Supporting Information

12 Pages, 7 Figures, 1 Table

Using lidar technology to assess urban air pollution and improve estimates of greenhouse gas emissions in Boston

Yanina D. Barrera[1], Thomas Nehrkorn[2], Jennifer Hegarty[2], Maryann Sargent[1], Joshua Benmergui[1], Elaine Gottlieb[1], Steven C. Wofsy[1], Phil DeCola[3,4], Lucy Hutyra[5], and Taylor Jones[1,3]*

[1]School of Engineering and Applied Sciences and Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, United States

[2] Atmospheric and Environmental Research Inc., Lexington, MA 02421, United States

3 Sigma Space Corp., Lanham, MD 20706, United States

[4] Department of Atmospheric and Oceanic Sciences, University of Maryland, College Park, MD 20742, United States

[5] Department of Earth and Environment, Boston University, Boston, MA 02215, United States

*Corresponding Author: E-mail: yaninabarrera@fas.harvard.edu, Telephone: +1617-4966361

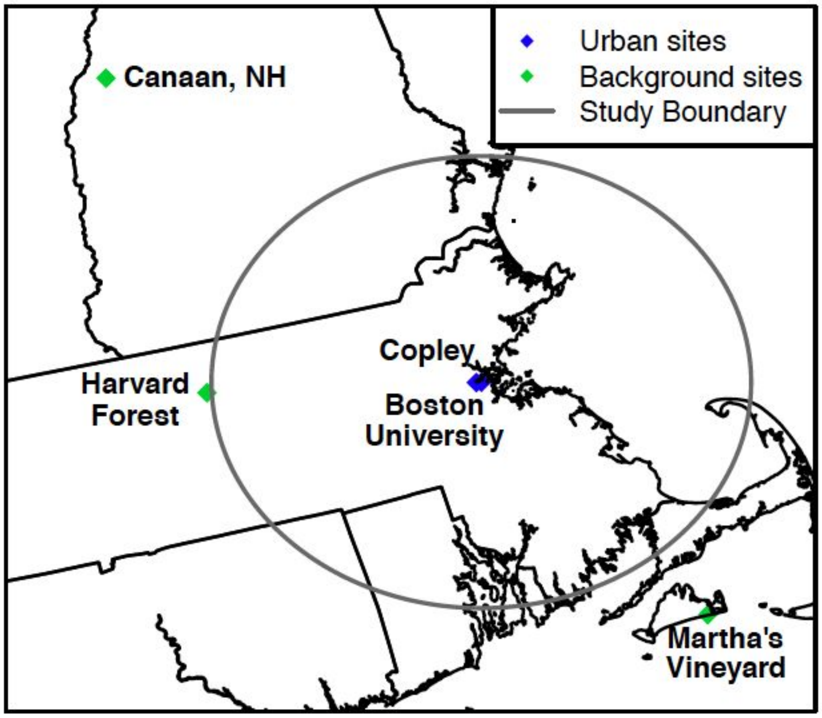


Figure S1. Background (green) and urban (blue) sites measuring carbon dioxide concentrations using Picarro instruments within and outside the Boston area. The study boundary is a 90-km radius around the Boston urban core where, estimates of carbon dioxides fluxes were conducted using two inverse models.

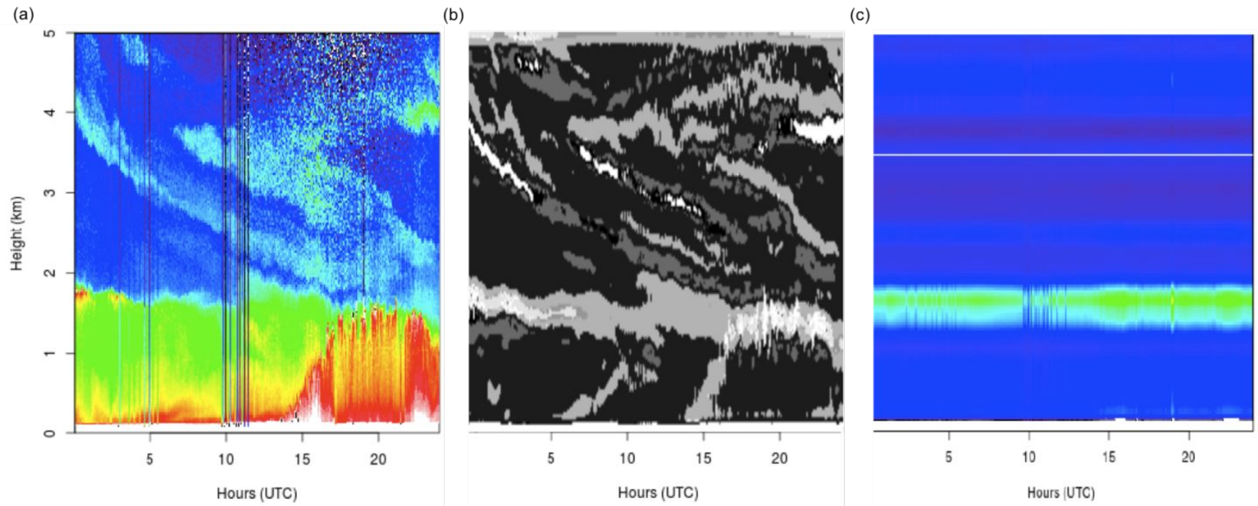
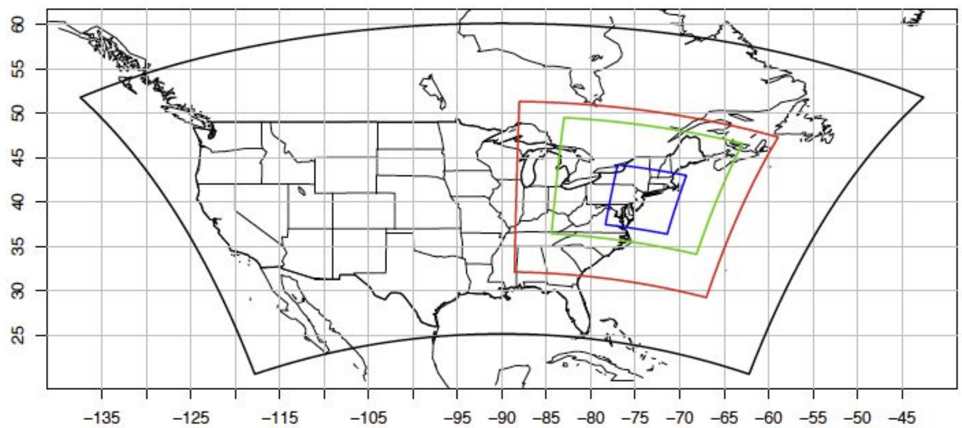


Figure S2. Lidar-retrieved NRB profile for August 21st, 2013 (a) was processed using first derivative Gaussian WCT followed by (b) edge detection and (c) singular value decomposition (SVD) to detect changes in NRB signals. With the human eye, one can see three layers in Figure S1a: the daytime PBL growth (red), residual layer (green), and the free atmosphere (blue). We see various layers (some thicker or longer than others), as well as various fluctuations in the NRB signal. By processing this image via edge detection, we remove some of the NRB signal fluctuations and enhance the edges of atmospheric structures. The grayscale compresses our image, making it easier to detect layers of interest at each time point. Our image recognition algorithm, which uses fuzzy logic, can then be used to identify atmospheric structures of interest such as the PBL and RL. SVD was used to retrieve the largest modes of our image (larger NRB signals) in our NRB profile, which are characteristics of the PBL and at times, the RL. On this day, we detected the RL at between 1.7km to 1.9km, roughly, using SVD.

(a)



(b)

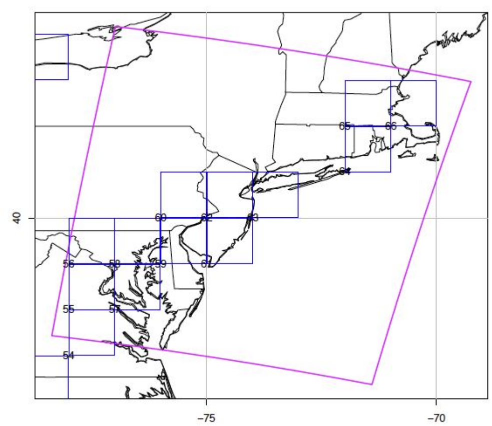


Figure S3. In the first phase of our NWP vs. miniMPL PBLH study, three WRF configurations were evaluated and the MYJ_v361 configuration agreed the most with miniMPL-retrieved PBLH. Four nested domains with the smallest in the horizontal direct at of 36km, 12km, 4km, and 1.33km (Figure S2a), for the WRF MYJ_v361 runs, were used with the primary focus in the Northeastern Corridor. Close-up of innermost WRF domain for Boston-DC WRF corridor runs are shown in Figure S2b. Overlaid in blue are the tiles of available urban parameter data (at 1km resolution) from WRF MYJ_v361.

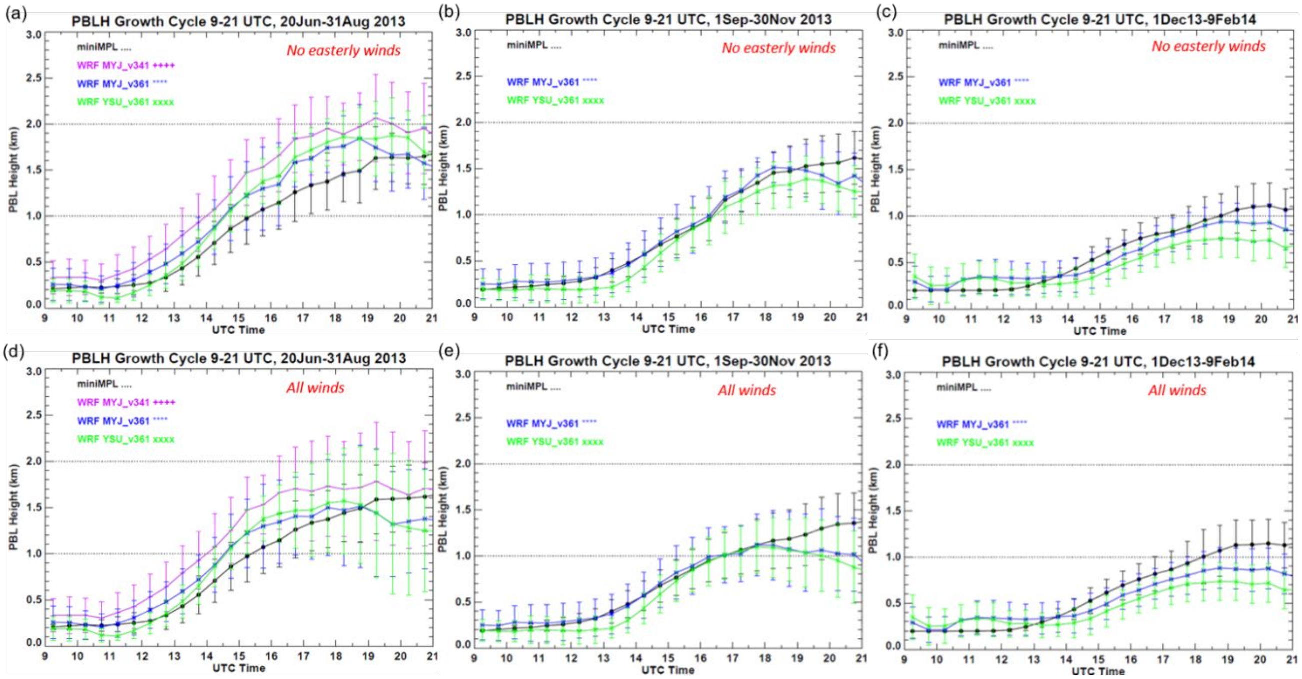


Figure S4. Three WRF versions MYJ_v361 vs. YSU_v361 vs. MYJ_v341, with miniMPL-retrieved PBL heights from 9-21 UTC for three periods: summer (June 20 – August 31, 2013), fall (September 1, 2013 – November 30, 2013), and winter (December 1 – February 28, 2014) with no easterly winds starting at 16 UTC (Figures S3 a, b, and c) and all winds (Figures S3 d, e, and f). The mean PBLH for each 30-minute is shown and error bars are the 1 standard deviation values. During the fall and winter periods, only MYJ_v361 and YSU_v361 were compared. PBL simulations agree the best for MYJ_v361 for all three seasons. During the late afternoon hours (18-21 UTC), the WRF model simulated the PBL collapsing early in comparison to the miniMPL-retrieved PBLH. Overall, agreements between all WRF PBL simulations and miniMPL-retrieved PBL heights improve when easterly afternoon winds (16 UTC and onward) data points are omitted from the analysis.

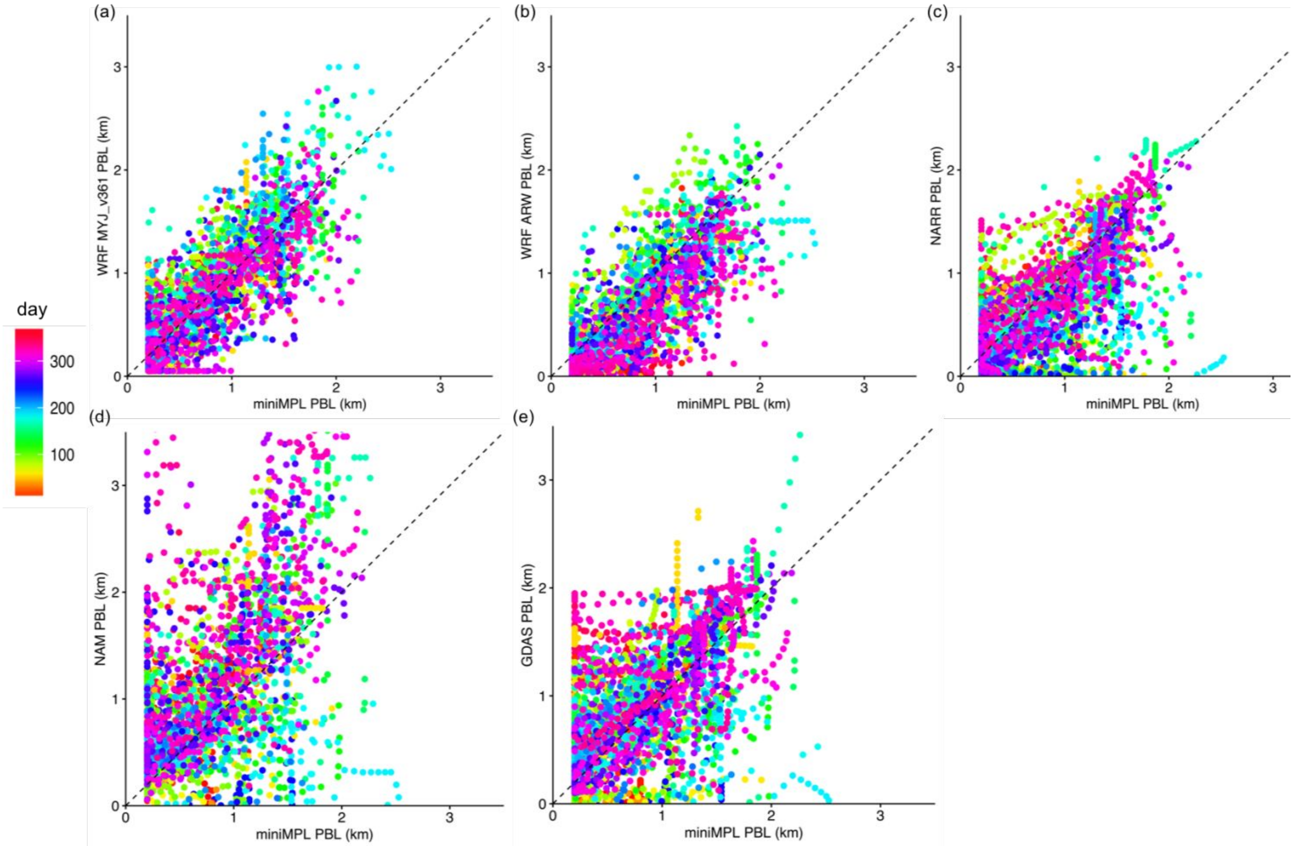


Figure S5. Scatter plots of NWP PBL simulations versus miniMPL-retrieved PBL heights from 9-21 UTC. Results are shown for Boston, MA, for each day of our CMS study period (September 1, 2013, to November 30, 2014), for the following NWP models: (a) MYJ_v361, (b) WRF ARW, (c) NARR, (d) NAM, and (e) GDAS. Data is color coded by day of period, ending from fall to fall (in red orange). The two forecast models (WRF MYJ_v361 at 1.33km and WRF ARW at 27km) agreed the most along the 1:1 line. The operational product NAM strongly deviates from the 1:1 line, when comparing PBL heights to miniMPL data.

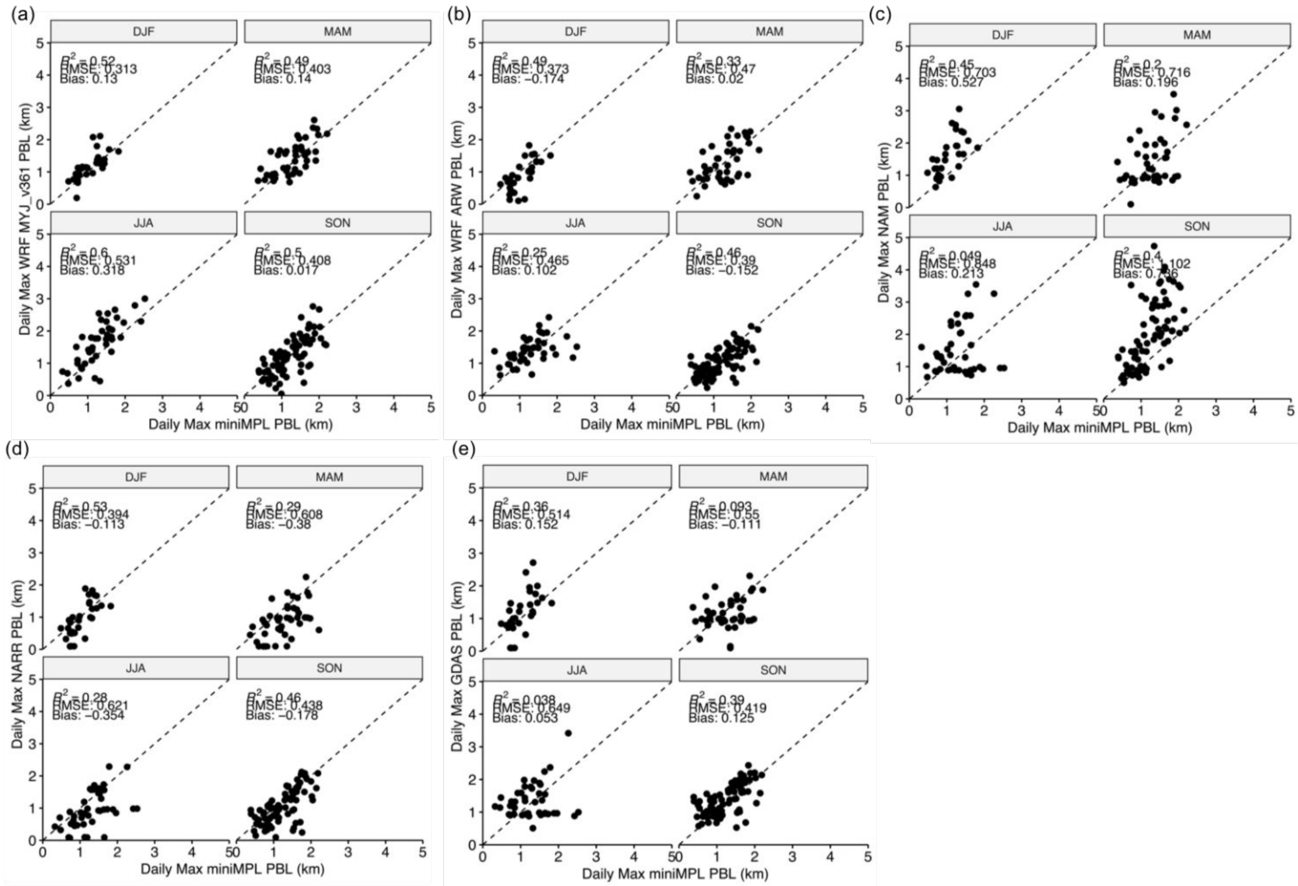


Figure S6. Seasonal scatter plots and correlation statistics of daily maximum PBL heights from NWP models and miniMPL data in Boston, MA. Results are shown for each season of our CMS study period (September 1, 2013, through November 30, 2104). The NWP models are as follows: (a) WRF MYJ_v361, (b) WRF ARW, (c) NAM, (d) NARR, and (e) GDAS. Overall, the forecast models (WRF MYJ_v361 and WRF ARW) strongly agreed with the miniMPL daily maximum PBLH, followed by the reanalysis product (NARR) and the operational product (GDAS). The operational product from NAM disagreed with the miniMPL-retrieved daily maximum PBLH for most seasons, with the largest PBLH differences in the summer and fall seasons (Figure S5c). For each NWP model, the strongest correlation with miniMPL-retrieved daily maximum PBLH varies by season. WRF MYJ_v361 agreed the strongest during the summer ($R^2 = 0.77$), WRF ARW ($R^2 = 0.70$) during the winter, NAM during the winter ($R^2 = 0.67$), NARR during the winter ($R^2 = 0.73$), and GDAS during the summer ($R^2 = 0.62$). Overall, the largest bias and mean error (RMSE) in daily maximum PBLH from NWP models occurred during the summer and fall months. The operational product from NAM biased the PBLH the most during our study CMS study period, 0.527km in the winter, 0.213km in the summer, 0.193km in the spring, and the largest bias of 0.736km in the fall. The operational product from GDAS and forecast model WRF MYJ_v361, however, biased the daily maximum PBLH the least for all seasons.

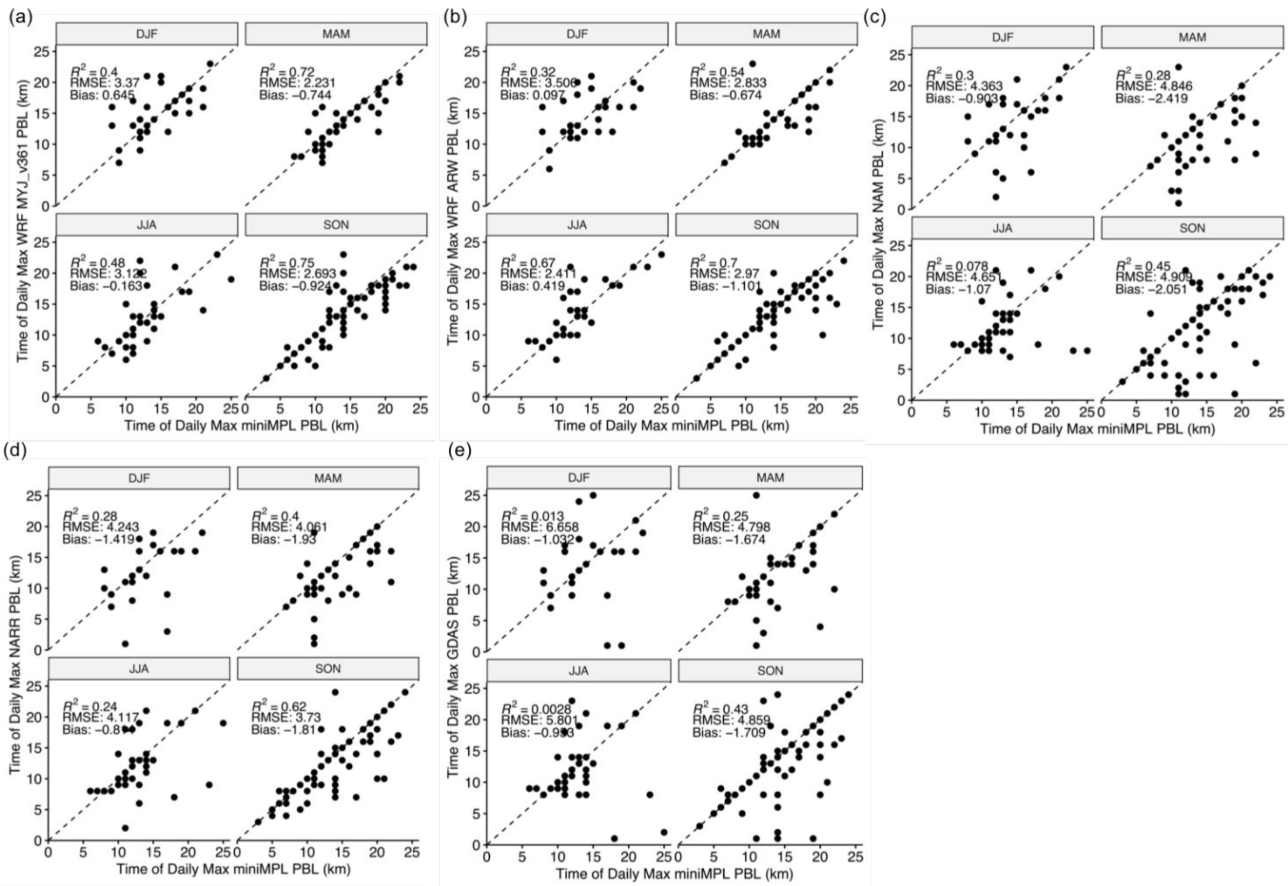


Figure S7. NWP vs. miniMPL time (UTC) of daily maximum PBL height quantized to the nearest hour between the hours of 9-21 UTC for each season of our CMS study period (September 1, 2013, through November 30, 2104), for the following models: (a) WRF MYJ_v361, (b) WRF ARW, (c) NAM, (d) NARR, and (e) GDAS. The time of the daily maximum PBLH retrieved from the miniMPL data agreed the strongest with the forecast models (WRF MYJ_v361 and WRF ARW) for all seasons, with highest agreement during summer (R^2=0.87, R^2=0.84) and least during winter (R^2=0.63, R^2=0.56). Overall, all NWP models agreed the most with the time where the PBL was fully developed, during the summertime. The largest mean bias for the time of daily maximum PBLH occurred during the fall and spring months: -0.924 UTC for WRF MYJ_v361 in the fall, -1.10 UTC for WRF ARW in the fall, -2.42 UTC for NAM in the spring, -1.93 UTC for NARR in the spring, and -1.709 UTC for GDAS in the fall. This result highlights that NWP model simulations of a fully developed PBL in Boston occur too soon in comparison to the miniMPL data. The operational (NAM and GDAS) and reanalysis (NARR) products showed the largest mean errors in time of daily maximum PBLH.

Model	Season	Mean Error	RMSE
GDAS	DJF	14.7%	0.513 km
GDAS	MAM	4.79%	0.550 km
GDAS	JJA	21.3%	0.649 km
GDAS	SON	21.3%	0.419 km
NAM	DJF	52.8%	0.703 km
NAM	MAM	25.5%	0.715 km
NAM	JJA	32.0 %	0.848 km
NAM	SON	63.3 %	1.10 km
NARR	DJF	-14.2%	0.394 km
NARR	MAM	-28.2%	0.608 km
NARR	JJA	-25.5%	0.621 km
NARR	SON	-11.8%	0.438 km
WRF ARW	DJF	-17.9%	0.373 km
WRF ARW	MAM	8.57%	0.470 km
WRF ARW	JJA	20.9%	0.465 km
WRF ARW	SON	-6.50%	0.390 km
WRF MYJ_v361	DJF	14.3%	0.313 km
WRF MYJ_v361	MAM	18.4%	0.403 km
WRF MYJ_v361	JJA	28.2%	0.531 km
WRF MYJ_v361	SON	4.75%	0.407 km

Table S1. The mean percent error in daily maximum PBLH (fully developed PBL) for each season, compared to miniMPL PBLH at the Boston site from September 2013 through November 2014. Data was filtered by WRF meteorological data and dates kept in miniMPL NRB profile analysis. Data with easterly winds (sea breezes) starting at 16 UTC were omitted from the analysis.