. Comparison of CO_2 Road Segment Median Source Contributions for SIBaR and Apte techniques. Points are colored by distance from highway.

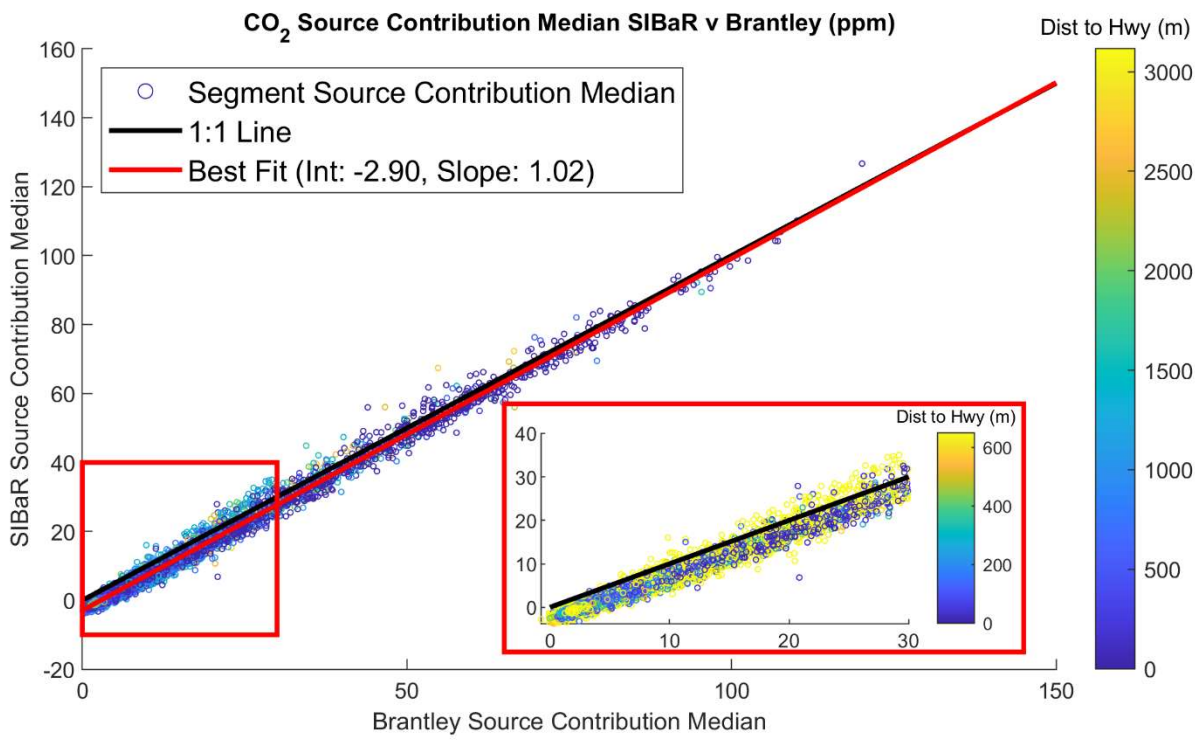

Figure S17. Comparison of CO_2 Road Segment Median Source Contributions for SIBaR and Brantley techniques. Points are colored by distance from highway.

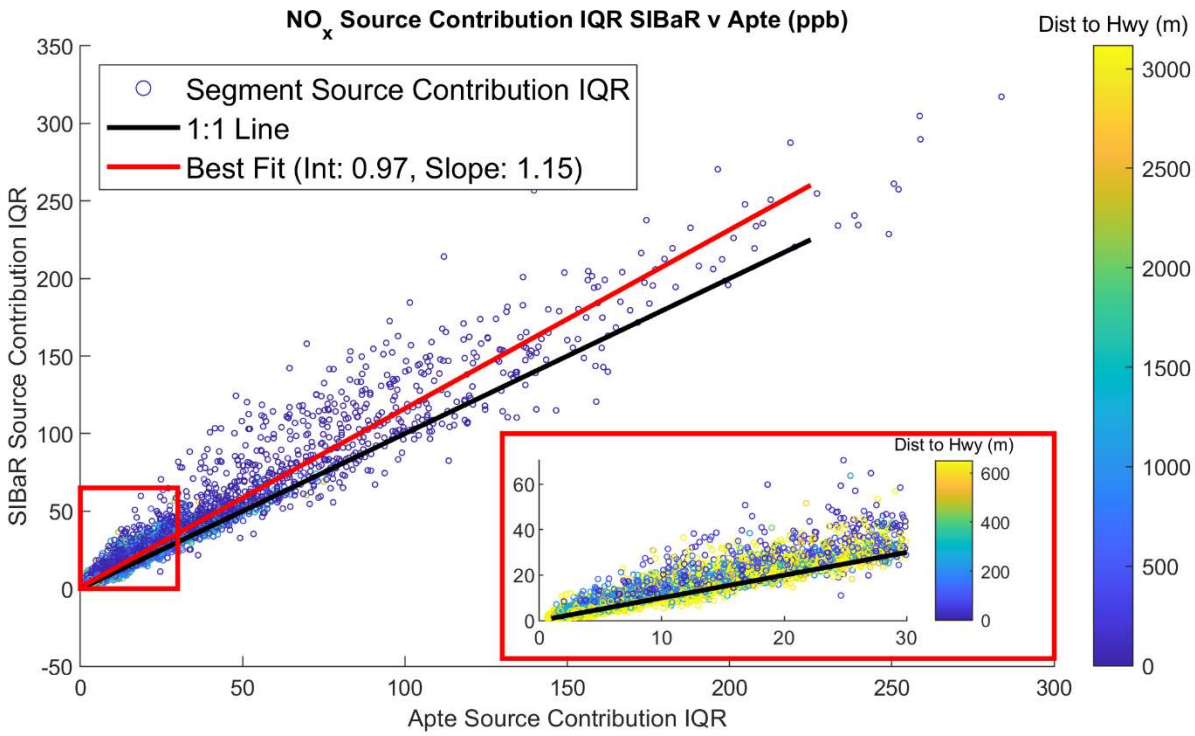

Figure S18. Comparison of NO_x Road Segment IQR Source Contribution for SIBaR and Apte. Points are colored by distance from highway.

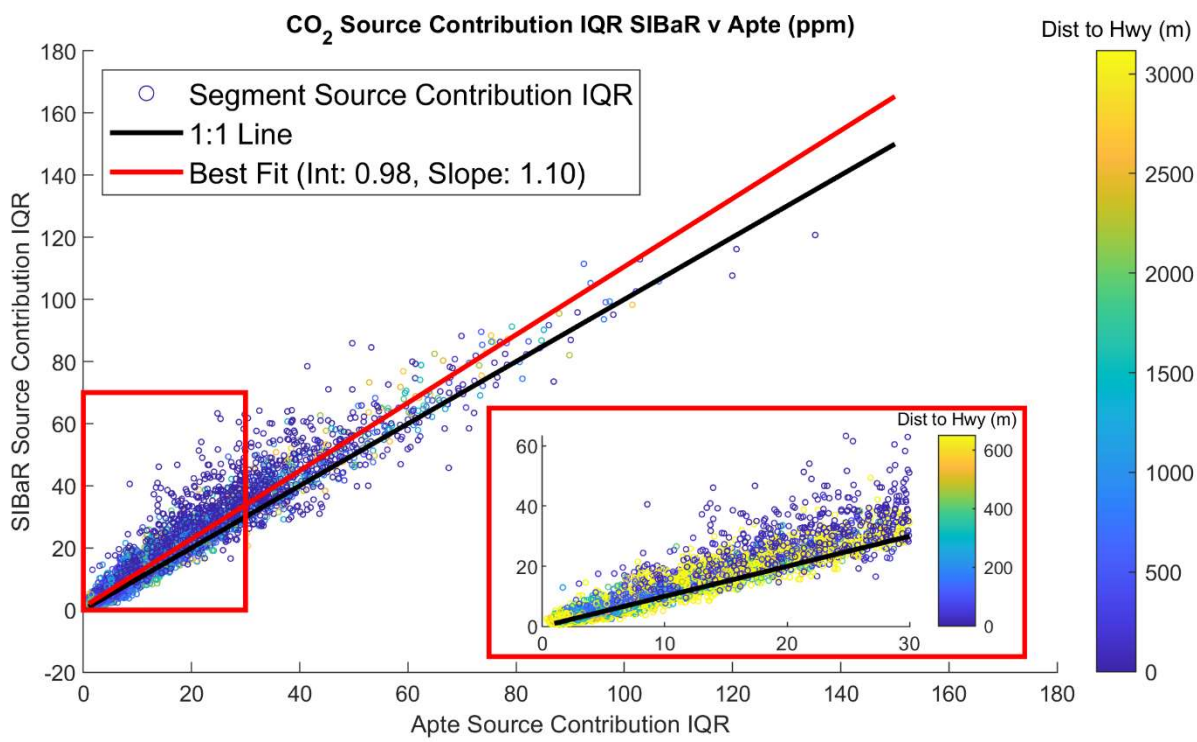

Figure S19. Comparison of CO_2 Road Segment IQR Source Contribution for SIBaR and Apte. Points are colored by distance from highway.

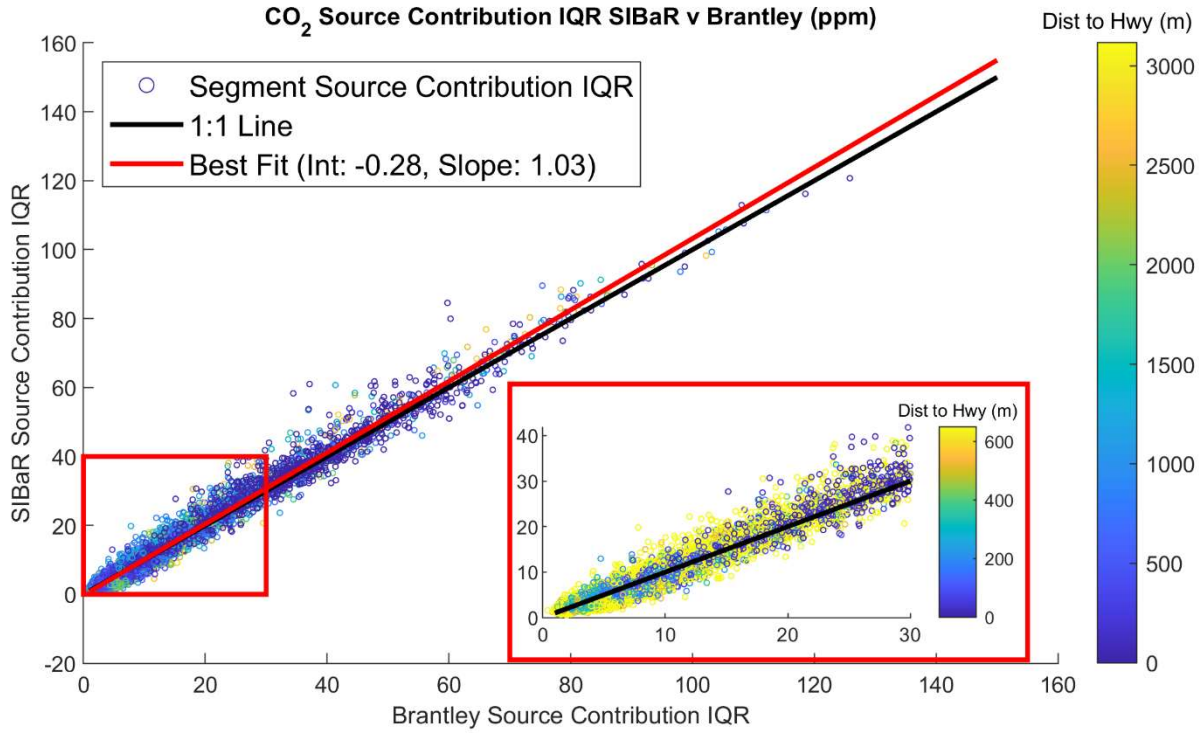

Figure S20. Comparison of CO_2 Road Segment IQR Source Contribution for SIBaR and Brantley.

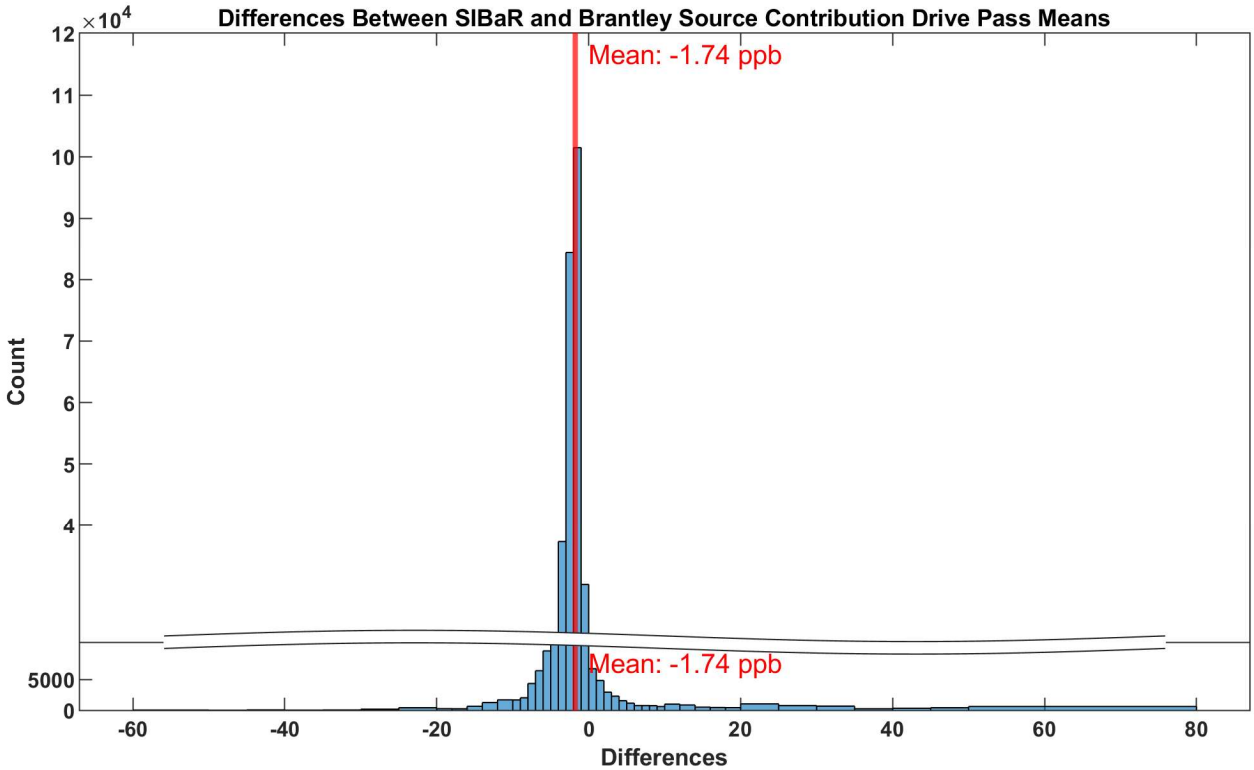

Figure S21. Histogram of differences between the SIBaR and Brantley drive pass means (SIBaR-Brantley) for NO_x. Bin spacing is 1 for differences in the interval [-10, 10), 2 for differences in intervals [-20,-10) and [10,20) and 5 for differences in intervals [-50,20) and [20,50). All differences less than -50 are put into a single bin, while differences greater than 50 are put into another.

Pseudocode for misclassification

Fitted Line Classifier

```
Set percentage threshold (50%)
Given time series:
Identify measurements where state transition takes place (evaluate differences in state vec and lagged state vec)
Store indices of transitions in vec indices
for index in indices:
    store timestamp at index
    average measurements at index-1,index,index+1
    store average
Fit line between timestamps, averages (averages ~ timestamps)
If num(background states above line)/num(states above line) *100 >50 OR num(source states below line)/num(states below line) > 50
     label "Misclassification"
Else
     label "Classified correctly"
Go to next time series
```

References

Census 2010 Tracts, [online] Available from: https://cohgis-mycity.opendata.arcgis.com/datasets/census-2010-tracts (Accessed 23 November 2020)

Finding pollution—and who it impacts most—in Houston, Environmental Defense Fund [online] Available from: https://www.edf.org/maps/airqualitymaps/houston/pollution-map/ (Accessed 23 November 2020).