Supplement of Atmos. Chem. Phys., 24, 6197–6218, 2024
https://doi.org/10.5194/acp-24-6197-2024-supplement

Atmospheric Chemistry and Physics Open Access

Supplement of

Technical note: Challenges in detecting free tropospheric ozone trends in a sparsely sampled environment

Kai-Lan Chang et al.

Correspondence to: Kai-Lan Chang (kai-lan.chang@noaa.gov) and Owen R. Cooper (owen.r.cooper@noaa.gov)

List of Tables

List of Figures

Supplementary analysis for Section 3.2 is provided as follows:

- Since our meteorological adjustments are made by colocating ozone and dewpoint observations (when selecting subsamples, meteorological variables used in different sampling schemes are coupled with ozone data, i.e., not only ozone but also dewpoint is varying in the sampling analysis), we conduct a further sensitivity test by fitting the same ozone data, but using different sources of dewpoint measurements, including: (1) coupled daily data, (2) monthly averages of all daily nighttime data, and (3) monthly averages of all 24-h hourly data. We randomly selected 4 days-per-month to carry out this test by 1000 iterations. The result shows much lower residual RMSD and MAD are obtained from the coupled data, followed by monthly nighttime averages, and 24-h averages have the greatest errors (see Figure S7, note that ENSO index is only available at monthly scale and thus has no effect on this test). Therefore, it provides strong evidence that a better correlation and predictive performance can be achieved by colocating ozone and meteorological variables at a finer scale.

- In Figure 2 we demonstrate that the sampling bias from weekly samples has a smaller impact on temperature trends at MLO. We show the extended result by different time periods and by each day of the week in Figure S8 (trends for meteorological variables are based on basic model (M1)). As expected, although some variability is shown in trends from weekly subsamples, the result is more consistent between different days of the week.

- Although our focus is the effect of sampling bias on ozone trends, it is also desirable to carry out the similar attribution analysis to the pure sampling deviations (defined as the differences between ozone daily value and its monthly mean). By removing the ozone variability and accounting for meteorology, we aim to identify any remaining patterns (see Table S2 for the regression output). The result shows that some seasonal differences are present and a higher sampling variability is more likely to occur from July to November. As expected, neither a clear difference between days of the week nor a trend was found in the sampling deviations. Overall, a moderate low R^2 of 0.39 was found, indicating that a large portion of the sampling deviations might merely be random noise.

Supplementary analysis for Section 3.3 is provided as follows:

In this section we aim to quantify the marginal improvement of sampling bias by increasing the number of samples per month (Strategy B in Table 2). A complete random sampling is generally an infeasible plan in a monitoring program, but compared with the regular sampling Strategy A, the randomness in Strategy B enables us to avoid any potential systematic bias and investigate the undercoverage bias under different sampling rates. As discussed in Section 2.1, since the MLO data set has limited missing values, it is reasonable to assume the resulting data characteristics and statistics are representative of the underlying population, i.e., not only monthly means, but also the trend estimate and its uncertainty derived from full sampling, should also be representative. Based on this rationale, it is possible to quantify the respective effect of sampling bias on (1) monthly mean, (2) trend estimate, and (3) trend uncertainty. The negative effect of sampling bias on trend estimate and uncertainty can be considered to be a deterioration of trend accuracy and precision, respectively. Since our focus is the marginal improvement, the sampling frequency is the only control variable in this analysis (different time periods should play a minor role), so we only show the result based on the mean trends over 1990-2021 (it is sufficiently long such that the discrepancy between the mean and median trends is not critical).

Figure S11 shows the marginal decrement of bias exceedance rate and percentage bias as the monthly sampling rate increases (step curves), the colocated meteorological adjustments are made for all the trends and associated uncertainty estimates from full and reduced sampling. This figure demonstrates that even if the weekly sampling frequency cannot be increased, any additional monthly samples can still reduce the sampling bias and improve the trend accuracy and precision, albeit the marginal improvement generally levels off at a certain point (as it is closer to the truth). Note that the step curves in Figure S11 are approximated results based on a random sampling over 10,000 iterations. Therefore, we use logistic regression to smooth out some heterogeneity and explicitly quantify the marginal improvement for each additional

monthly sample. Let p_x be the bias exceedance rate or percentage bias derived from x profiles-per-month, then logistic regression can be expressed as $\log(p_x/(1-p_x)) = a + bx$, where a and b are coefficients to be fitted. Since the coefficient b in logistic regression cannot be directly interpreted as a slope or marginal effect, an average derivative is used to represent the marginal effect (by calculating a derivative of the curve for each x and then taking the average, Kleiber and Zeileis (2008)). An alternative approach is to use odds ratio (e^b), but odds ratio is often misinterpreted as probability, and it is not a standard measure in the atmospheric science literature, so odd ratio is not adopted here. In summary, the marginal improvement is roughly consistent between monthly mean, trend estimate and uncertainty, each additional monthly sample corresponds to an average reduction of bias exceedance rate by 3.3% (SD=1.6%) and percentage bias by 0.6% (SD=0.2%).

Note that an offset can be observed for the 5% bias exceedance rate in Figure S11(b), due to variability in the colocated dewpoint observations (standard regression model assumes that covariates are measured without error), which can cause some weak inconsistency in trend estimates (if we use the basic trend model, such an offset can be removed, see Figure S12). Nevertheless, since the other metrics are not affected, there is no need to particularly adjust the 5% bias exceedance rate in this analysis. If necessary, errors-in-variables models can be applied (Gleser, 1981; Li, 2002).

The exercise in Figure S11 can be used as a reference to determine the minimum number of samples necessary for accurate and precise trend detection. For example (based on logistic regression fits), at least 12 (and 23) samples per month are required to reduce MAPD trend bias to 10% (and 5%); and at least 15 (and 26) samples are required if RMSPD is a criterion. Based on Figure S11 and compared with previous studies, (1) we found that a roughly 10% (and 5%) sampling uncertainty in monthly mean (MAPD) is associated with a sampling frequency of 4 (and 12) samples per month, which is consistent with the IAGOS profiles at 700 hPa above Frankfurt, Germany (Saunois et al., 2012); (2) 25 samples per month are necessary for producing the mean value within ±2% MAPD bias, which is consistent with the regional study above western North America (Cooper et al., 2010); (3) Chang et al. (2020) found that 18 samples per month are required for the MAPD trend bias to be less than 5% above Europe; under the same frequency the MAPD trend bias is 6.8% at MLO, and it takes an additional 5 monthly samples to reach the 5% threshold, as the marginal improvement eventually becomes less effective.

Supplementary analysis for Section 3.4 is provided as follows:

To better understand the diverse result between monthly mean bias decreases and trend bias increases in Strategy C, Figure S13 demonstrates one of the most extreme cases in MAM: We first show the contrast of time series between full sampling and Friday sampling, and then show another scenario that a time series is produced from a mixture from full sampling for MAM and Friday sampling for other months. Under this particular circumstance, we show that the trends are consistent between full and once-per-week sampling, but the trend bias turns out to be inflated if we deliberately increase samples in MAM. On the other hand, even though sampling enhancement in MAM tends to cause a low bias in the overall trends (Figure 7), we show that a very high bias could also occur by simply selecting a different day of the week sampling (Figure S14). Therefore, the direction of trend bias is not necessarily connected to the seasonal variability.

Supplementary analysis for Section 3.5 is provided as follows:

Previous trend studies of free tropospheric ozone profiles and/or columns were typically conducted without considering other covariates (apart from the basic trend model (Tiao et al., 1986; Oltmans et al., 2006)) or by only incorporating large-scale circulations, such as ENSO and QBO (Logan, 1994; Oltmans et al., 2013; Chang et al., 2022). No previous trend studies (to the best of our knowledge) have thoroughly investigated the attributions of free tropospheric ozone profile data variability to meteorological variates. Therefore, while we aim to investigate the consistency between the Hilo ozonesondes and the MLO nighttime averages (subsampled to the colocated dates), it is also desirable to consider meteorological influences on ozonesonde trends. Note that relative humidity sensors on the older sondes were not as reliable as modern sensors (Fujiwara et al., 2003), so the records before July 1991 were excluded from this analysis.

Since the once-per-week sampling scheme at Hilo has too few profiles to perform the resampling analysis (as we did for the MLO record in Figures 4-6), we are not able to properly quantify the improvement of trend accuracy due to covariate adjustments, so we focus on the reductions of fitted residuals (an indication of the overall fitted quality) before and after incorporating covariates. The results are shown in Figure S18:

- An overall strong correlation can be found between individual Hilo ozonesondes (680 hPa) and the colocated MLO nighttime averages.
- Our previous findings show that meteorological adjustments on average reduce the fitted residuals by 27% and trend uncertainty by 35% at MLO. Consistent improvements can be found at the corresponding level (680 hPa) above Hilo, by 24% and 34%, respectively, further demonstrating that the free tropospheric ozone variability can be attributed to colocated meteorological influence (i.e., dewpoint variability in this analysis).
- Nevertheless, highly consistent trends are still not observed between the Hilo ozonesonde (680 hPa) and the colocated MLO record. The reason behind this warrants further detailed investigation, but the combined effect of measurement uncertainty and intra-daily variability is expected to play a major role.

Table S1: A list of the months that have not met the 50% data coverage criterion (MLO ozone record).

Year	Month	# daily values
1984	Apr	0
1987	Apr	0
1987	May	0
1987	Jun	0
1987	Jul	8
2000	Sep	0
2002	May	14
2003	Aug	14
2004	Dec	10
2005	Jan	0
2005	Feb	0
2005	Mar	8

Table S2: Numerical output from the multiple regression fit to the sampling deviations (differences between daily ozone value and its monthly mean), where each month and day of the week are treated as discrete factors (e.g., to investigate which months are more likely to have a stronger sampling variability).

	Estimate	Std. Error	SNR	P-value
Intercept [ppbv]	-1.99	1.13	-1.76	0.08
trend [ppbv/dec.]	<0.01	0.09	0.04	0.97
dewpoint [°C]	-0.53	0.03	-18.65	<0.01
relative humidity [%]	-0.09	0.01	-8.42	<0.01
temperature [°C]	-0.89	0.05	-17.62	<0.01
wind direction [degree]	-0.01	<0.01	-3.03	<0.01
wind speed [m/s]	-0.04	0.03	-1.37	0.17
factor(Feb) [ppbv]	-0.04	0.41	-0.10	0.92
factor(Mar) [ppbv]	1.22	0.40	3.05	<0.01
factor(Apr) [ppbv]	0.21	0.40	0.52	0.60
factor(May) [ppbv]	2.07	0.41	5.06	<0.01
factor(Jun) [ppbv]	2.70	0.42	6.36	<0.01
factor(Jul) [ppbv]	4.81	0.42	11.34	<0.01
factor(Aug) [ppbv]	5.90	0.44	13.55	<0.01
factor(Sep) [ppbv]	5.15	0.43	11.91	<0.01
factor(Oct) [ppbv]	4.33	0.41	10.50	<0.01
factor(Nov) [ppbv]	3.83	0.41	9.43	<0.01
factor(Dec) [ppbv]	1.32	0.40	3.29	<0.01
factor(Mon) [ppbv]	-0.11	0.30	-0.35	0.73
factor(Tue) [ppbv]	-0.49	0.30	-1.61	0.11
factor(Wed) [ppbv]	-0.34	0.30	-1.12	0.26
factor(Thu) [ppbv]	-0.10	0.30	-0.34	0.73
factor(Fri) [ppbv]	-0.33	0.30	-1.10	0.27
factor(Sat) [ppbv]	-0.10	0.30	-0.33	0.74

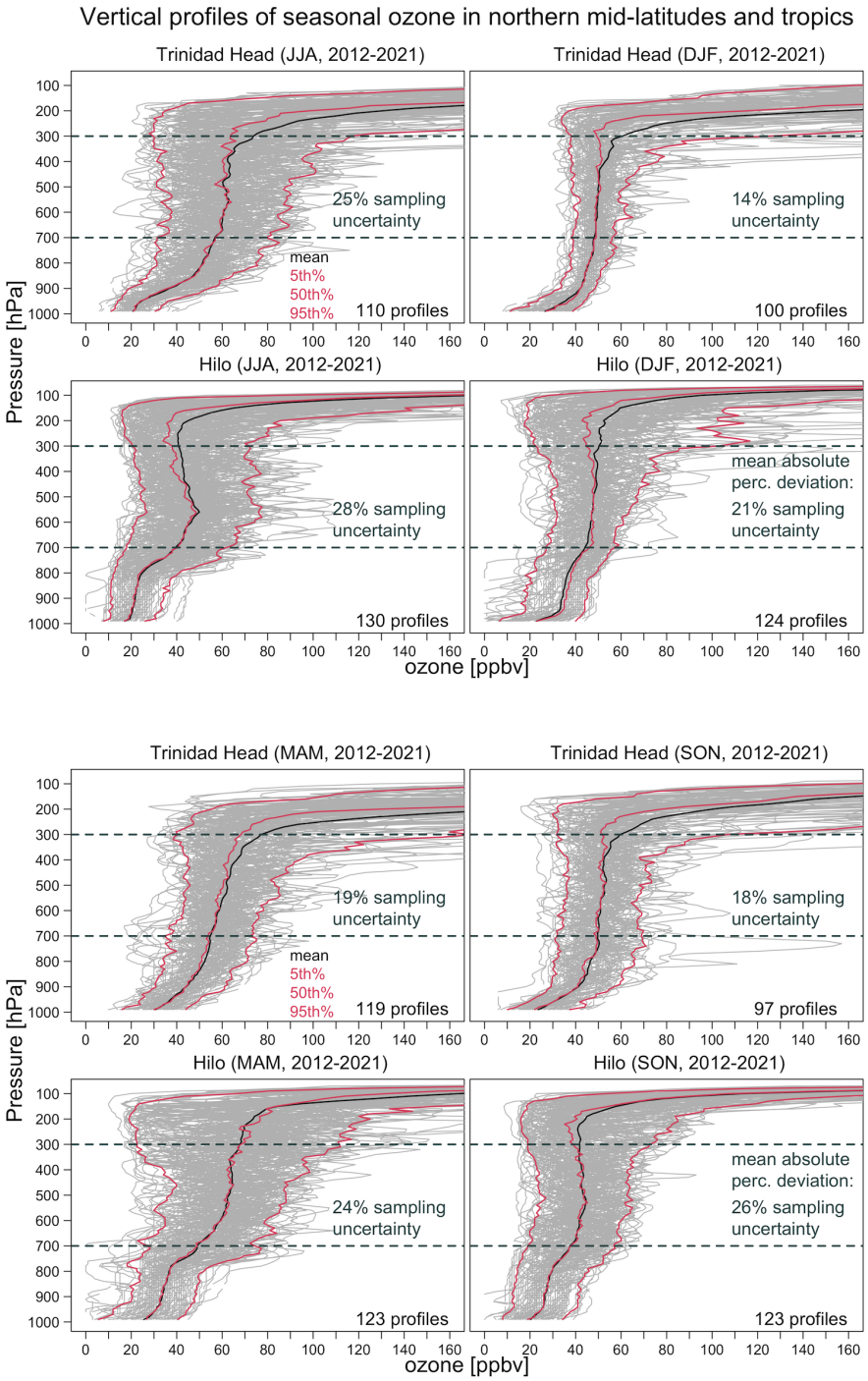

Figure S1: Vertical profiles of seasonal ozone in the northern mid-latitudes (Trinidad Head, California) and the tropics (Hilo, Hawaii) over 2012-2021: gray lines represent individual sondes, black lines represent the mean, and red lines represent the 5th, 50th and 95th percentiles. Sampling uncertainties are evaluated by mean absolute percentage deviation at 10 hPa resolution layers within 700-300 hPa.

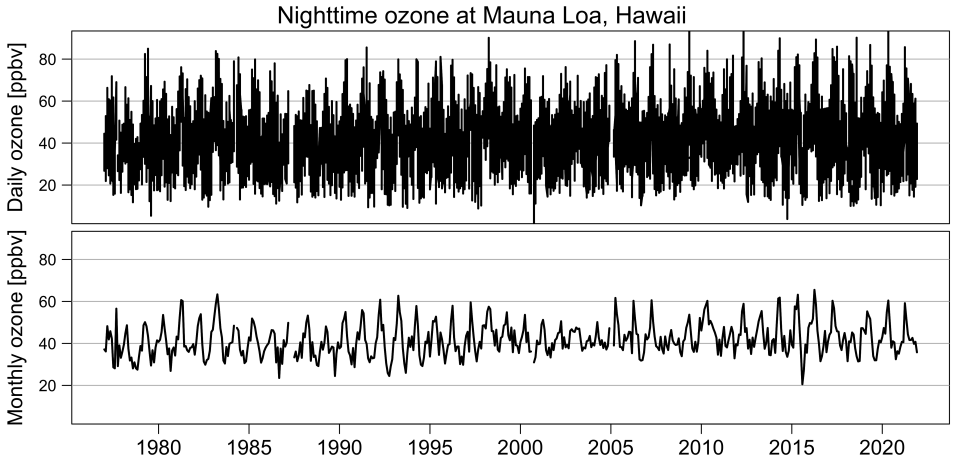

Figure S2: Daily and monthly nighttime ozone mean time series at Mauna Loa, Hawaii.

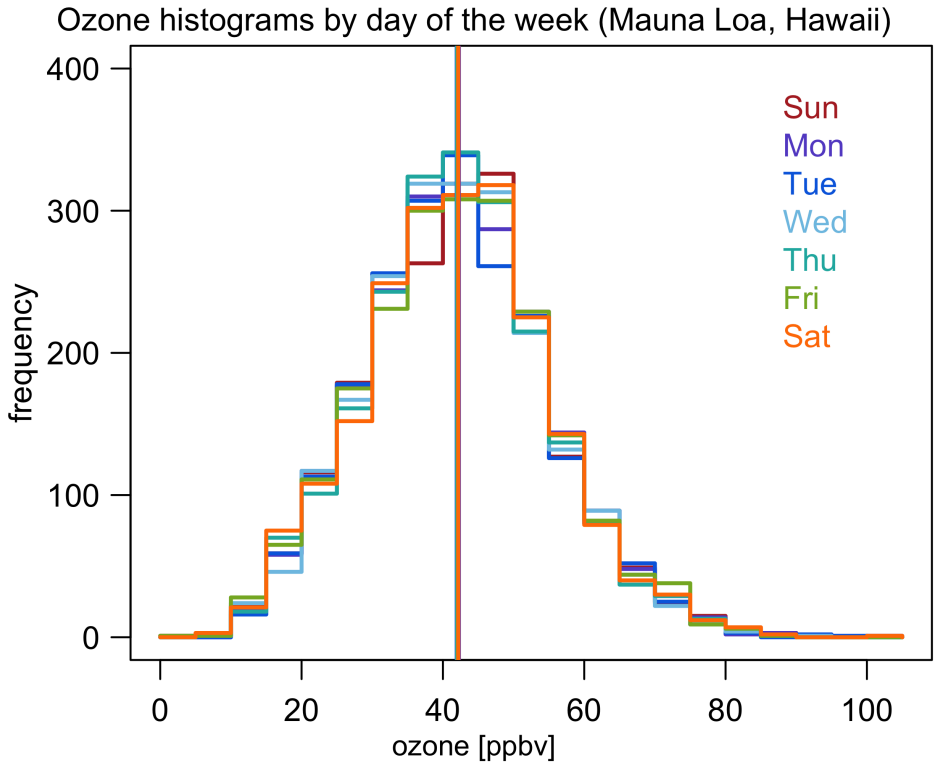

Figure S3: Daily nighttime ozone histograms by each day of the week at Mauna Loa, Hawaii. No distinguishable difference can be observed in the average of each histogram (as indicated by the vertical line).

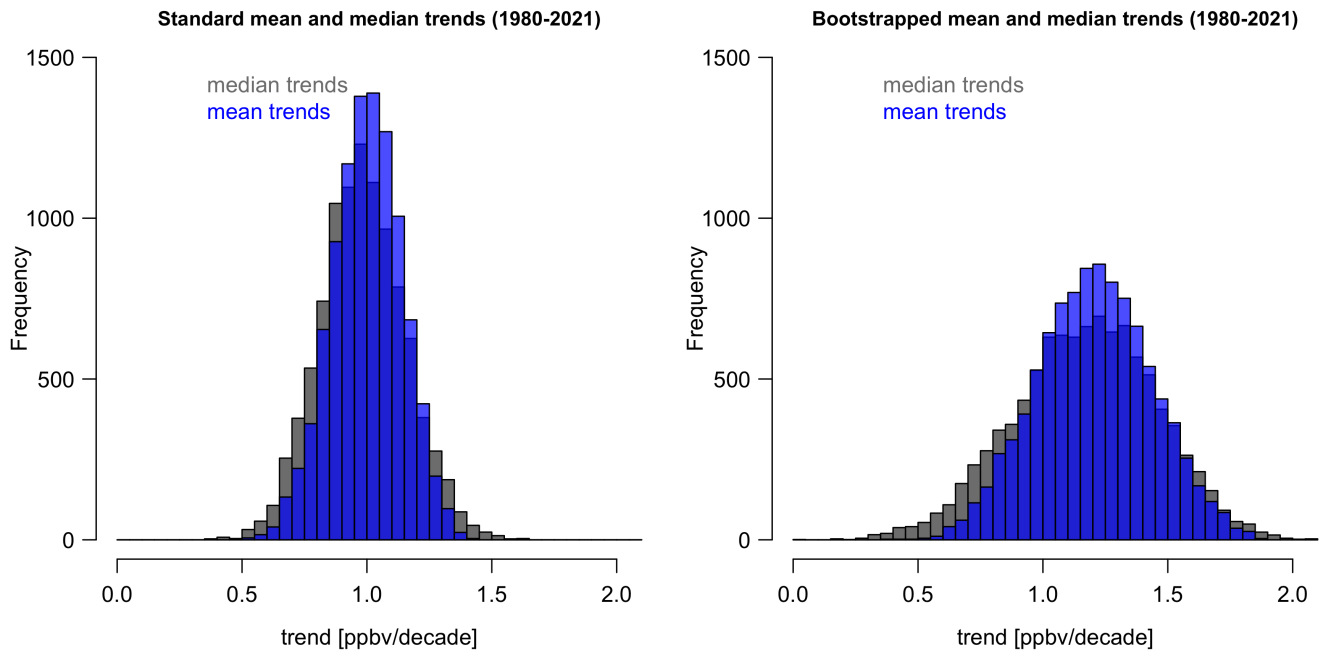

Figure S4: Resampling distributions of the median/mean trends based on standard LAD/LS fits (left) and moving block bootstrap (right). A total of 10,000 iterations is made and for each iteration, 4 samples per month are randomly selected and then the median/mean trends are fitted to the same subsamples. The result shows that the mean-based regression tends to have a narrower uncertainty than the median-based regression, and moving block bootstrap (accounted for autocorrelation) tends to have a greater uncertainty than the standard regression fits. Trends and associated uncertainty estimates are meteorologically adjusted (MLO nighttime ozone, 1980-2021).

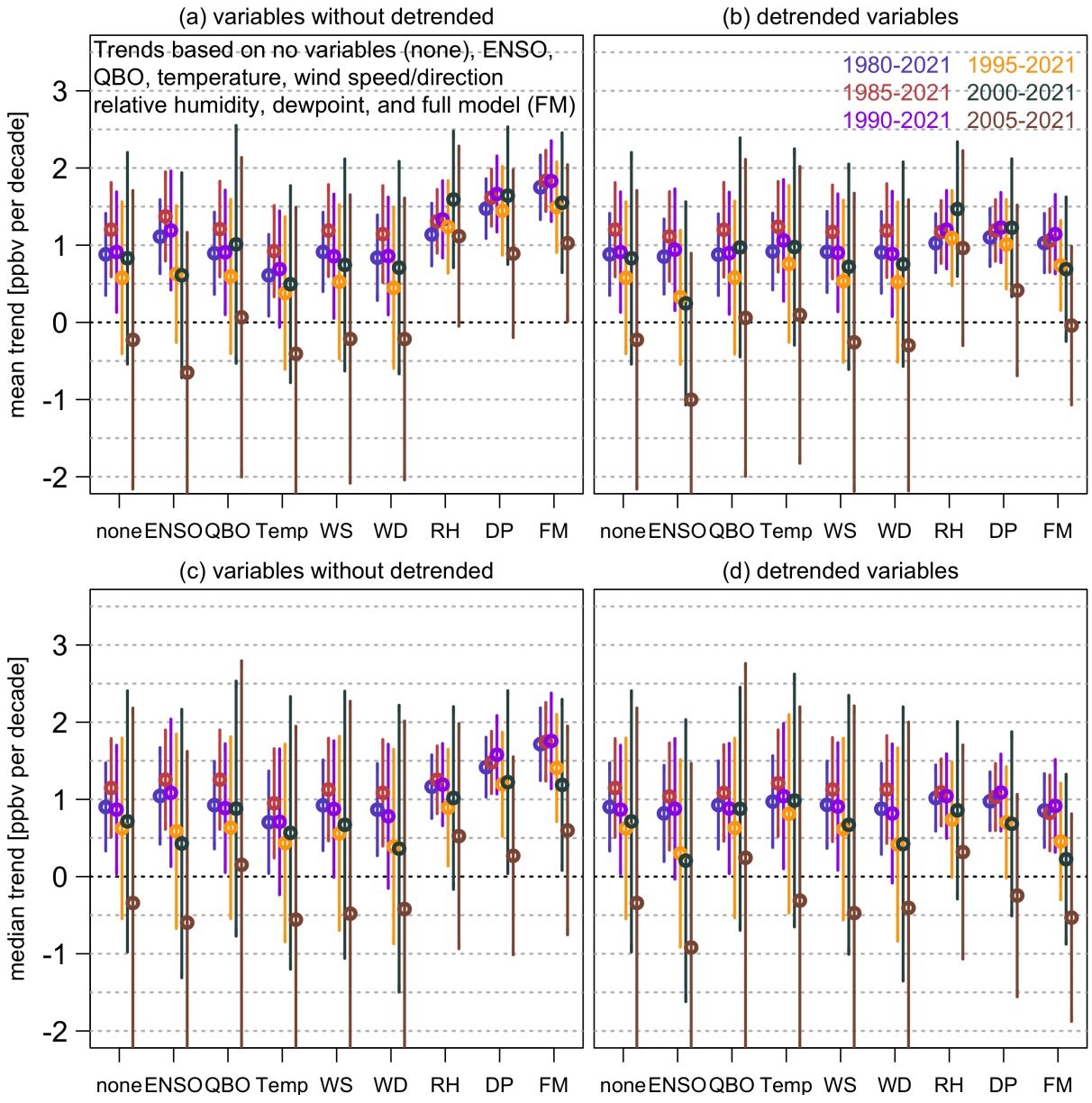

Figure S5: Comparison of the impact of climate indices and meteorological variables on MLO ozone trend estimates and 2-sigma intervals: Ozone trends are based on the mean (a & b) and median (c & d) estimators, and derived from variables without detrended (a & c) and detrended variables (b & d), respectively. In each panel the results are based on the basic model (left), each individual variable (middle), and full model (right), as well as different time periods. Variables include El Niño-Southern Oscillation (ENSO), quasi biennial oscillation (QBO), temperature, wind speed (WS), wind direction (WD), relative humidity (RS), and dewpoint (DP). Note that the peak correlation between ENSO and ozone is found where the ENSO index shifts forward by 5 months, so here a lagged ENSO correlation is considered, albeit no noticeable impact on trends is found when using the peak or zero-lag ENSO correlation (not shown).

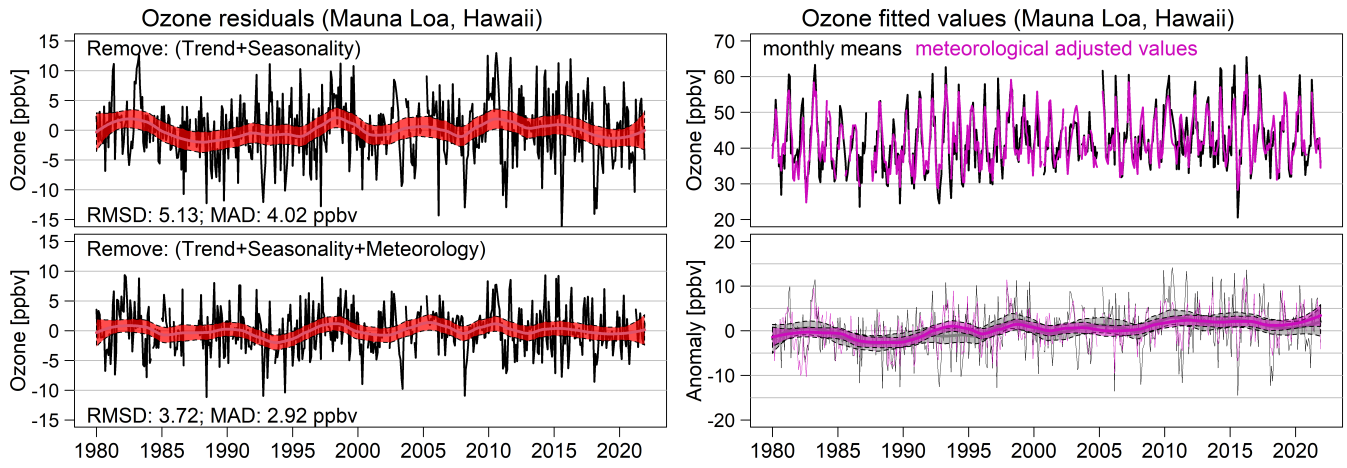

Figure S6: Residuals from least squares regression models without and with meteorological adjustments (left), and meteorological adjusted ozone values and anomalies (right). Shaded curves indicate the Lowess smoother [$\pm$ 2-σ].

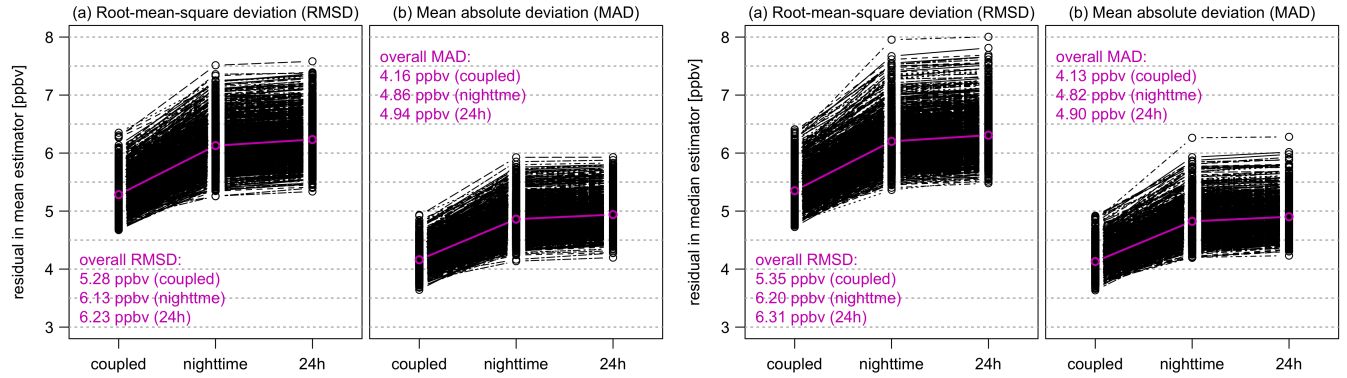

Figure S7: Residual RMSD and MAD from different sources of meteorological observations, based on the mean (left) and median (right) estimators: Trend models are fitted through meteorological variables selected from colocated sampling dates (coupled), from all nighttime averages (nighttime), and from all hourly averages (24h), respectively. Each black line represents an outcome from resampled data, and the purple line represents the average over 1000 iterations. The result shows a better predictive performance can be achieved by colocating dependent and independent variables at a finer scale.

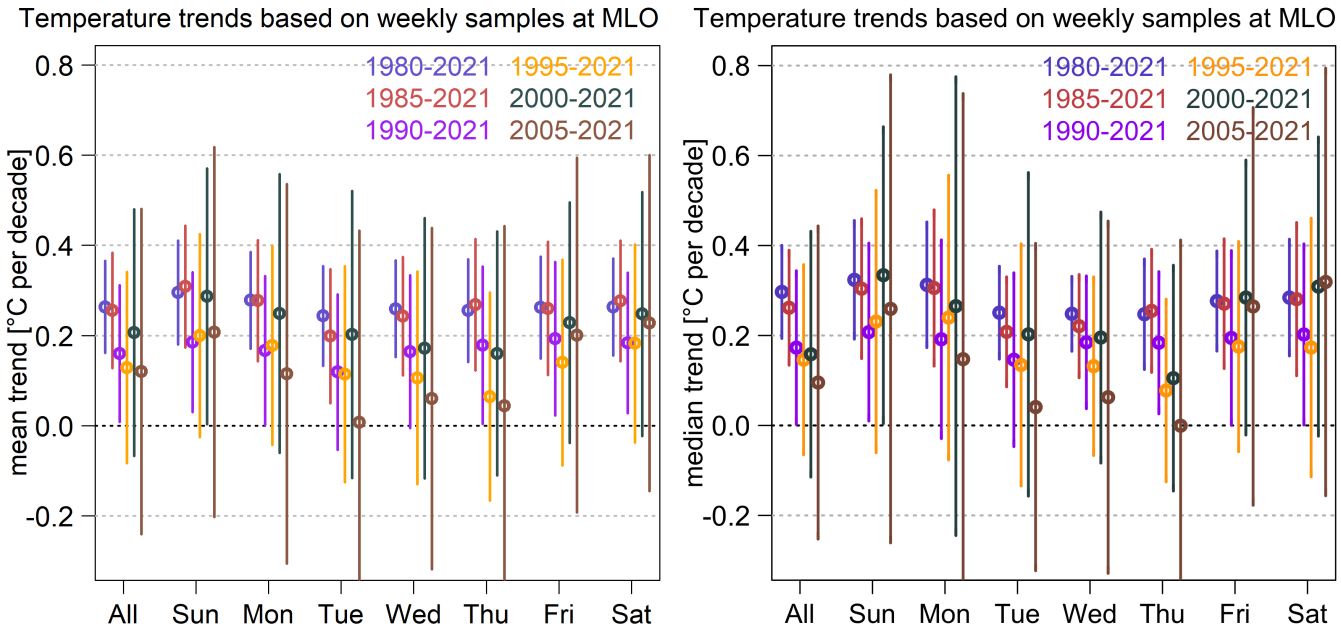

Figure S8: MLO nighttime temperature trends based on the mean (left) and median (right) estimators.

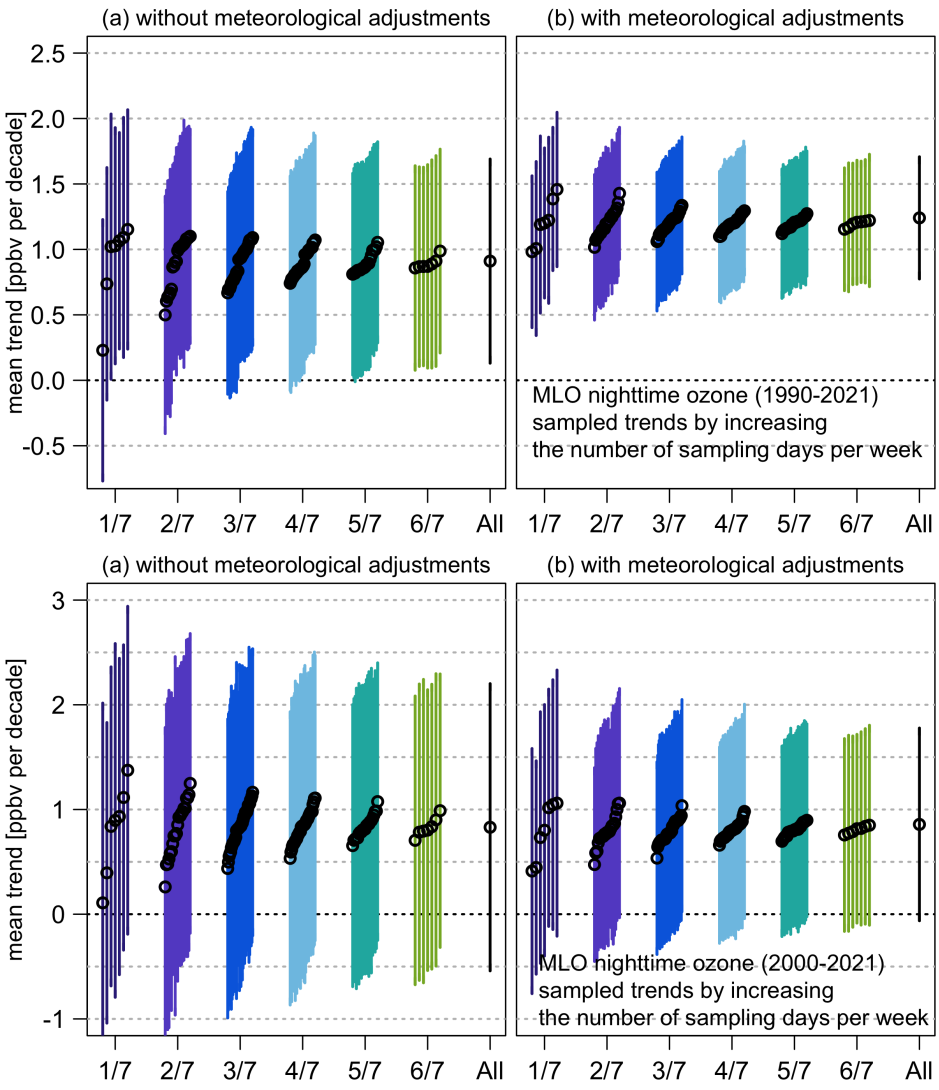

Figure S9: Impact of increasing weekly sampling frequency on trends over 1990-2021 (upper panel) and 2000-2021 (lower panel): The possible combinations are different for each sampling scheme (i.e., a total of 7 sets for 1 day/week and 6 days/week, 21 sets for 2 days/week and 5 days/week, and 35 sets for 3 days/week and 4 days/week). For each scheme, results are sorted from the lowest to the highest sampled trend values (MLO nighttime ozone).

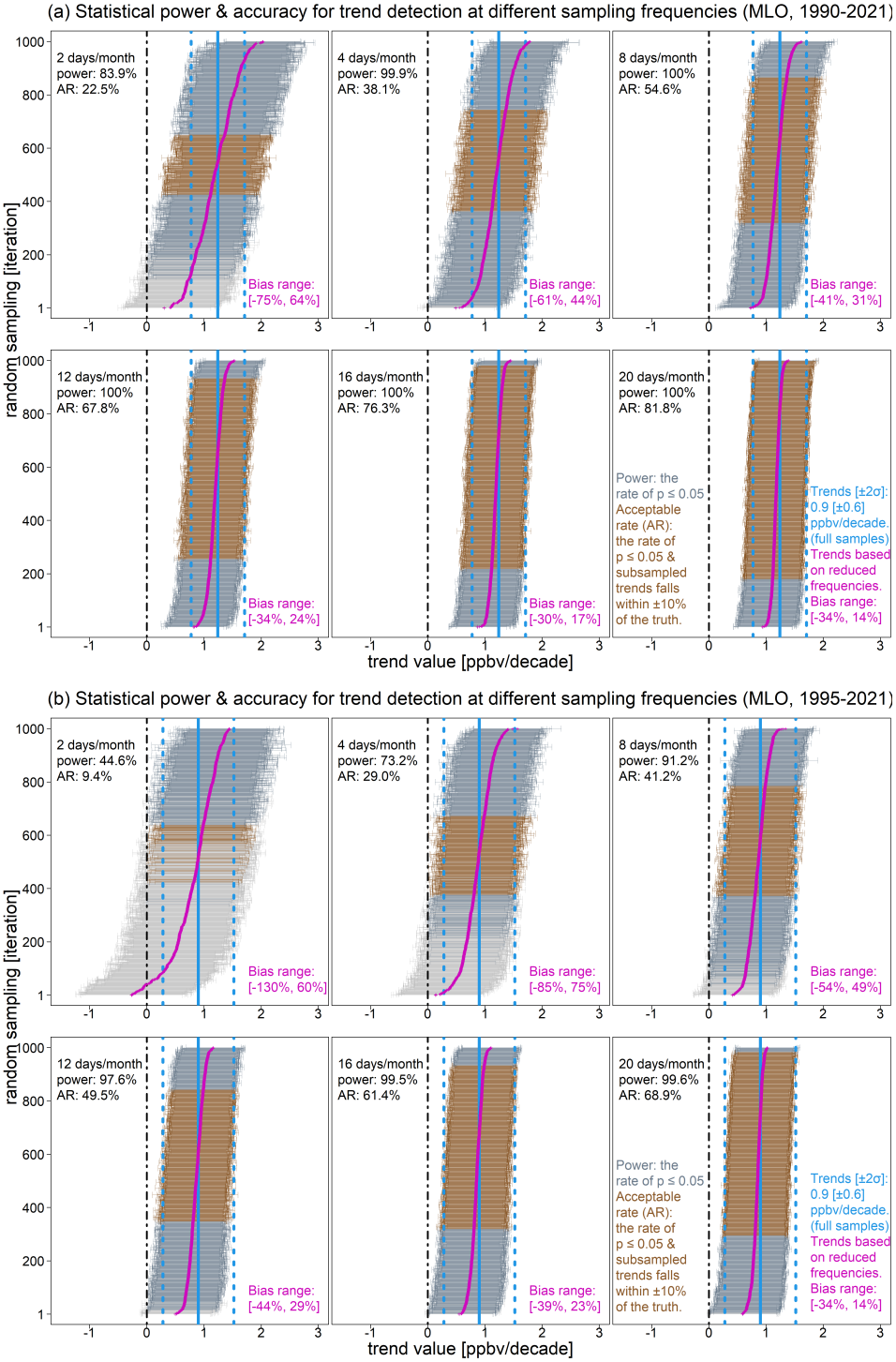

Figure S10: Same as Figure 6, but (a) for 1990-2021 with meteorological adjustments (to show the scenario when a stronger trend and SNR are present), and (b) for 1995-2021 with meteorological adjustments (to show the scenario that when a similar trend is present, a lower sigma can also yield a reasonable statistical power). Note that meteorology is accounted for in these panels, so the acceptable rate at a low sampling frequency is greater than Figure 6 (see Section 3.2 for detailed discussions of meteorological impact on trend accuracy and precision), however, the results are less satisfactory compared to scenarios with higher sampling rates.

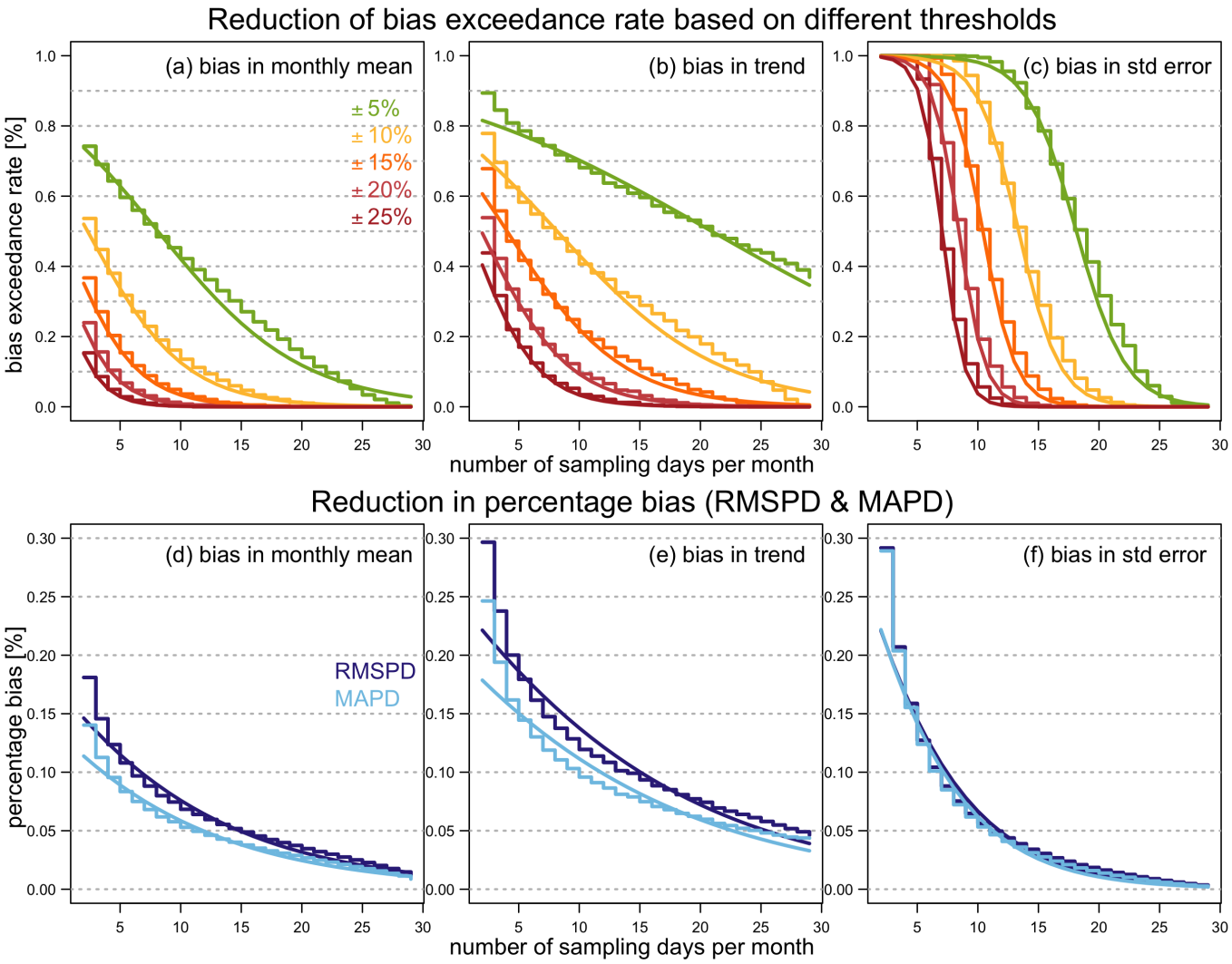

Figure S11: Marginal decrement of the bias exceedance rate (upper panel), and RMSPD and MAPD (lower panel) in monthly means, trend estimates, and trend uncertainties, according to different sampling frequencies per month (MLO nighttime ozone, 1990-2021): Step curves represent the results obtained from resampling method, and smooth curves represent the logistic regression model fit in order to quantify the marginal improvement. Trends and associated uncertainty estimates are meteorologically adjusted.

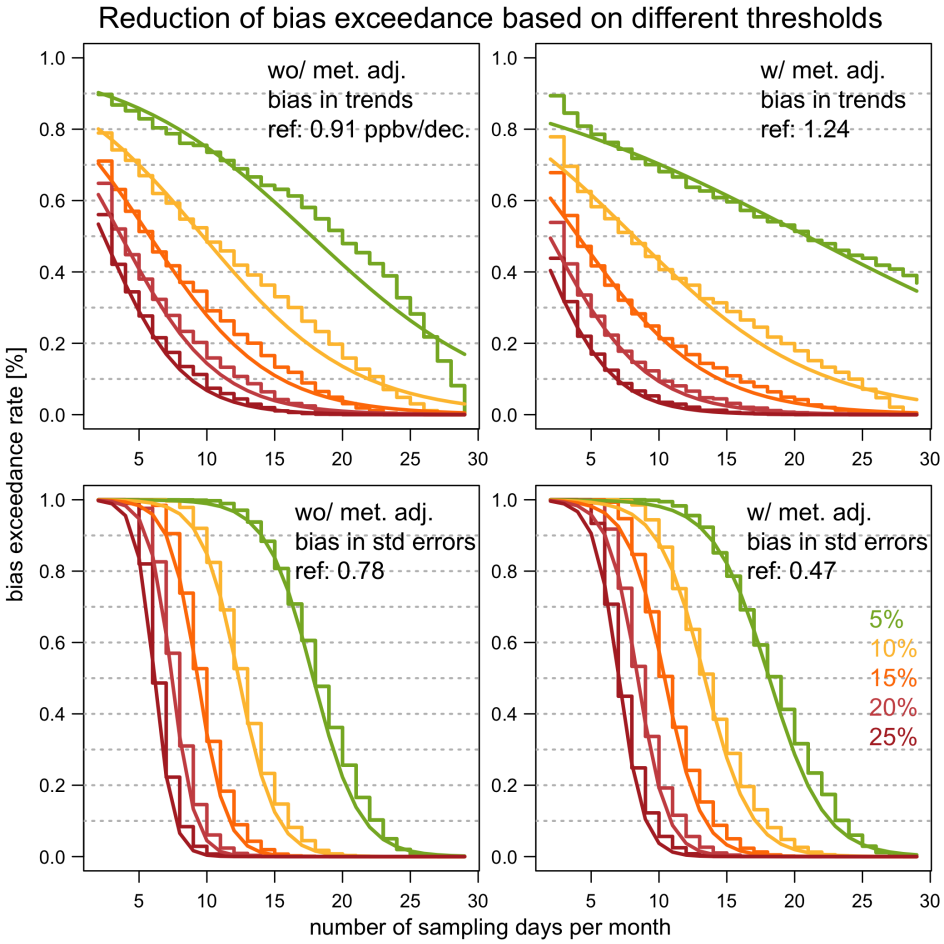

Figure S12: Same as Figure S11, but based on without (left) and with (right) meteorological adjustments. The bias exceedance rates in each panel are estimated based on different reference numbers. For example, the green curve in the upper left panel is determined by how often do $|s_{k,x} - 0.91|/0.91$ exceeds 0.05 in 10000 resampling (where $s_{k,x}$ is the bootstrapped trend value, $k = 1, ..., 10000$ iterations and $x = 2, ..., 29$ days/month). This result demonstrates that since there is an uncertainty in the colocated meteorological variables, the marginal decrement for the 5% bias exceedance rate is less efficient when the sampling frequency increases (but the other metrics are not affected).

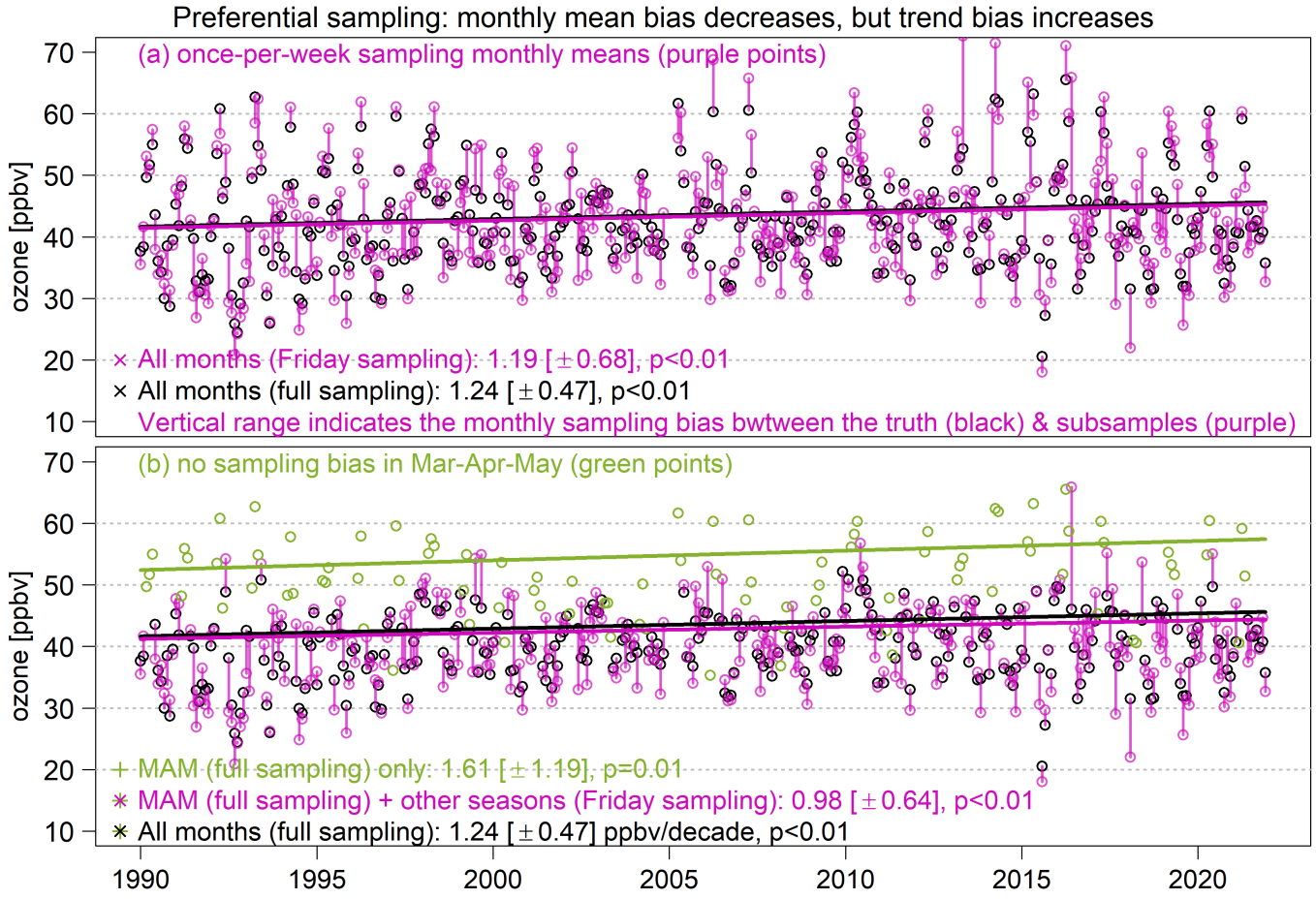

Figure S13: Demonstration of preferential sampling for ozone time series: The upper panel shows the magnitude of monthly sampling bias between full sampling (black) and once-per-week sampling on Friday (purple). The lower panel is based on the same scheme, but under the assumption of no sampling bias in monthly means over Mar-Apr-May (green). Trends and associated uncertainty estimates are meteorologically adjusted. This example indicates that if we deliberately enhance sampling during certain months, imbalanced sampling can result in a stronger bias in the overall trend.

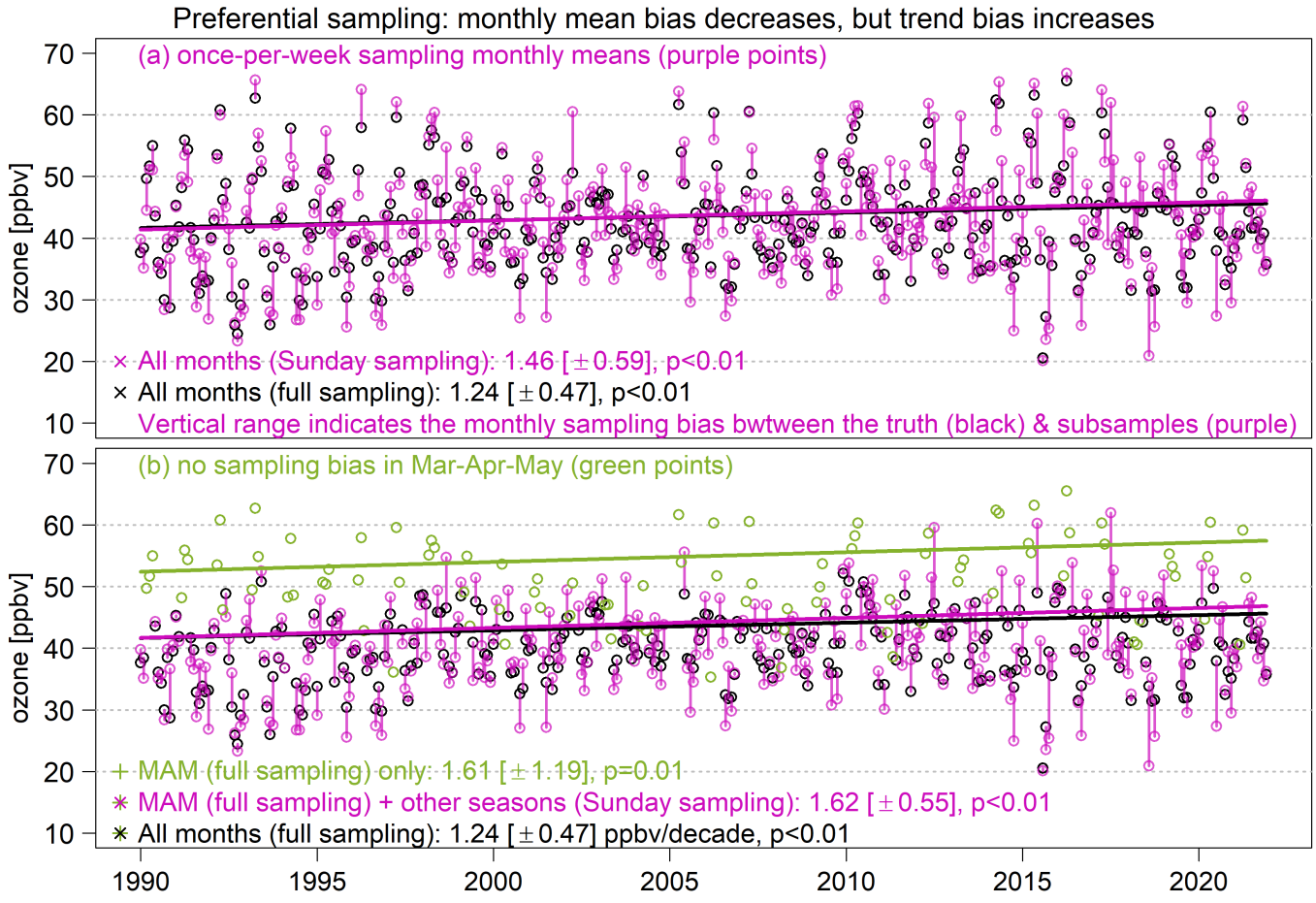

Figure S14: Same as Figure S13, but replacing Friday with Sunday.

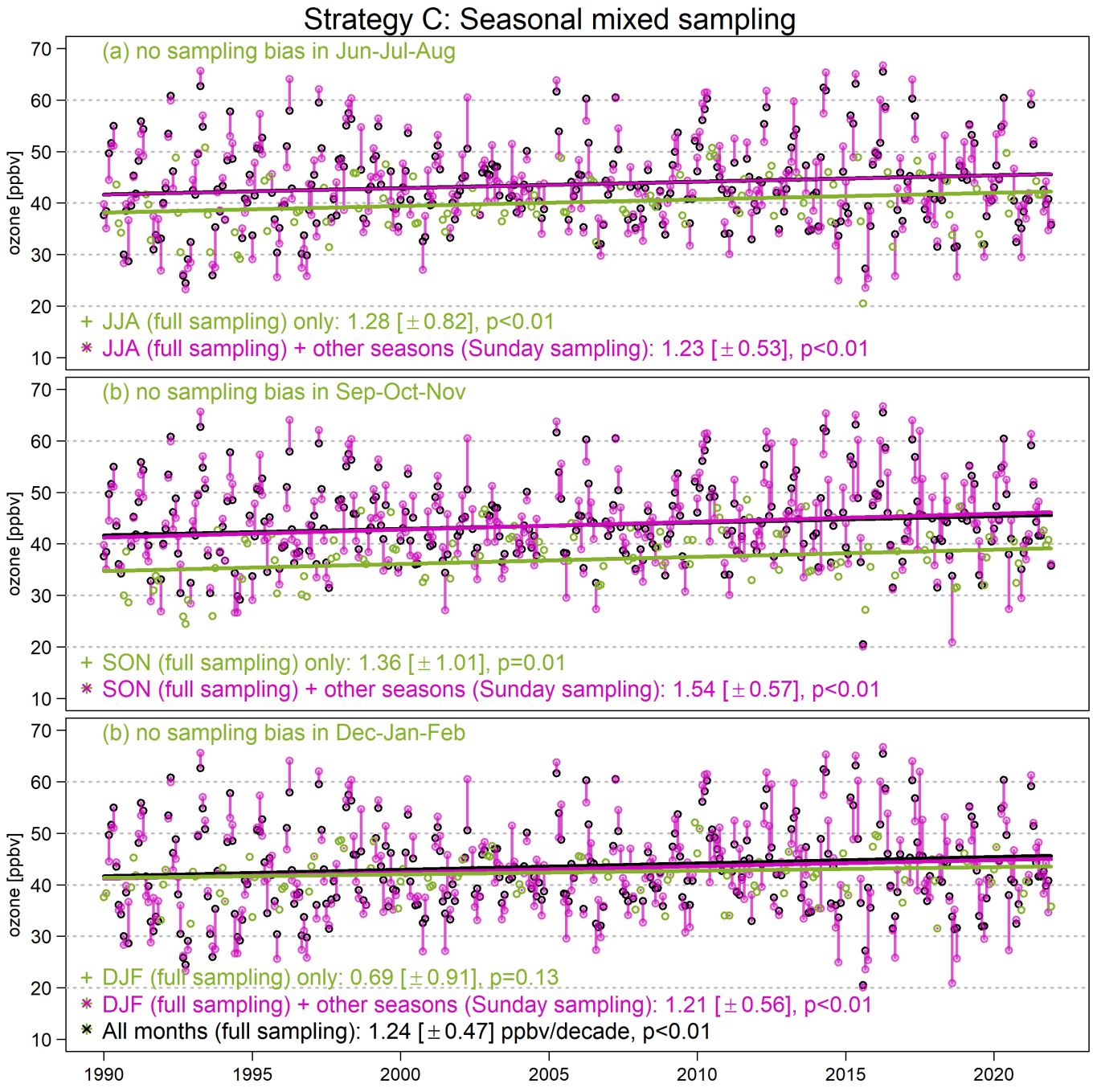

Figure S15: Same as Figure S14, but replacing MAM with other seasons.

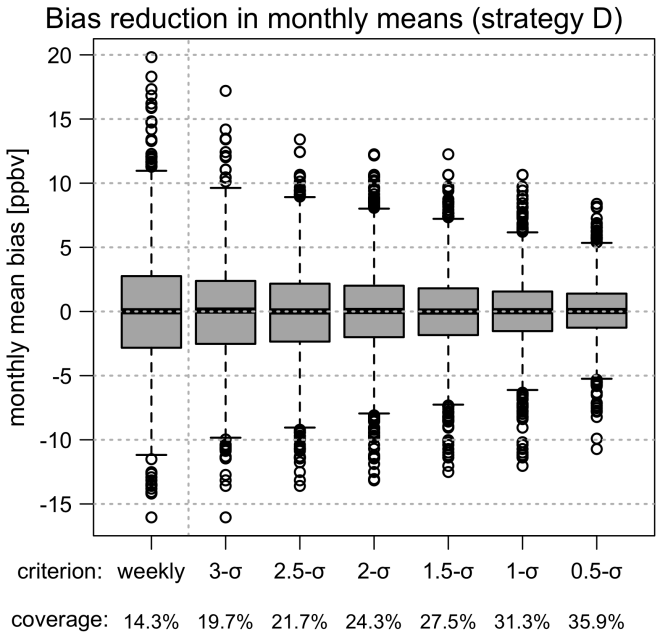

Figure S16: Monthly mean bias (in units of ppbv) for Strategy D: Result is based on at most 3 samples per week scenario and different tolerance ranges (MLO nighttime ozone, 1990-2021).

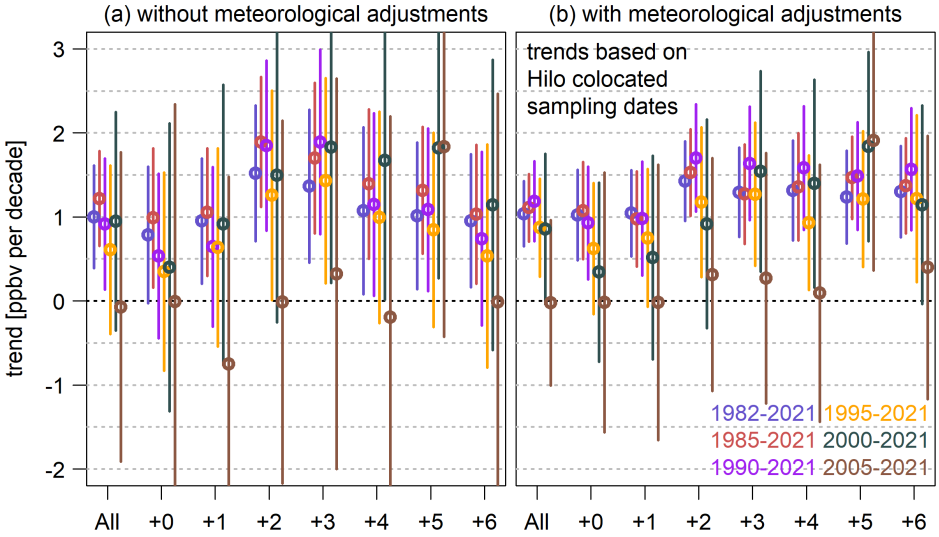

Figure S17: MLO ozone trends based on Hilo ozonesonde sampling dates (labeled as +0), where +1 indicates the trends based on data taken from one day after Hilo ozonesonde sampling dates, and so on.

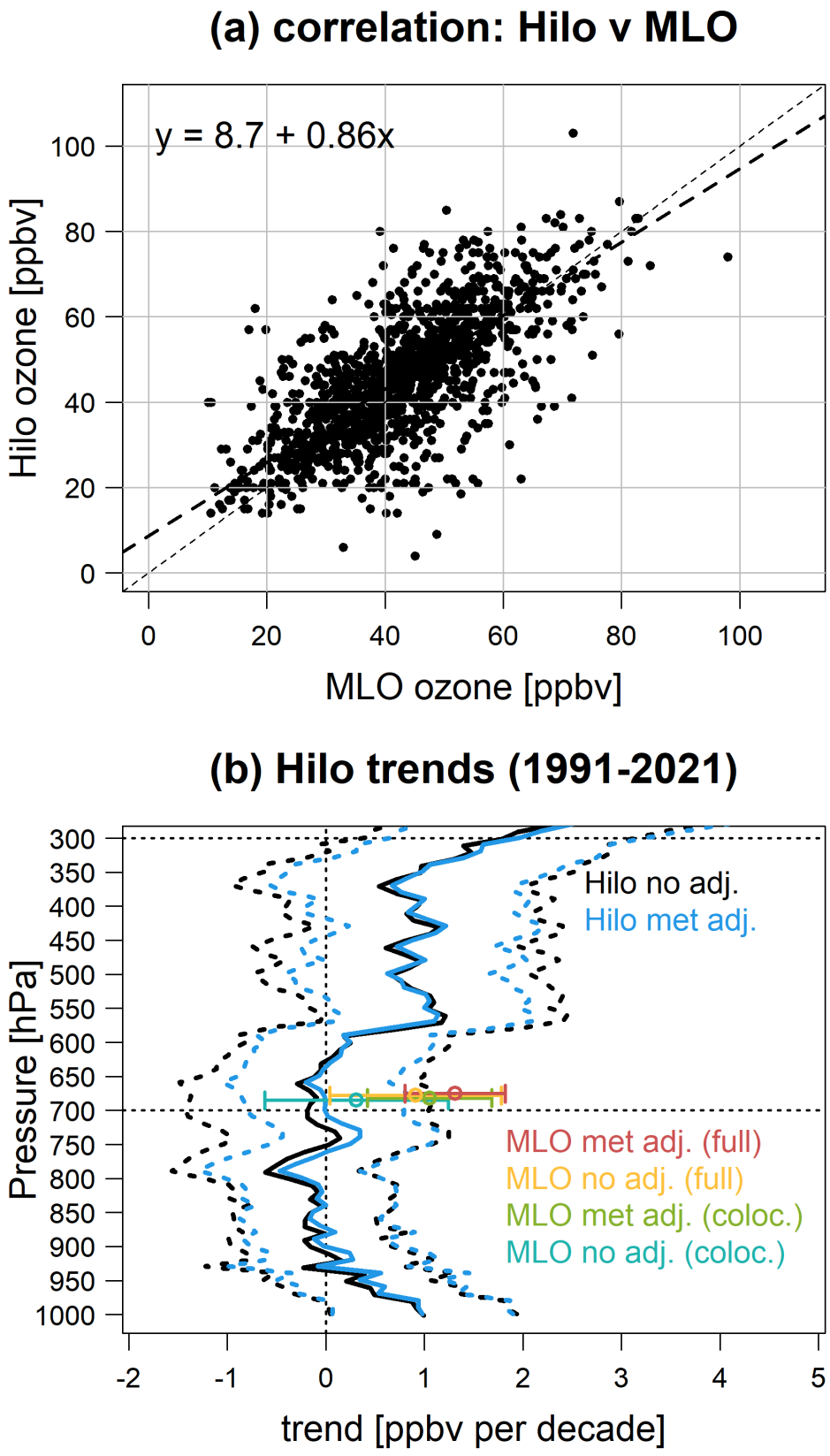

Figure S18: (a) Measurement correlation between individual Hilo ozonesondes (680 hPa) and their colocated MLO nighttime average over 1991-2021; and (b) the Hilo trend profiles, along with the MLO trends (full or colocated record), with or without the meteorological adjustments.

References

Chang, K.-L., Cooper, O. R., Gaudel, A., Allaart, M., Ancellet, G., Clark, H., Godin-Beekmann, S., Leblanc, T., Van Malderen, R., Nédélec, P., Petropavlovskikh, I., Steinbrecht, W., Stübi, R., Tarasick, D. W., and Torres, C. (2022). Impact of the COVID-19 economic downturn on tropospheric ozone trends: an uncertainty weighted data synthesis for quantifying regional anomalies above western North America and Europe. *AGU Advances*, e2021AV000542.

Chang, K.-L., Cooper, O. R., Gaudel, A., Petropavlovskikh, I., and Thouret, V. (2020). Statistical regularization for trend detection: An integrated approach for detecting long-term trends from sparse tropospheric ozone profiles. *Atmospheric Chemistry and Physics*, 20(16):9915–9938.

Cooper, O. R., Parrish, D. D., Stohl, A., Trainer, M., Nédélec, P., Thouret, V., Cammas, J.-P., Oltmans, S., Johnson, B. J., Tarasick, D., Leblanc, T., McDermid, I. S., Jaffe, D. A., Gao, R., Stith, J., Ryerson, T., Aikin, K., Campos, T., Weinheimer, A., and Avery, M. A. (2010). Increasing springtime ozone mixing ratios in the free troposphere over western North America. *Nature*, 463(7279):344–348.

Fujiwara, M., Shiotani, M., Hasebe, F., Vömel, H., Oltmans, S. J., Ruppert, P. W., Horinouchi, T., and Tsuda, T. (2003). Performance of the meteolabor ?snow white? chilled-mirror hygrometer in the tropical troposphere: Comparisons with the vaisala rs80 a/h-humicap sensors. *Journal of Atmospheric and Oceanic Technology*, 20(11):1534–1542.

Gleser, L. J. (1981). Estimation in a multivariate "errors in variables" regression model: large sample results. *The Annals of Statistics*, pages 24–44.

Kleiber, C. and Zeileis, A. (2008). *Applied econometrics with R*. Springer Science & Business Media.

Li, T. (2002). Robust and consistent estimation of nonlinear errors-in-variables models. *Journal of Econometrics*, 110(1):1–26.

Logan, J. A. (1994). Trends in the vertical distribution of ozone: An analysis of ozonesonde data. *Journal of Geophysical Research: Atmospheres*, 99(D12):25553–25585.

Oltmans, S., Lefohn, A., Shadwick, D., Harris, J., Scheel, H., Galbally, I., Tarasick, D., Johnson, B., Brunke, E.-G., Claude, H., et al. (2013). Recent tropospheric ozone changes–a pattern dominated by slow or no growth. *Atmospheric Environment*, 67:331–351.

Oltmans, S. J., Lefohn, A. S., Harris, J., Galbally, I., Scheel, H., Bodeker, G., Brunke, E., Claude, H., Tarasick, D., Johnson, B., et al. (2006). Long-term changes in tropospheric ozone. *Atmospheric Environment*, 40(17):3156–3173.

Saunois, M., Emmons, L., Lamarque, J.-F., Tilmes, S., Wespes, C., Thouret, V., and Schultz, M. (2012). Impact of sampling frequency in the analysis of tropospheric ozone observations. *Atmospheric Chemistry and Physics*, 12(15):6757–6773.

Tiao, G., Reinsel, G., Pedrick, J., Allenby, G., Mateer, C., Miller, A., and DeLuisi, J. (1986). A statistical trend analysis of ozonesonde data. *Journal of Geophysical Research: Atmospheres*, 91(D12):13121–13136.