Supplementary Material

Section S1. Sampling sensitivity test for in-situ measurements

Even though high temporal and spatial variability in ozone is well recognized, the positive impact of abundant sample sizes on detectability of trends is often under-appreciated. In terms of detecting trends in the free troposphere several previous studies concluded that a sampling frequency of once per week generally fails to produce accurate monthly mean and trend values (Logan, 1999; Saunois et al., 2012; Chang et al., 2020).

Since the in-situ sampling scheme is infrequent and sporadic at most locations in this study, we use the IAGOS dataset collected above Africa which has the highest measurement density (more than 30 profiles in some individual months) to explore the impact of sample size on trend detection in the tropics. In order to provide a baseline reference at northern mid-latitudes, we also analyze the IAGOS data collected above Frankfurt, Germany. Table S2 provides monthly sample sizes from Africa and Frankfurt. Even though Africa has the most abundant IAGOS data in the tropics, the overall sample sizes are still small compared to Frankfurt. The following analysis focuses on observations in the free troposphere (700-300 hPa).

We compute mean absolute percentage error (MAPE) between the ozone trend inferred from the complete data record and from an ensemble of trend estimates for randomly subsampled data sets. As in Chang et al. (2020), trend estimates are defined as accurate once MAPE falls below 5% with increasing sampling frequency r. Table S3 provides the MAPE obtained from 1000 random subsamples composed of a fixed number of r profiles per month. The sampling strategy can be summarized as follows: if a given month has n profiles and the requested monthly sampling frequency is r, then 1) if $n \leq r$, we select all the profiles, this is fixed in each iteration; 2) if $n > r$, we select r profiles randomly in each iteration. From the table we see that 19 profiles per month are required to produce an accurate trend estimate over Frankfurt, which is consistent with Chang et al. (2020). However, over Africa, the decrease of the trend MAPE is slow and MAPE remains high even when the considered sampling frequency is increased because there is a sufficient number of IAGOS profiles (n>r) for just a small fraction of individual months.

The above finding is limited by the fact that we cannot meet the predetermined criterion for most cases in Africa (and the 5% criterion cannot be met). To determine the threshold for minimum sampling frequency for basic trend detection in the tropics, we further investigate the relationship between the magnitude of trends and the sampling frequency. In this case, basic trend detection refers to enough profiles to determine if there is a trend at a 2-sigma level, based on either the interquartile range (i.e. the 75% percentile) or tail (i.e. the worst-case scenario) of the sampling distribution, but it is not ideal for an accurate trend quantification. Figure S1 shows the distribution of median trends for a sampling frequency of 2, 4, 6, 8, 10 and 12 profiles a month, from 800 to 300 hPa with a 50 hPa vertical resolution. We can see the range of sampled trends becomes smaller when the sampling frequency is increased. Figures S2 and S3 show how the signal-to-noise ratios (i.e the ratio between the trend value and its uncertainty) of sampled trends vary with different sampling frequencies at 800 to 300 hPa. These figures reveal many considerations regarding the relationship between sampling frequency and the magnitude of trends:

1. If the magnitude of the trend is strong (e.g. > 3 nmol mol^{-1}/decade at 800 hPa), the trend can be detected at a low sampling frequency: 2 and 6 profiles per month are required for basic trend detection in 75% samples and the worst-case scenario (i.e. even for the worst case, the trend can be detected), respectively.
2. If the magnitude of the trend is moderate (e. g. between 1 and 2 nmol mol^{-1}/decade at 600 hPa):
   7 and 15 profiles per month are required for basic trend detection in 75% samples and the worst-case scenario, respectively.
3. If the trend is weak (e.g. around 1 nmol mol^{-1}/decade at 700 hPa), a high sampling frequency is required to detect the weak signal: 14 profiles per month are required for basic trend detection in 75% samples, and the worst-case scenario cannot be prevented in this analysis.
4. For pressure surfaces with weak and highly uncertain trends (e.g. 350 and 500 hPa, Figure S1), the same conclusion can be generally drawn at either low or high sampling rates.

Based on the above discussion, a typical sampling frequency of once per week is only sufficient for detection of very large trends (e.g. > |3| nmol mol^{-1}/decade), which are not common in the free troposphere. We also conclude that a sampling frequency of 7 profiles per month is sufficient for basic trend detection of tropospheric ozone in the tropics, when the magnitude of a trend is above |1| nmol mol^{-1}/decade, but additional data are required for accurate quantification. It should be noted that natural variability also plays a role in trend detection, but its impact is expected to be more pronounced when we conduct sensitivity analyses on varying lengths of the data record, which is beyond the scope of the current analysis. Even though the influence between natural variability and sampling frequency is typically inseparable, by focusing on the same data set and same data length, the impact of natural variability should be weak on this sensitivity analysis. In monitoring long-term changes, the first problem is to detect a trend (as we investigated in this analysis). Once the presence of a trend is established, any additional information will help us to improve the accuracy and precision of trend detection.

Section S2. The OMI/MLS measurements and drift corrections

The Ozone Monitoring Instrument (OMI) and Microwave Limb Sounder (MLS) are two of four instruments on board the Aura spacecraft which is flown in a sun-synchronous polar orbit at 705 km altitude with a 98.2°inclination. The Aura spacecraft was launched 15 July 2004 and has an equatorial crossing time of about 1:45 pm (ascending node). Both OMI and MLS instruments are still providing ozone measurements as of late 2023 which has yielded a nearly 20-year record of tropospheric ozone for evaluating global trends and other applications. In this study we focus on the 2004-2019 time period.

OMI/MLS tropospheric column ozone (TCO) is derived using the residual technique of Fishman et al. (1990). Fishman et al. (1990) originally subtracted Stratospheric Aerosols and Gas Experiment (SAGE) stratospheric column ozone (SCO) from Total Ozone Mapping Spectrometer (TOMS) total ozone measurements. We apply the same approach where Aura MLS SCO is subtracted from coincident Aura OMI total column ozone to derive TCO. The OMI/MLS algorithm is discussed in detail by Ziemke et al. (2006), And here we only summarize this method. First, along-track measurements of daily MLS profile ozone are vertically integrated in pressure from the top of the atmosphere down to the tropopause pressure to measure SCO. National Centers for Environmental Prediction (NCEP) re-analyses are incorporated for tropopause pressure using the standard World Meteorological Organization (WMO) 2K km^{-1} lapse-rate definition. Next, a spatial 2D (Gaussian + longitudinal) interpolation is used to fill in between the MLS SCO orbital-track measurements. Daily TCO is then determined by subtracting these SCO fields from OMI total column ozone fields. Finally, OMI/MLS TCO daily maps are averaged monthly to produce the final TCO product.

The OMI/MLS product has a high sampling frequency, as shown in Figure S4 and Figure S5. Prior to 2009 each 5° x 5° grid cell had 300-500 measurements per month in the tropics; this number decreased to 200-400 measurements per month after the row anomaly took effect (described below). The measurement uncertainty (one standard deviation) of the OMI/MLS product is approximately 7 DU for a daily retrieval at 1° x 1.25° resolution, or approximately 2 DU at 5° x 5° resolution. It is reasonable to ask if this measurement uncertainty impacts the calculation of long-term trends from the OMI/MLS product. This question is addressed by the statistical field of error analysis (Grubbs, 1973; Taylor and Thompson, 1982; Moffat, 1988; Rabinovich, 2006; Buonaccorsi, 2010; Hughes and Hase, 2010). According to error analysis theory, if measurement uncertainty occurs randomly then the errors across a large sample size will cancel out and have little impact on the mean; in our case we are considering monthly mean values based on 200-400 OMI/MLS retrievals across a 5° x 5° grid cell. Given the very large sample size of the 5° x 5° OMI/MLS product the errors associated with measurement uncertainty cancel out and have little impact on the mean, and therefore little impact on the trend. Figure S6 below illustrates this concept using the ozonesondes above Debilt, The Netherlands (one profile per week), Uccle, Belgium (three profiles per week), and the IAGOS aircraft profiles above Europe (multiple profiles per day). All three data sets report clear positive trends for the period 1994-2019 based on monthly means produced from all available profiles (Figure S6a). In the next step random errors of 10% (representing measurement uncertainty) are imposed on all profiles. Figure S6b shows that the uncertainty of the monthly means increases slightly at Debilt and Uccle, but the uncertainty is almost unchanged for the IAGOS ensemble (due to the far greater sampling rates); the trend values at all three locations are almost unchanged, with only very slight increases in the 95% confidence intervals and *p*-values. In the final step random errors of 20% are imposed on all profiles (Figure S6c). These errors produce greater uncertainty of the monthly means for all three records, but the impact is greatest at Debilt which has the lowest sampling rate. Even though the imposed errors are relatively high, the overall trend values remain almost unchanged. The uncertainty of the trend values increases at all three sites, but the *p*-values remain below 0.05, and the impact is least for the IAGOS ensemble.

The OMI/MLS TCO measurements over time have encountered instrument drift and other long-term quality issues including an OMI row-anomaly which became a large problem in late

January 2009 and which still continues (Torres et al., 2018, and references therein). The row anomaly was caused by a physical obstruction in the optical path of the OMI instrument, resulting in about 1/3 of pixel measurements being flagged for non-use beginning in January 2009. Ziemke et al. (2019) discusses previous ozonesonde evaluation of offset and drift/row-anomaly corrections for the OMI/MLS TCO product. For that study corrections were made to include a mean +2 DU offset adjustment and a global -1.0 DU decade^{-1} drift adjustment. We have recently made a further adjustment to the OMI/MLS TCO long record guided by comparisons with ozonesondes, ground-based Brewer/Dobson total ozone, and OMI convective-cloud differential (CCD) tropical TCO measurements. The sonde analyses indicated that an additional correction of about -0.6 ± 0.38 DU decade^{-1} over the long record be applied to all latitudes. The ozonesonde analysis includes measurements from the Southern Hemisphere Additional OZonesondes (SHADOZ) network (Thompson et al. 2017; Witte et al. 2017, 2018; Sterling et al., 2017) and measurements from the World Ozone and Ultraviolet Radiation Data Center (WOUDC) and Network for the Detection of Atmospheric Composition Change (NDACC) (deMazière et al., 2018). After combining this -0.6 DU decade^{-1} correction with the previous $^{-1}$ DU decade^{-1} correction, the total drift correction is now equivalent to about -3 DU total change over the long record. This overall -3 DU drift correction coincides closely with calculated drift error for OMI total ozone of about +3 DU (+1%) from ground-based Brewer and Dobson total ozone comparisons (G. Labow, personal communication, 2023). As an additional cross-check for the new adjustment, we also included comparisons with OMI CCD tropical TCO (Ziemke et al., 1998); this suggested an additional drift correction of about -0.5 ± 0.30 DU decade^{-1} which is close to the ozonesonde comparisons. Thus, all three of these analysis methods for evaluating positive drift in OMI/MLS TCO agree.

Section S3. Confidence scale for in situ trends

This section provides a detailed description of the factors taken into consideration when assigning a confidence level to the in situ ozone trends reported in Table 1.

- **Western Africa** (1994-2019): Data coverage is moderate (high sampling rate and moderate data gaps), combined with a low p-value associated with a strong trend, therefore **high confidence** is assigned to this region.
- **India** (1994-2019): Data coverage is moderate (moderate data gaps and moderate sampling rates), combined with a low p-value associated with a strong trend. According to Table A1, high confidence should be assigned. However, since both the number of data gaps and sampling rates are on the fuzzy area around our criteria between low and moderate data availability, **moderate confidence** is assigned to this region.
- **Samoa** (1994-2019): Data coverage is low (limited data gaps and low sampling rates), combined with a high p-value, so **very low confidence** is assigned to this region.
- **Natal + Ascension Island** (1994-2019): Data coverage is low (limited data gaps but low sampling rates), combined with a low p-value, so **moderate confidence** is assigned to this region.

- **Americas** (1994-2019): Data coverage is moderate (limited data gaps and moderate sampling rates), combined with a high p-value, so **low confidence** is assigned to this region.
- **Southeast Asia** (1994-2019): Data coverage is moderate (moderate data gaps and moderate sampling rate), combined with a low p-value, so **high confidence** is assigned to this region.
- **Malaysia/Indonesia** (1994-2019): Data coverage is low (moderate data gaps but low sampling rate), combined with a low p-value, so **moderate confidence** is assigned to this region.
- **Western Africa, India and Samoa** (2004-2019): Data coverage is low (short time period), combined with a high p-value, so **very low confidence** is assigned in these regions.
- **Natal + Ascension Island**, **Americas**, **Southeast Asia** and **Malaysia/Indonesia** (2004-2019): Data coverage is low (short time period), combined with a low p-value, so **moderate confidence** is assigned to these regions.

Section S4. Tables and Figures

Table S1. Description of IAGOS fleet with the airlines, airports and number of profiles for the three focus time periods in this study (1995-2019 for long-term trends calculation, 2004-2019 for satellite evaluation, and 2014-2019 for present-day ozone distribution). All data above all airports listed in the table are used to quantify the distribution and trends of tropical tropospheric ozone as well as the evaluation of the satellite data.

Region	Airline	Airport	N profiles [1995-2019]	N profiles [2004-2019]	N profiles [2014-2019]
Americas	Austrian	Punta Cana	11	0	0
	Lufthansa	Bogota	560	356	347
	Air France				
	Lufthansa	Saint Martin	89	75	32
	Air France				
	Lufthansa	Panama City	14	14	14
	Iberia	Guayaquil	4	2	2
	Lufthansa				
	Lufthansa	Lima	24	0	0
	Air France				
	Lufthansa	Maracay	1	0	0
	Lufthansa	San Juan	45	0	0
	Lufthansa	Antigua	31	0	0
	Iberia	San Jose	32	32	32
	Lufthansa				
	Lufthansa	Caracas	1214	633	85
	Air France				
	Lufthansa	Mexico City	52	3	3

	Air France				
	Air France	Cayenne	216	0	0
	Lufthansa Air France	Quito	72	1	1
	Lufthansa	Cali	2	0	0
	Air France	Recife	25	0	0
	Lufthansa	Santo Domingo	2	0	0
	Lufthansa	Porlamar	2	0	0
	Austrian	Puerto Plata	12	0	0
	Lufthansa Air France	Malabo	182	182	182
	Air France	Yaounde	47	16	6
	Lufthansa Air France	Libreville	31	5	2
	Lufthansa Air France	Abuja	376	355	351
	Air France	Ndjamena	25	25	23
	Air France Lufthansa Sabena	Abidjan	233	38	35
	Air France	Bamako	48	48	40
	Sabena Lufthansa Air France	Lagos	761	441	396
	Air France	Ouagadougou	122	113	74
	Lufthansa	Tahoua	2	2	2
	Air France	Djibouti	11		9
	Lufthansa Air France	Port Harcourt	190	188	185
	Air France Lufthansa Sabena	Dakar	101	12	0
	Lufthansa	Bamenda	1	1	1
	Sabena	Entebbe	75		0
	Air France	Nouakchott	91	62	58
	Lufthansa	Khartoum	272		14
	Air France Air Namibia Lufthansa	Accra	139	66	43
	Air France	Niamey	123	113	56
	Lufthansa	Freetown	22	22	18

	Air France				
	Lufthansa	Jeddah	95		95
	Sabena	Douala	215	87	53
	Air France				
	Sabena	Lome	103	72	58
	Air France				
	Sabena	Cotonou	104	76	68
	Air France				
	Air France	Conakry	74	49	45
	Lufthansa				
	Sabena				
	Air France	Pointe-noire	28	28	28
	Sabena	Kigali	64		0
	Air Namibia	Brazzaville	40	31	29
	Air France				
	Air France	Kinshasa	102	19	17
	Sabena				
	Lufthansa	Luanda	254	210	89
	Air Namibia				
	Air France				
India	Cathay Pacific	Chennai	680	437	209
	Lufthansa				
	Sabena				
	Air France	Bangalore	32	32	32
	Austrian	Male	80	4	4
	LTU				
	Lufthansa				
	Austrian	Colombo	58	19	19
	Lufthansa				
	LTU				
	Austrian	Mumbai	177	56	28
	Cathay Pacific				
	Lufthansa				
	Air France				
	Cathay Pacific	Hyderabad	552	552	12
	Lufthansa				
Southeast Asia	Cathay Pacific	Cebu	18	18	18

	Lufthansa Air France Austrian Cathay Pacific China Airlines	Bangkok	1535	895	598
	China Airlines Austrian Cathay Pacific Lufthansa	Manila	191	191	146
	China Airlines Cathay Pacific Lufthansa Air France	Ho Chi Minh City	367	231	182
	China Airlines	Guam	8	8	8
	Cathay Pacific China Airlines	Hong Kong	80	80	80
Malaysia/Indonesia	Lufthansa	Paya Lebar	1	0	0
	Austrian	Darwin	3	0	0
	China Airlines Cathay Pacific Lufthansa Air France	Jakarta	113	86	61
	Cathay Pacific China Airlines	Surabaya	18	18	18
	China Airlines Cathay Pacific	Kuala Lumpur	208	192	139
	China Airlines	Denpasar	32	32	32

	Cathay Pacific				
	China Airlines	Singapore	265	143	92
	Cathay Pacific				
	Lufthansa				
	Air France				

Table S2. Number of IAGOS profiles by year and month above Africa (left panel) and Frankfurt, Germany (right panel).

	1	2	3	4	5	6	7	8	9	10	11	12
1997	0	0	0	6	52	37	48	65	40	58	28	8
1998	9	23	9	2	4	2	4	2	24	6	8	10
1999	22	18	13	15	12	20	32	8	2	30	15	12
2000	22	10	20	44	38	18	54	40	16	36	23	6
2001	18	14	32	66	0	0	0	4	2	29	10	12
2002	20	45	37	0	0	0	2	0	1	18	8	2
2003	19	32	4	14	30	12	38	47	42	66	46	10
2004	34	18	22	0	2	13	20	4	2	32	0	0
2007	12	6	0	0	0	0	0	0	0	0	0	4
2008	6	2	4	2	2	0	2	2	2	2	0	2
2009	8	2	22	4	0	0	8	0	0	0	0	0
2010	0	3	0	0	0	0	0	0	0	5	0	2
2011	0	4	6	0	0	0	21	2	4	2	6	10
2012	0	6	4	7	0	0	14	2	17	32	30	54
2013	70	21	21	8	0	33	34	6	2	29	32	36
2014	34	9	0	13	4	16	0	0	0	0	0	0
2015	0	4	14	44	38	0	10	19	38	56	33	68
2016	55	61	49	37	42	42	57	36	33	16	23	44
2017	62	40	18	54	80	86	51	60	86	38	56	19
2018	4	15	96	51	52	64	49	7	5	0	0	16
2019	6	14	28	24	28	31	29	30	24	28	24	14

	1	2	3	4	5	6	7	8	9	10	11	12
1994	0	0	0	0	0	0	0	62	101	96	98	116
1995	122	141	121	119	115	125	141	109	150	64	100	84
1996	139	123	39	123	153	128	121	158	101	102	145	103
1997	168	115	134	107	156	180	192	185	192	162	114	143
1998	185	146	158	137	165	126	121	187	181	160	141	143
1999	113	148	131	132	169	182	219	201	191	206	190	198
2000	177	163	172	218	209	167	196	168	151	155	135	131
2001	134	109	140	150	101	92	127	118	117	119	83	88
2002	68	109	104	46	45	13	74	96	117	126	130	57
2003	45	90	97	121	118	189	188	215	178	197	169	180
2004	158	128	136	112	108	172	212	184	111	126	58	142
2005	138	105	146	121	131	120	96	98	125	89	105	103
2006	71	68	73	70	78	78	129	83	109	117	113	116
2007	75	85	68	66	71	81	90	76	80	20	49	67
2008	80	67	56	29	69	78	54	70	64	70	88	115
2009	130	99	118	63	111	110	147	80	53	10	47	56
2010	9	46	6	2	4	4	4	4	15	40	49	54
2011	53	90	109	102	58	73	171	98	95	122	98	102
2012	16	63	85	138	38	13	110	86	104	79	109	145
2013	159	91	83	84	104	139	125	94	86	77	90	134
2014	109	85	71	126	103	132	104	63	36	55	36	10
2015	58	31	35	92	86	58	102	124	90	114	63	61
2016	64	74	52	95	148	111	150	145	120	89	85	135
2017	119	94	50	125	134	143	139	147	125	101	114	52
2018	46	23	66	62	78	89	85	59	42	51	0	47
2019	41	45	49	59	70	105	83	53	20	46	44	34

Table S3. The mean absolute percentage error (MAPE) values between the trend value derived from the full dataset and sampled trends are reported, based on quantile regression and free tropospheric observations (700-300 hPa) above Frankfurt and Africa. Sampled trends are generated by one thousand random samples for each of a predetermined number of profiles per month.

#profile per mon	1	2	3	4	5	6	7	8	9	10
Frankfu	69.1	39.3	28.1	18.0	16.7	15.3	14.3	12.4	11.7	9.1
Africa	93.9	85.0	79.6	71.9	64.3	55.2	54.9	53.1	50.1	48.9
#profile per mon	11	12	13	14	15	16	17	18	19	20
Frankfu	9.0	8.7	8.3	7.6	7.2	5.9	5.6	5.2	4.8	3.9
Africa	47.8	46.4	42.3	40.5	38.9	34.0	31.6	26.4	26.0	23.2

Table S4. Summary of the TTCO trends in nmol mol^{-1} decade^{-1} from IAGOS

The sampling column reports three numbers for the in situ data: i) the number on the top refers to the average number of profiles per months taking into account all the months with profiles, ii) the number in the middle refers to the percentage of months with data for the studied time-period (1994-2019 or 2004-2019), iii) the number in the bottom refers to the total number of profiles for the studied time period (1994-2019 or 2004-2019). We provide these numbers for a reference, but, for these three IAGOS regions, our final conclusions are based on the confidence scale for the fused (IAGOS + SHADOZ) results (See Table 1 in the main manuscript).

		1994-2019			2004-2019		
		Trends±2σ (nmol mol^{-1} decade^{-1})	p-value	Sampling	Trends±2σ (nmol mol^{-1} decade^{-1})	p-value	Sampling
IAGOS	Americas	2.07±0.51	<0.01	10.8 71.5% 2403	0.64±1.21	0.29	9.8 66.1% 1248
	Southeast Asia	4.66±0.46	<0.01	15.3 62.7% 2194	4.07±1.60	<0.01	16.2 61.1% 1423
	Malaysia/Indonesia	6.44±1.13	<0.01	8.0 44.4% 636	8.62±2.29	<0.01	9.3 47.2% 475

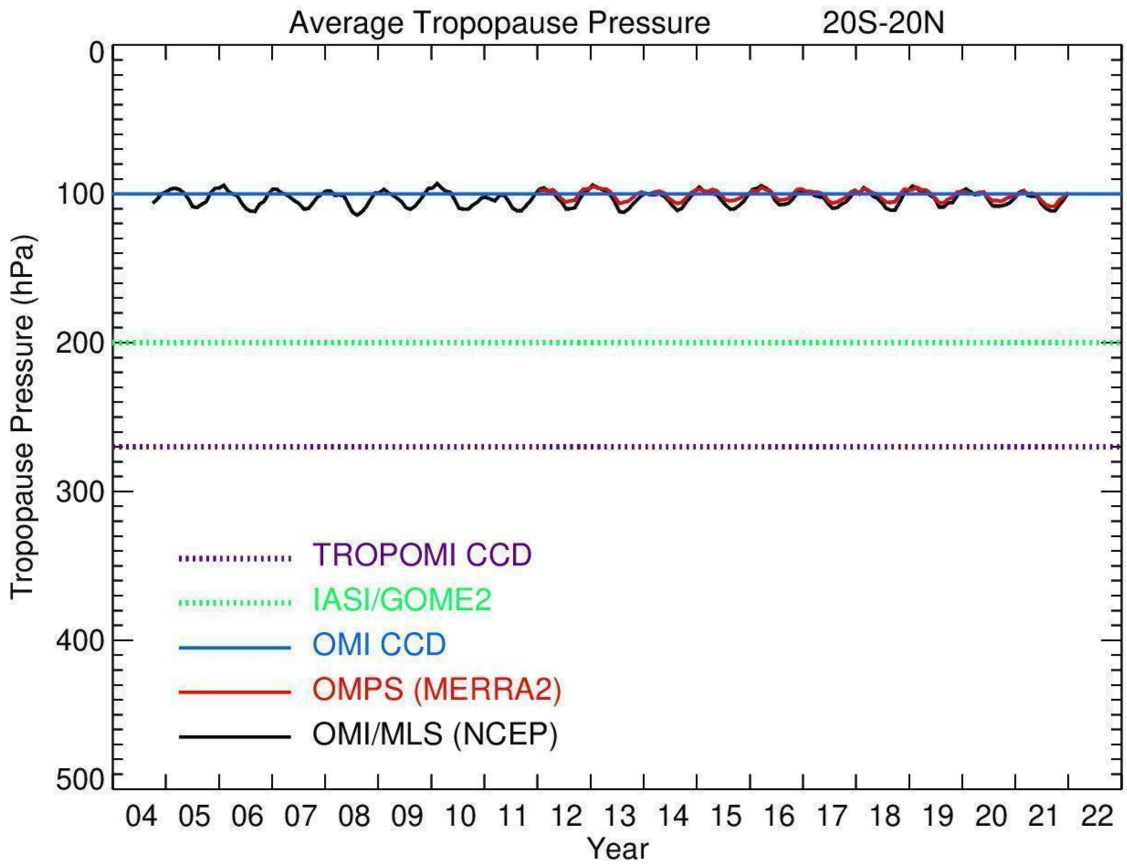


Figure S1. Time series of the monthly mean of the tropopause pressure level used to define the tropical tropospheric column ozone (TTCO) with satellite data.

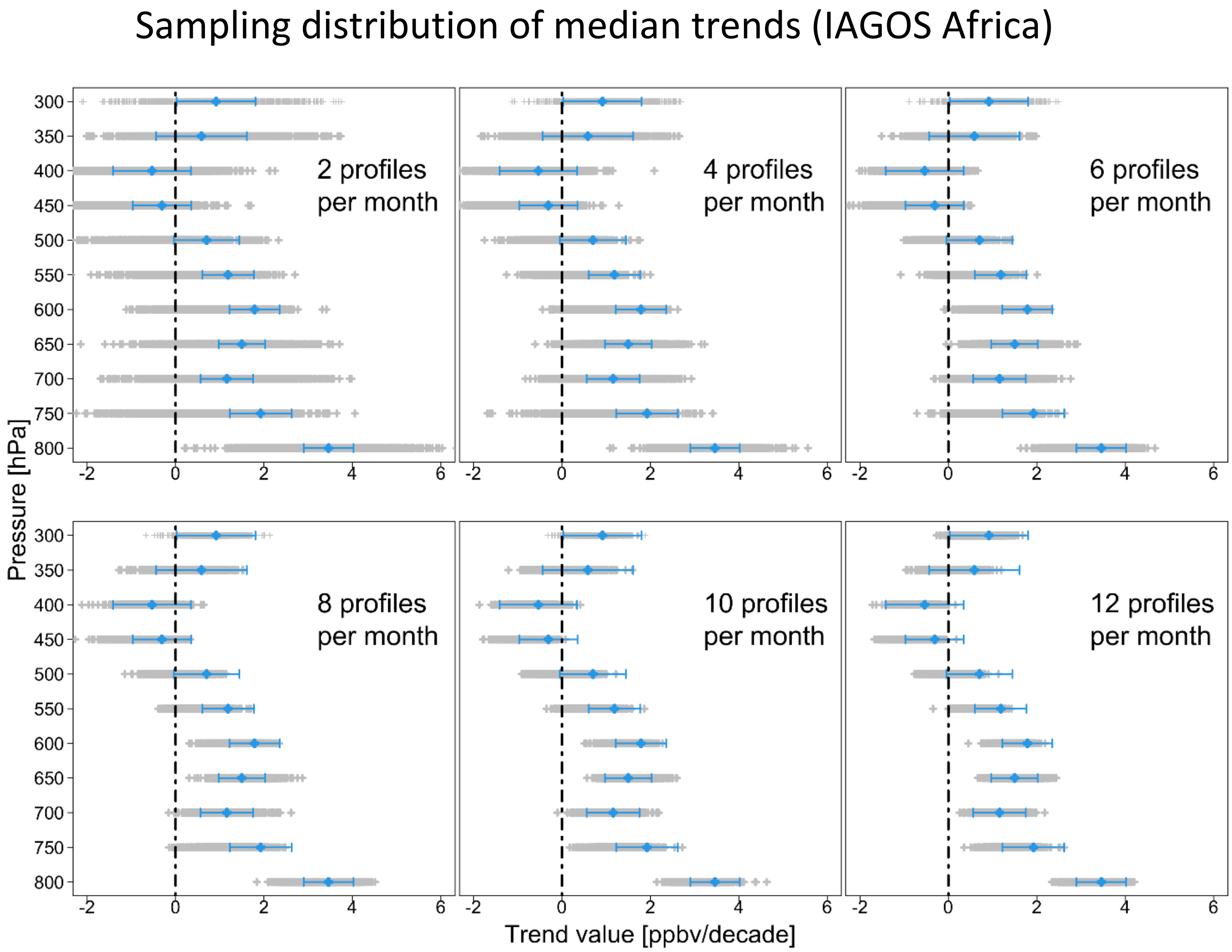


Figure S2. Sampling distributions of trends for a sampling frequency of 2 to 12 profile-per-month above Africa. The estimation is based on quantile regression. Blue crosses are the median trend estimates derived from all available data, the horizontal blue bars indicate the 95% confidence interval of the trend with full sampling, each gray cross represents the estimate produced by a random sampling from 1000 iterations.

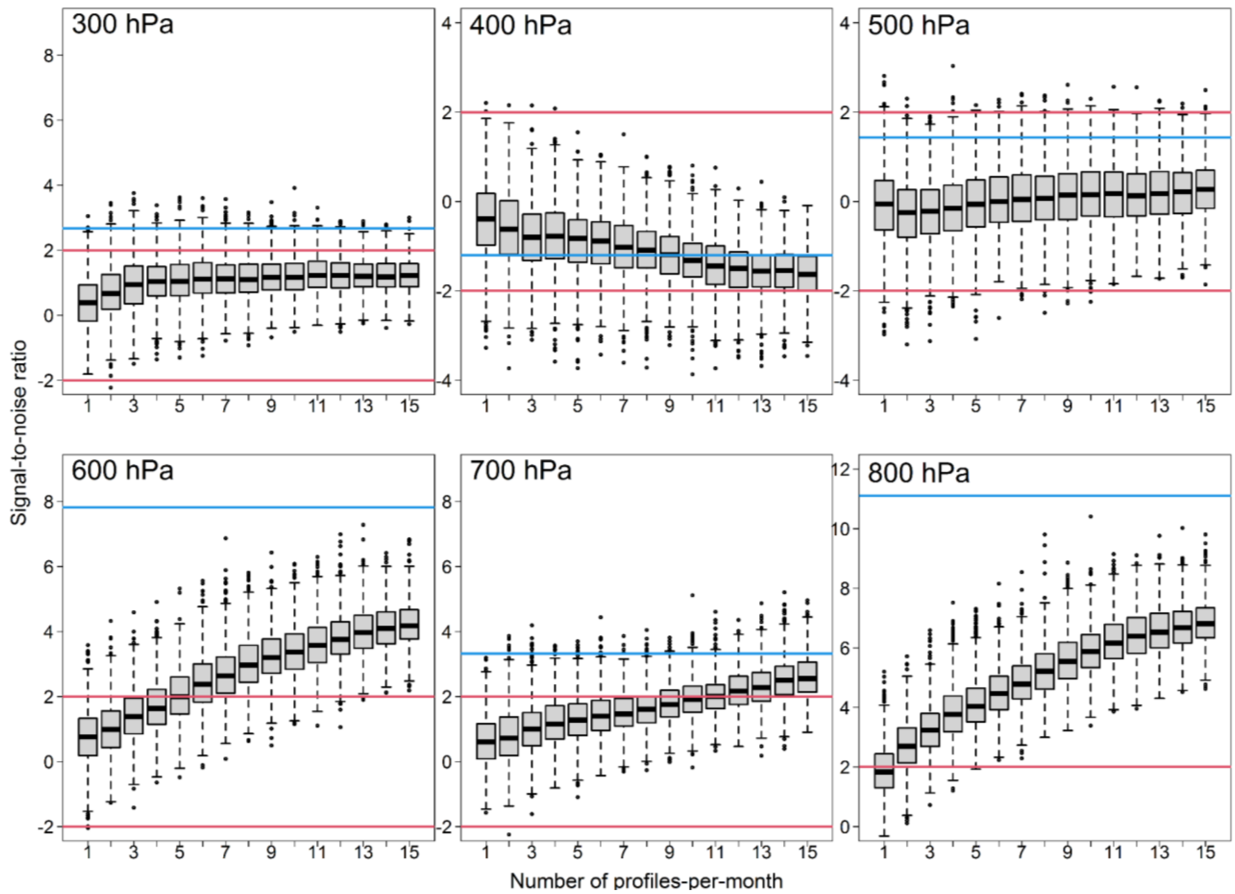


Figure S3. Sampled signal-to-noise ratios of the ozone trends. The ratios vary with different sampling frequencies at 800 to 300 hPa above Africa. Red lines (at signal-to-noise ratios of 2 and -2) represent the conventional trend detection threshold (i.e. 95% confidence level), and blue lines represent the SNR derived from all available data.

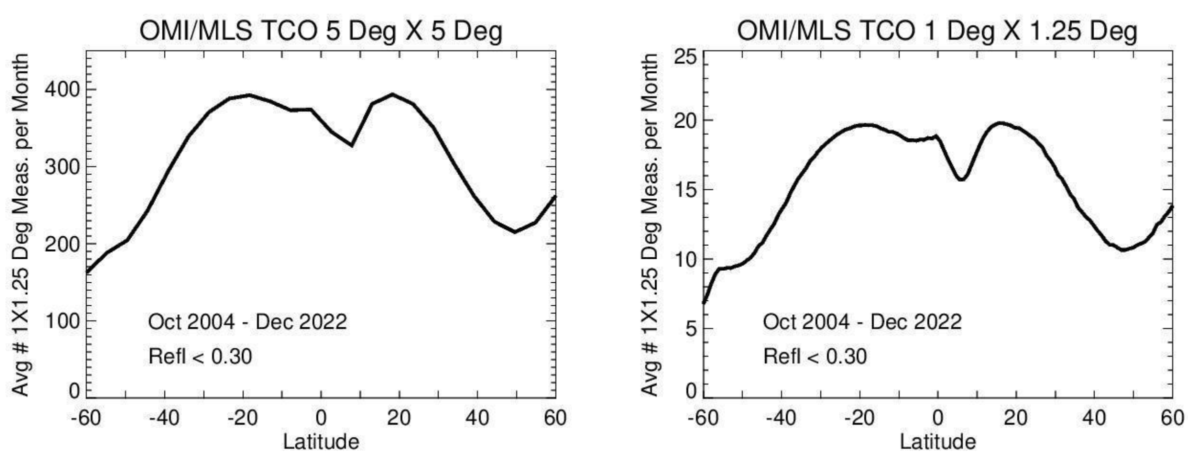


Figure S4. Number of 1° x 1.25° resolution OMI/MLS tropospheric column ozone measurements per month in a 5°x5° grid cell (left panel) or a 1° x 1.25° grid cell (right panel), by latitude. The data have been cloud-filtered using a low reflectivity threshold of R < 0.30, and the results are averaged across October 2004 to December 2022.

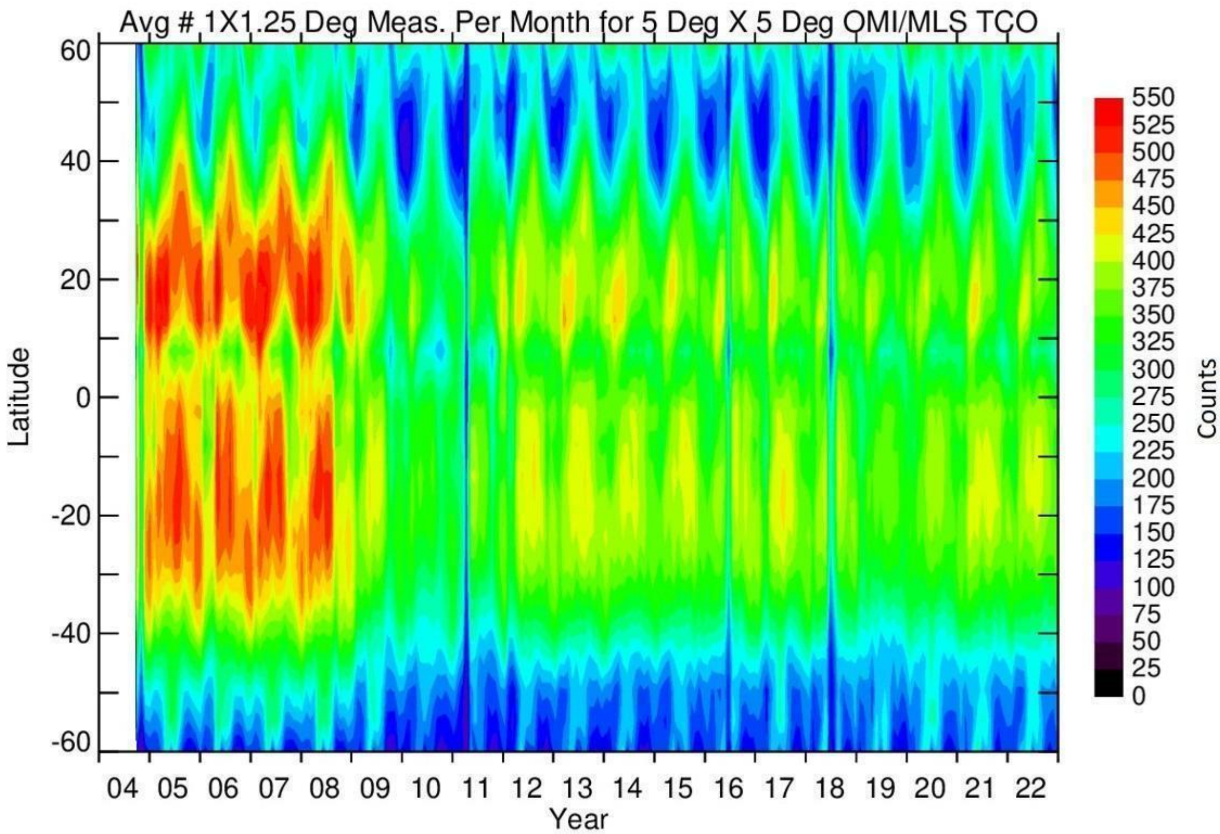


Figure S5. Hovmoller plot of average number of daily 1°x1.25° tropospheric column ozone (TCO) measurements per month within each 5°x5° grid cell, following R<0.30 cloud filtering. Starting January 2009 there are fewer measurements due to the row anomaly problem

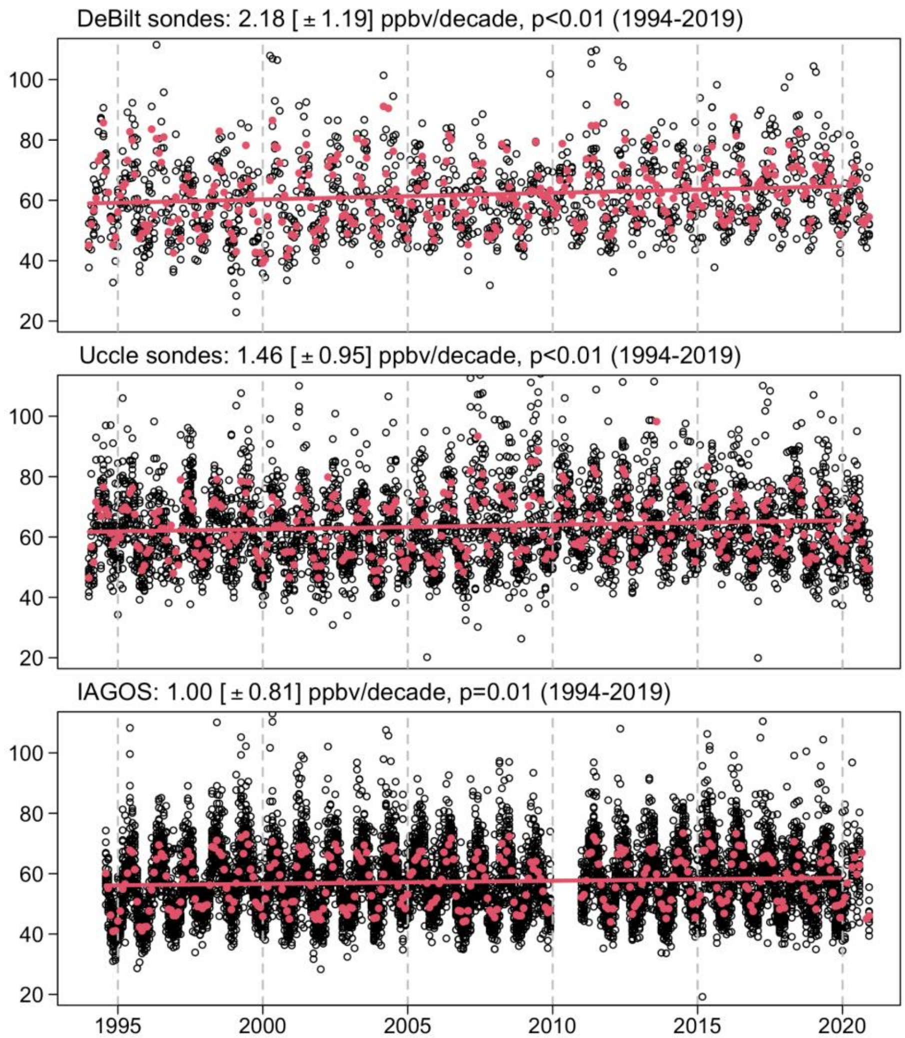


Figure S6a. Mid-tropospheric (700-300 hPa) ozone trends (1994-2019) at three European locations: Debilt, The Netherlands (top), Uccle, Belgium (center) and an ensemble of all IAGOS profiles above Europe (bottom). Each black point represents a mid-tropospheric observation (averaged over 700-300 hPa) from a single profile, while the red points represent monthly means (under the assumption of no measurement uncertainty). Also shown are the linear trends for 1994-2019, with 95% confidence intervals and *p*-values.

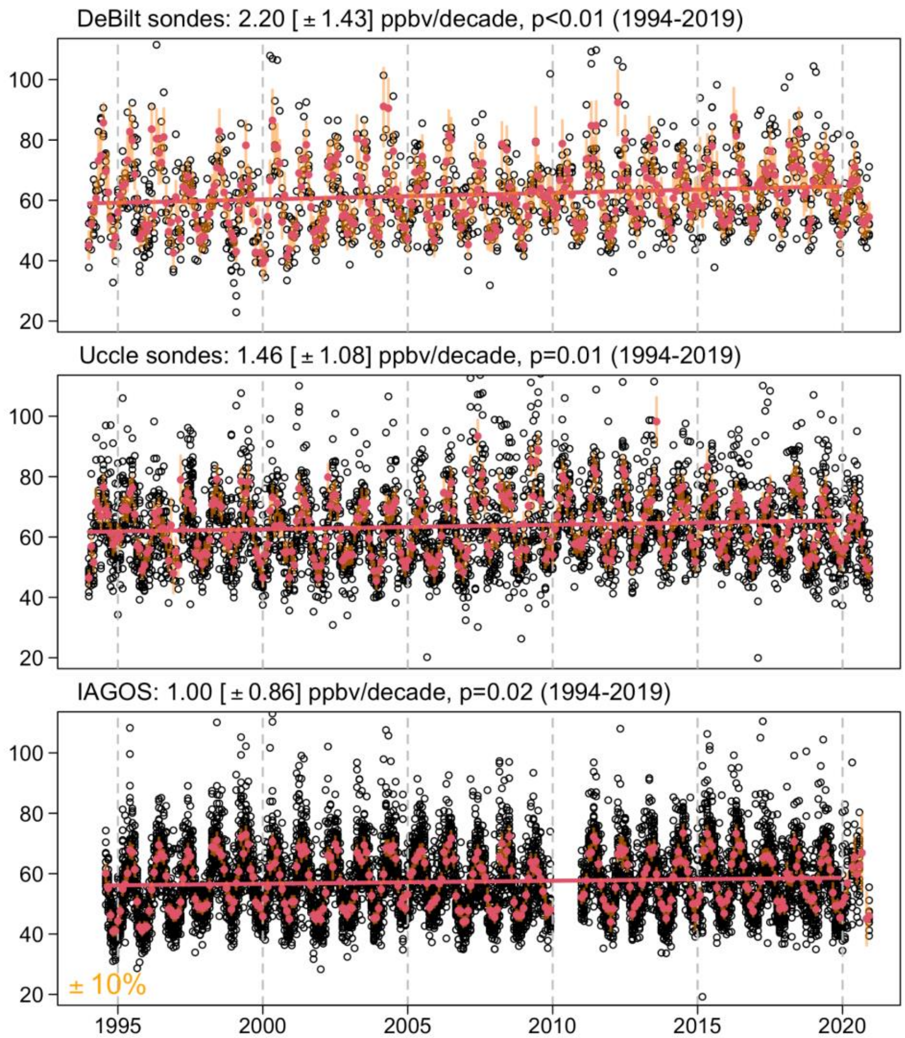


Figure S6b. As in Figure S6a, but with random noise of 10% imposed on each individual ozone value (the noise value is randomly selected from a normal distribution with 0 mean and SD as *y**0.10 (e.g. for a desired uncertainty of ±10%)). For each month, monthly means are produced from the corresponding noise-added observations. This procedure is repeated 10,000 times, and the 2.5th and 97.5th percentiles from the 10,000 noise-added monthly means indicate the 95% confidence interval of the means (shown with the orange bars on each monthly mean). Finally, 10,000 trend values are produced, and the mean and standard deviation become the final trend and sigma uncertainty reported in each panel of the figure.

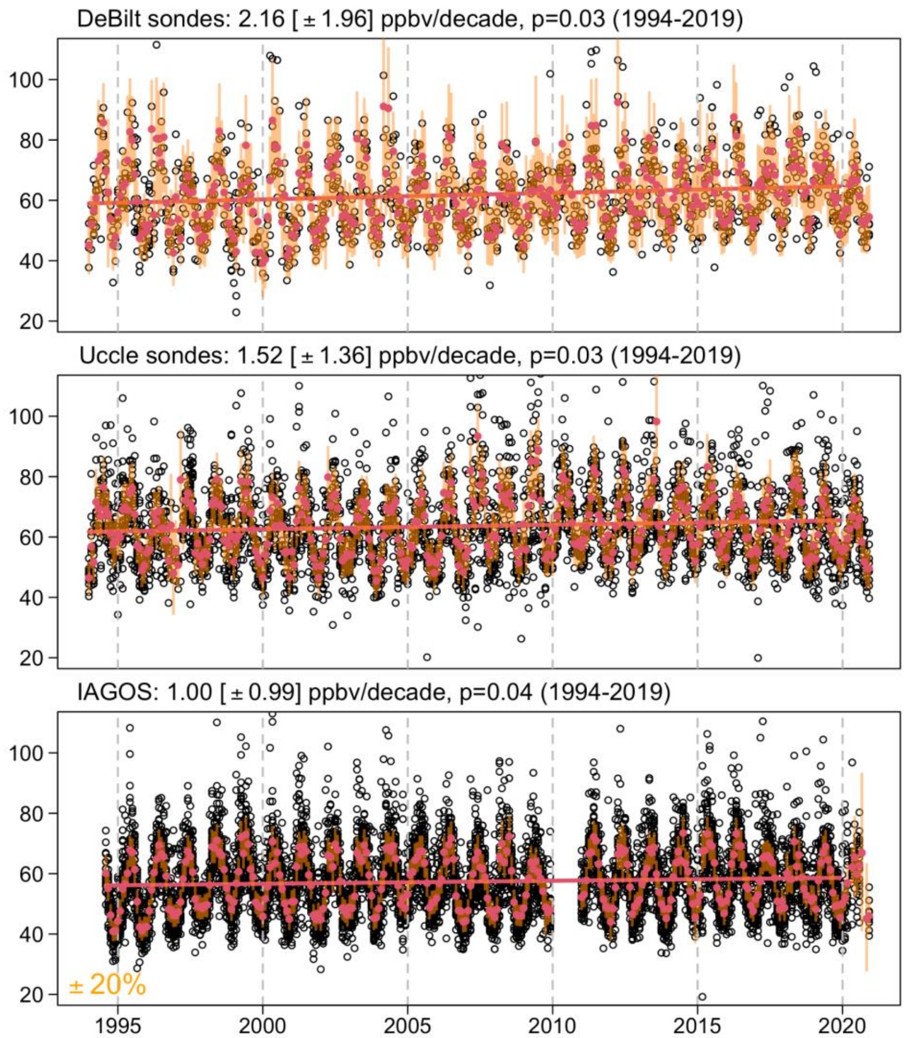
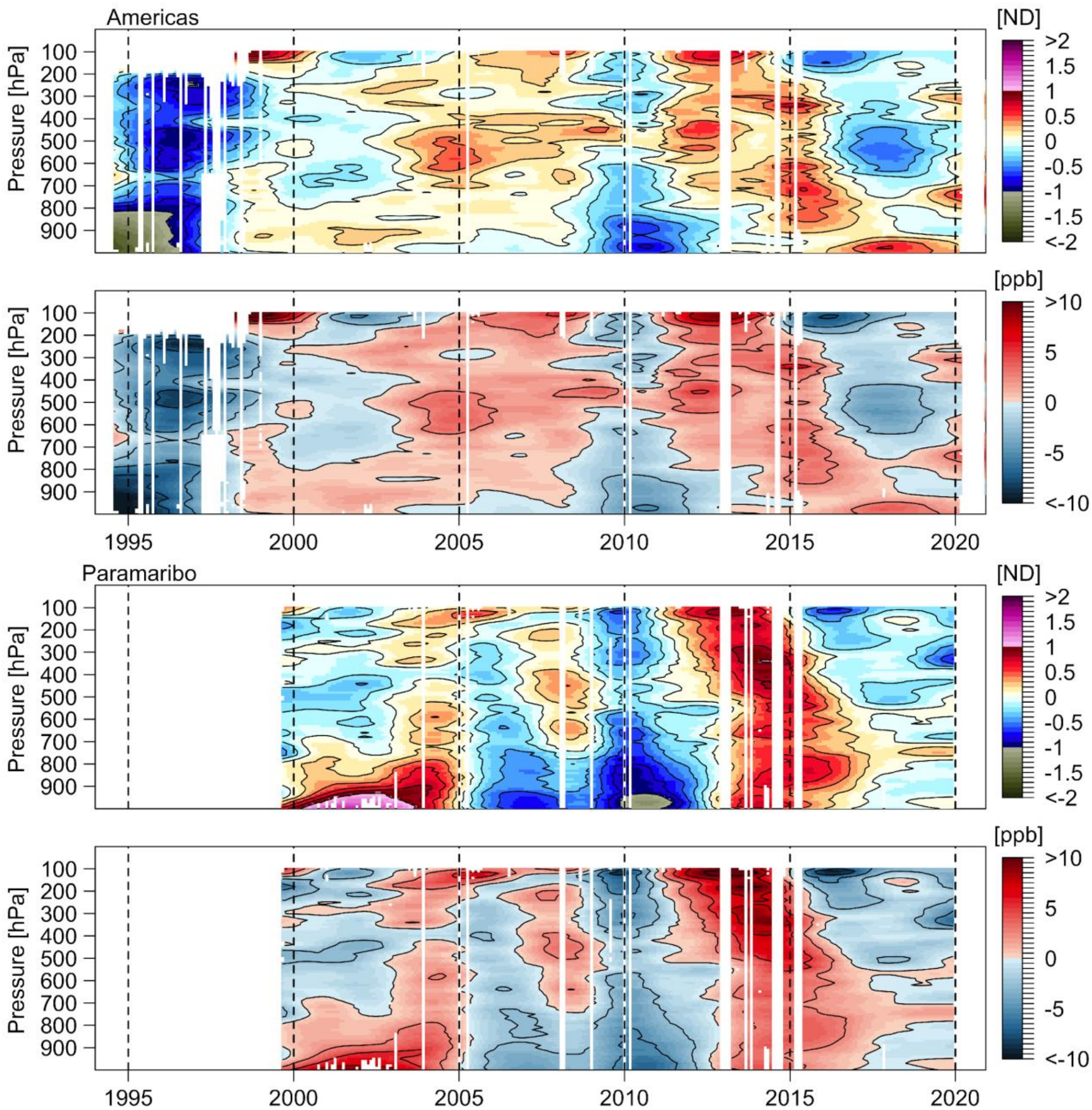


Figure S6c. As in Figure S6b but with 20% random noise imposed on each individual ozone value.



Americas
Paramaribo
Pressure [hPa]
[ND]
[ppb]
>2
1.5
1
0.5
0
-0.5
-1
-1.5
<-2
>10
5
0
-5
<-10
100
200
300
400
500
600
700
800
900
1995
2000
2005
2010
2015
2020

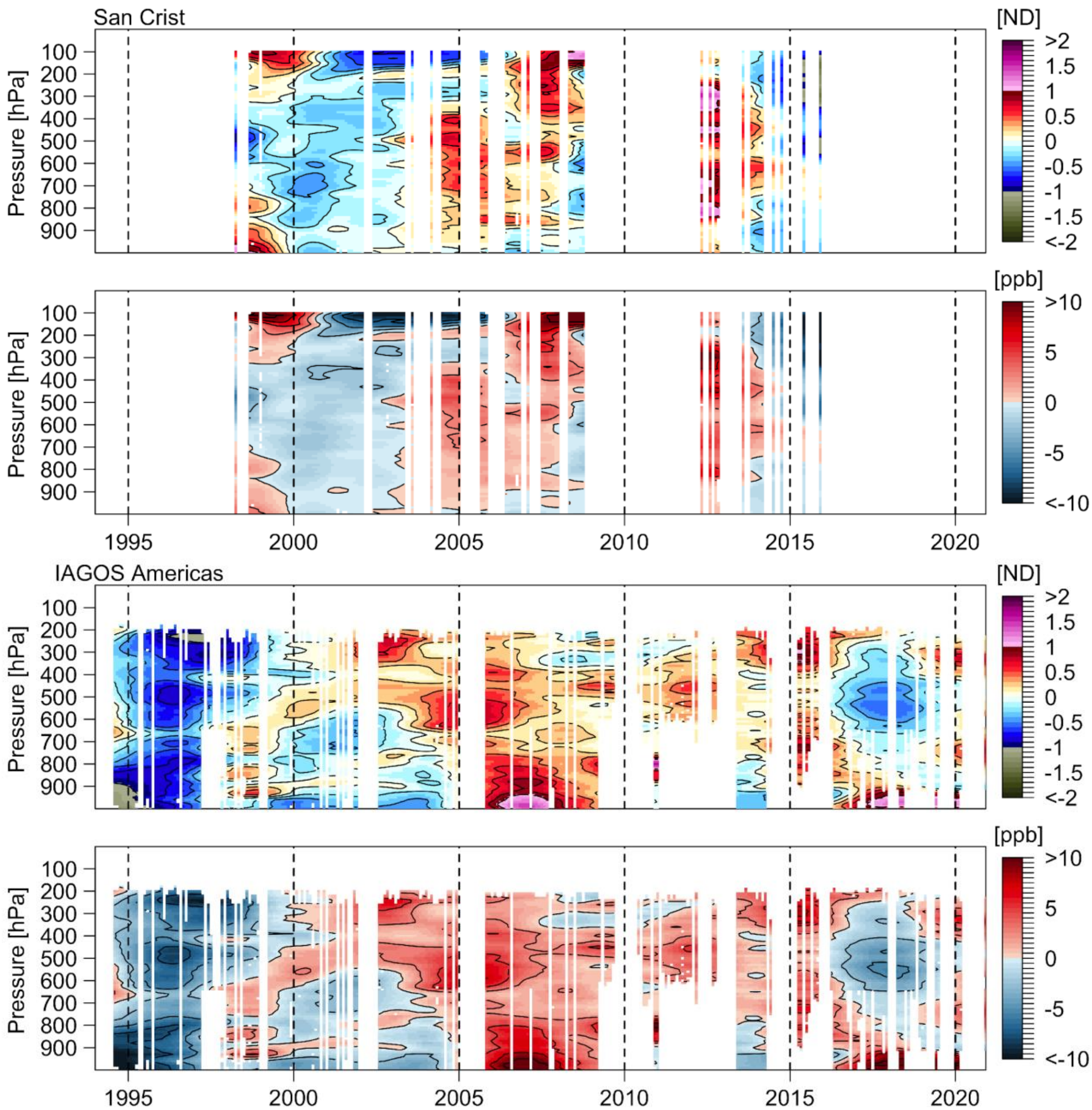
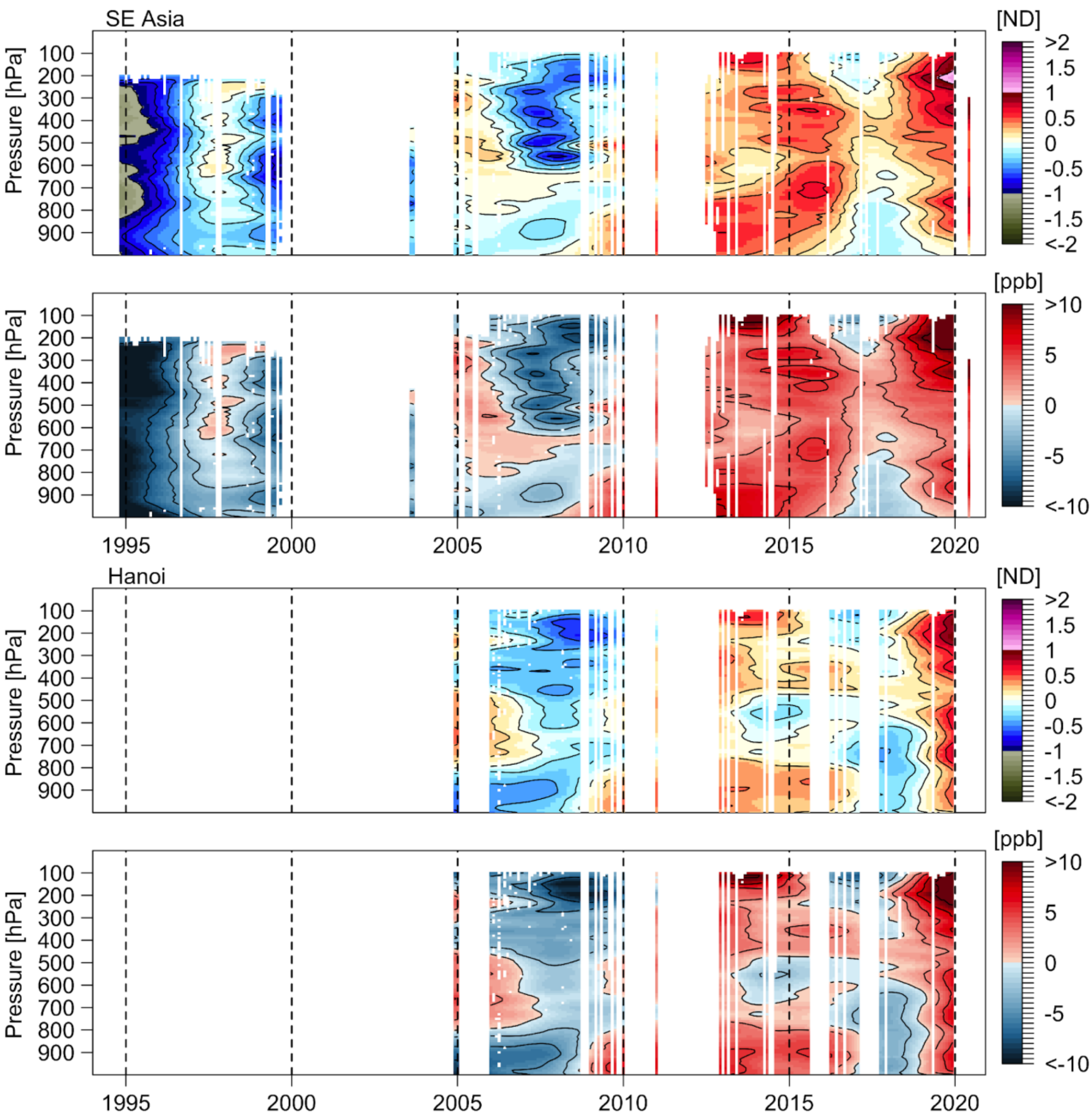


Figure S7. Ozone mean distributions above Americas based on normalized deviation. Panels show the results of the fused data set from IAGOS and SHADOZ (top), and for the SHADOZ individual sites (Paramaribo and San Cristobal) and IAGOS region (Americas).



SE Asia
Hanoi
[ND]
[ppb]
Pressure [hPa]
100
200
300
400
500
600
700
800
900
>2
1.5
1
0.5
0
-0.5
-1
-1.5
<-2
>10
5
0
-5
<-10
1995
2000
2005
2010
2015
2020

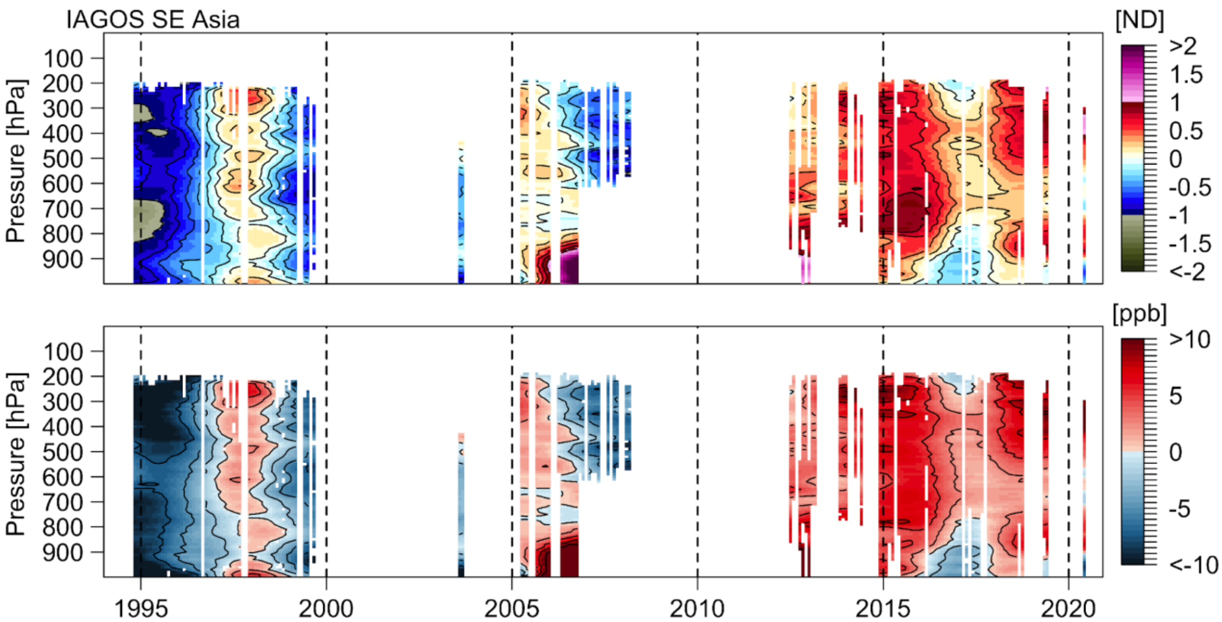
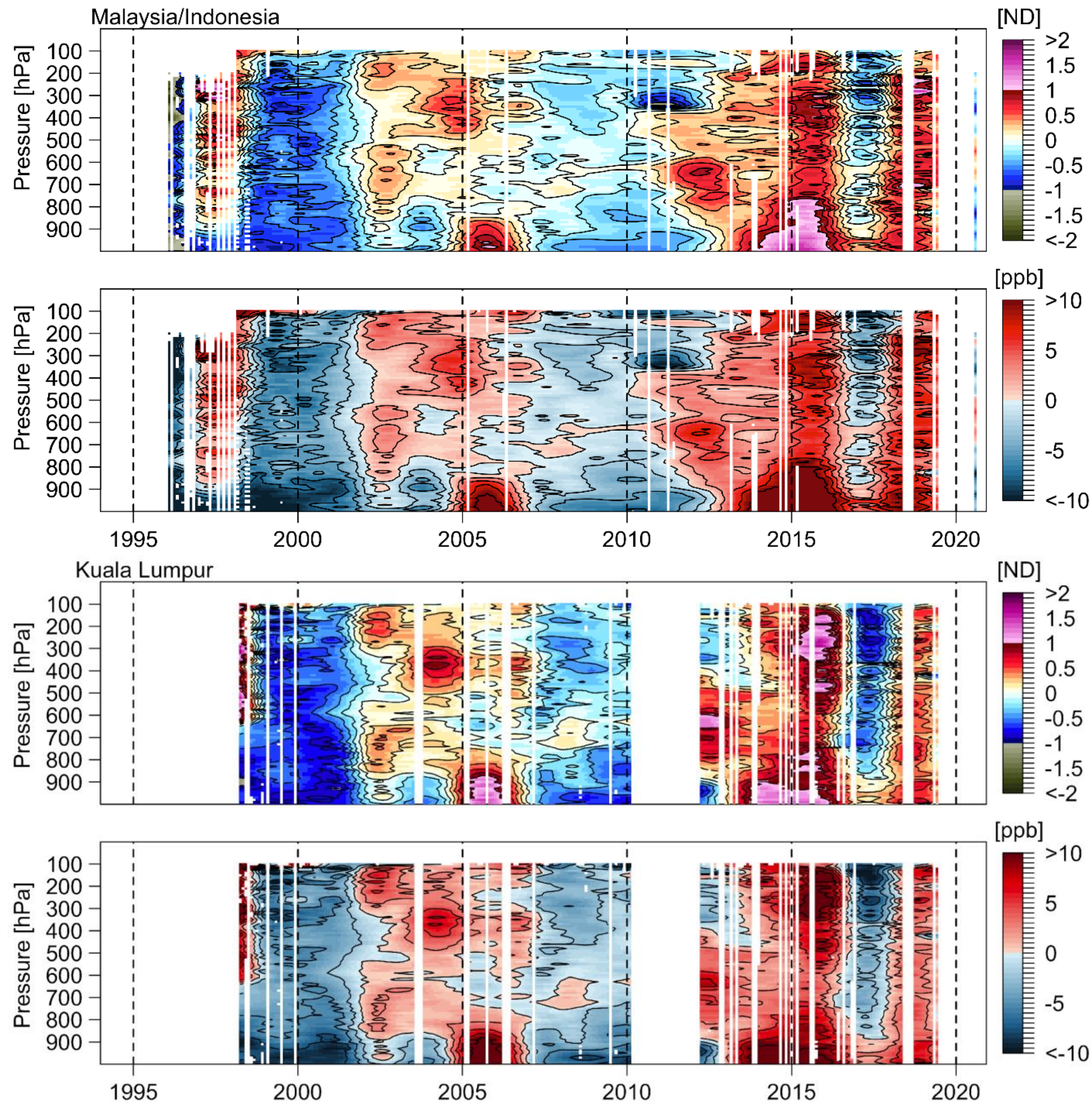


Figure S8. Same as Figure S7 but above Southeast Asia. The SHADOZ individual site used for the fused data is Hanoi.



Malaysia/Indonesia
Kuala Lumpur
Pressure [hPa]
100
200
300
400
500
600
700
800
900
[ND]
>2
1.5
1
0.5
0
-0.5
-1
-1.5
<-2
[ppb]
>10
5
0
-5
<-10
1995
2000
2005
2010
2015
2020

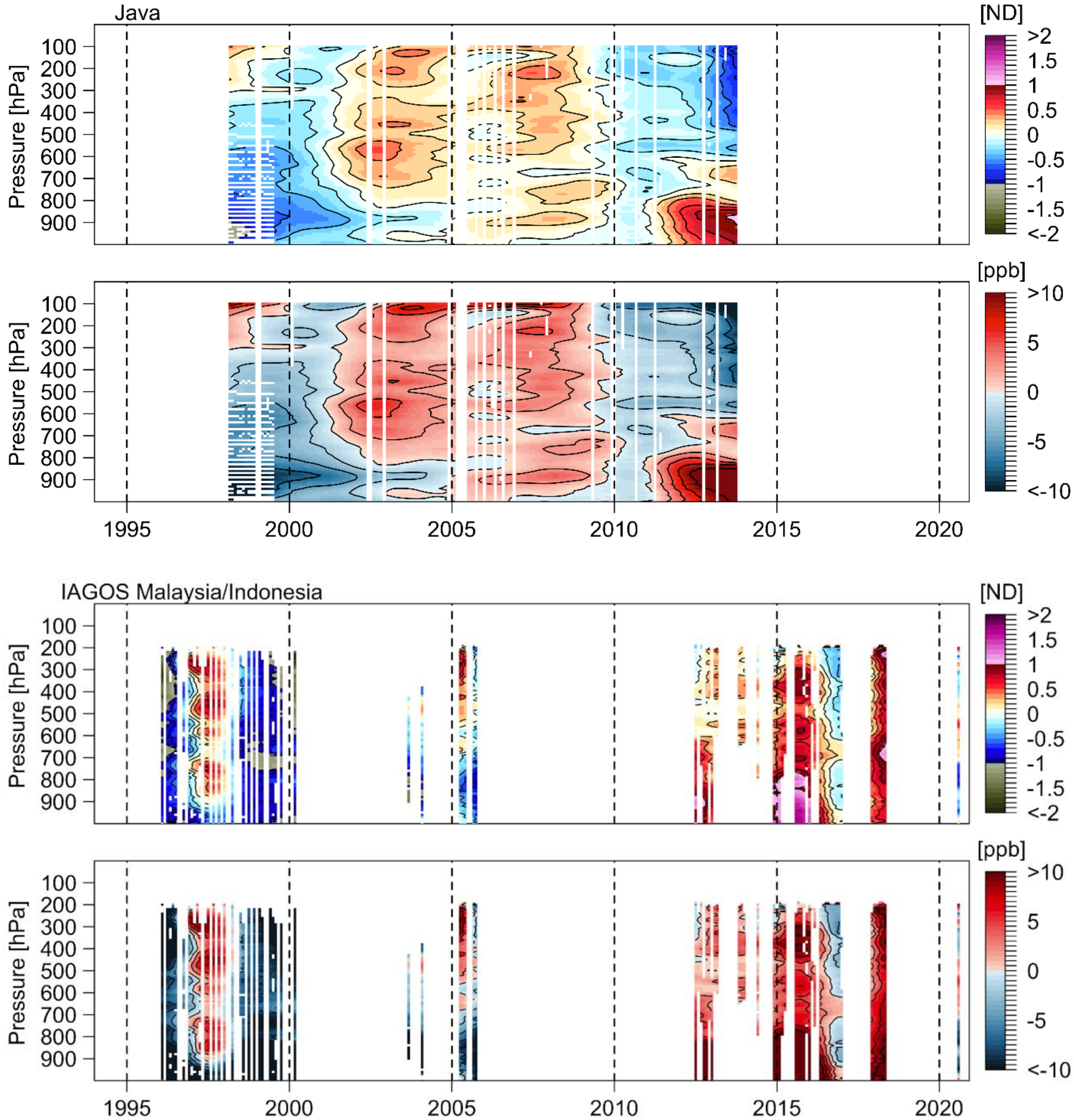


Figure S9. Same as Figure S7 but above Malaysia/Indonesia. The SHADOZ individual site used for the fused data is Kuala Lumpur and Watukosek (Java).

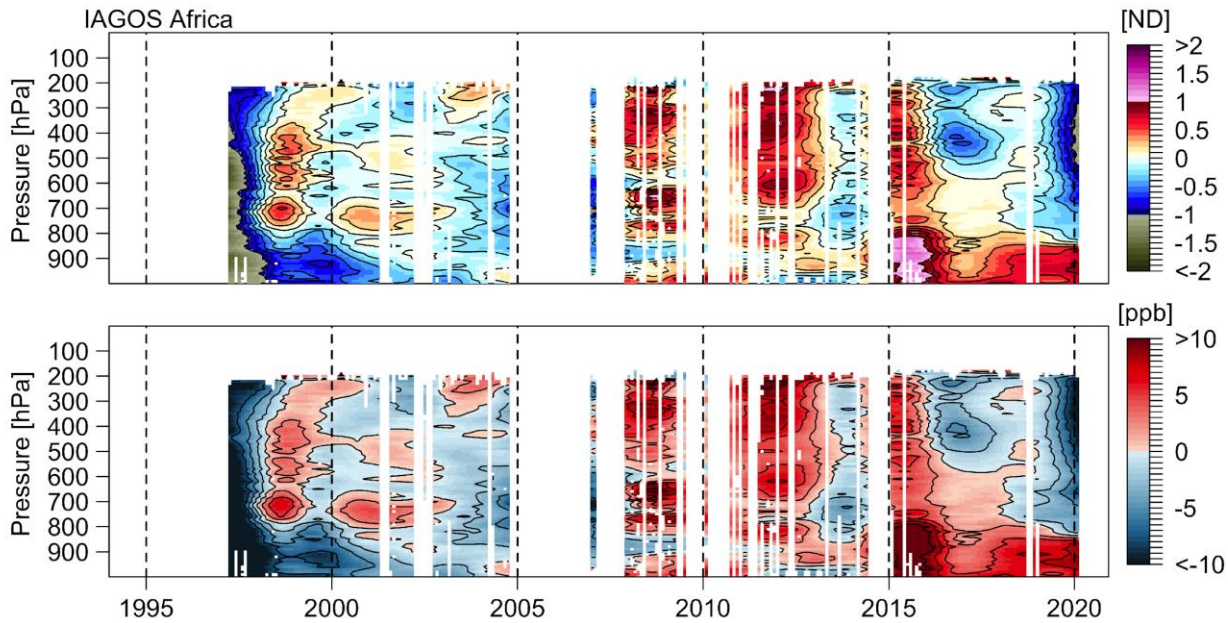


Figure S10. Same as Figure S7 but above western Africa. There are no SHADOZ data available in this region. We use only IAGOS data.

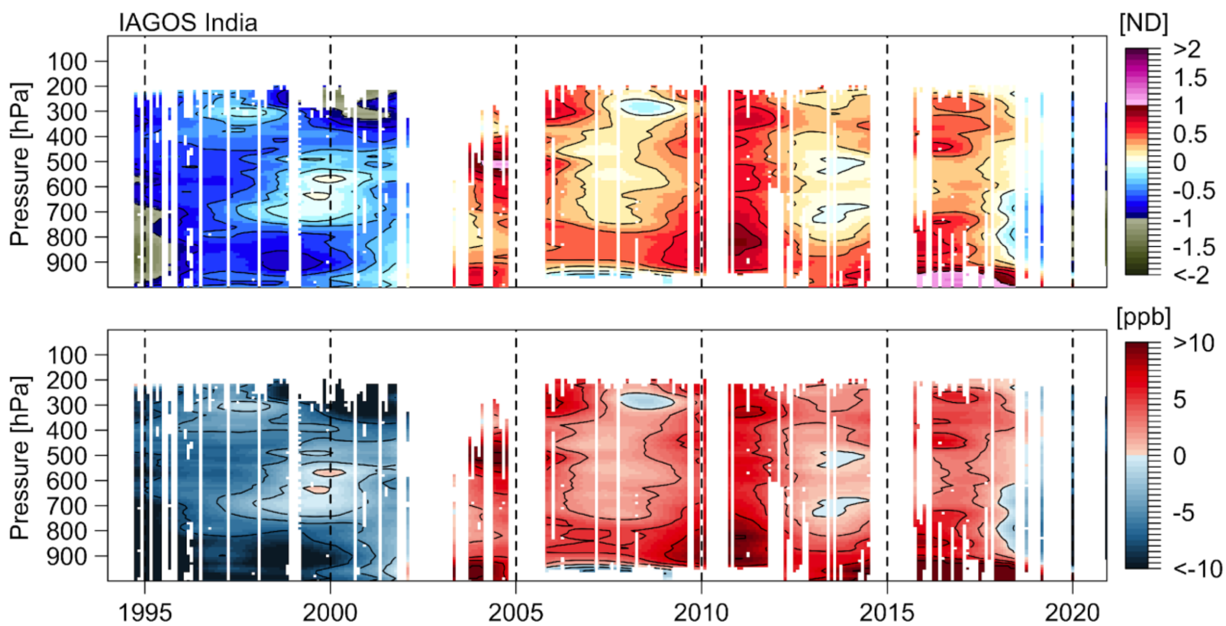


Figure S11. Same as Figure S7 but above India. There are no SHADOZ data available in this region. We use only IAGOS data.

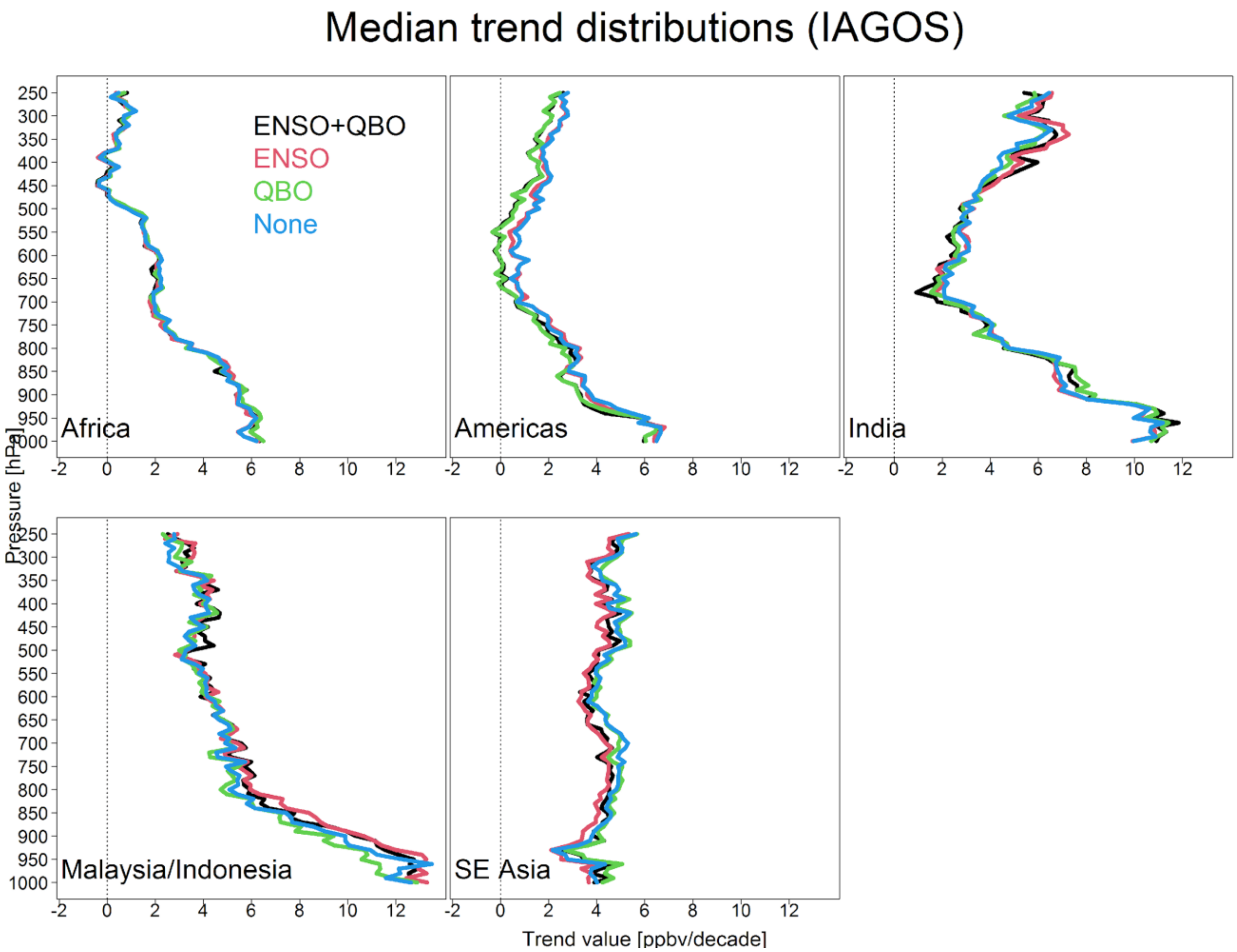
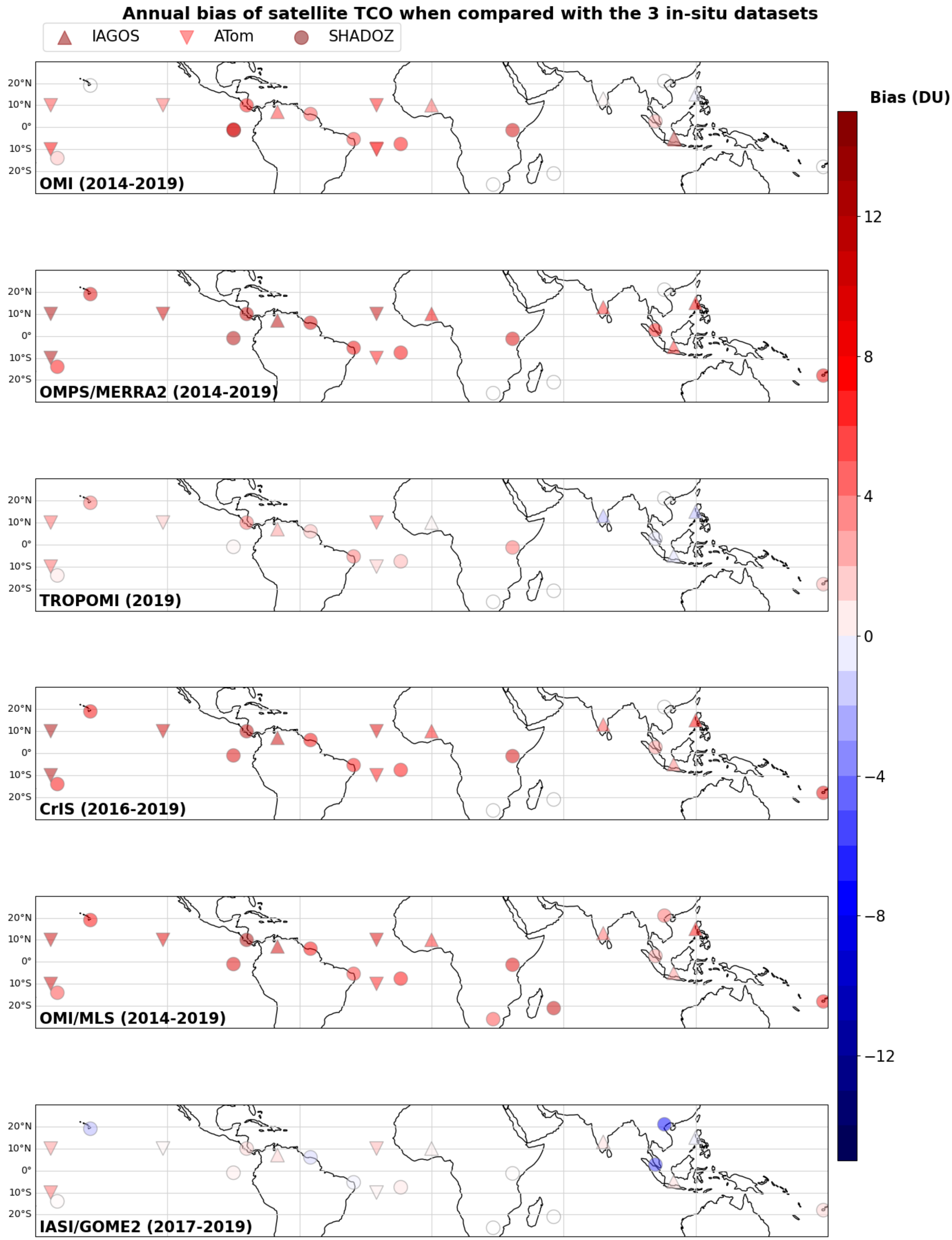


Figure S12. Vertical profiles of ozone trends (nmol mol^{-1}/decade) from a linear regression that considers climate variability such as ENSO (El Niño-Southern Oscillation) and QBO (quasi-biennial oscillation). The trends are reported over the five IAGOS regions: Africa, Americas, India, Malaysia/Indonesia and Southeast Asia.



Annual bias of satellite TCO when compared with the 3 in-situ datasets
IAGOS
ATom
SHADOZ
20°N
10°N
0°
10°S
20°S
OMI (2014-2019)
OMPS/MERRA2 (2014-2019)
TROPOMI (2019)
CrIS (2016-2019)
OMI/MLS (2014-2019)
IASI/GOME2 (2017-2019)
Bias (DU)
12
8
4
0
−4
−8
−12

Figure S13. Absolute annual mean biases of tropical tropospheric column ozone (TTCO in DU) of the six satellite products: OMI (2014-2019), OMPS/MERRA2 (2014-2019), TROPOMI (2019), CrIS (2018-2019), OMI/MLS (2014-2019) and IASI/GOME2 (2017-2019) against the three in situ TTCO up to 270 hPa: IAGOS (2014-2019), SHADOZ (2014-2019) and ATom (2016-2018).

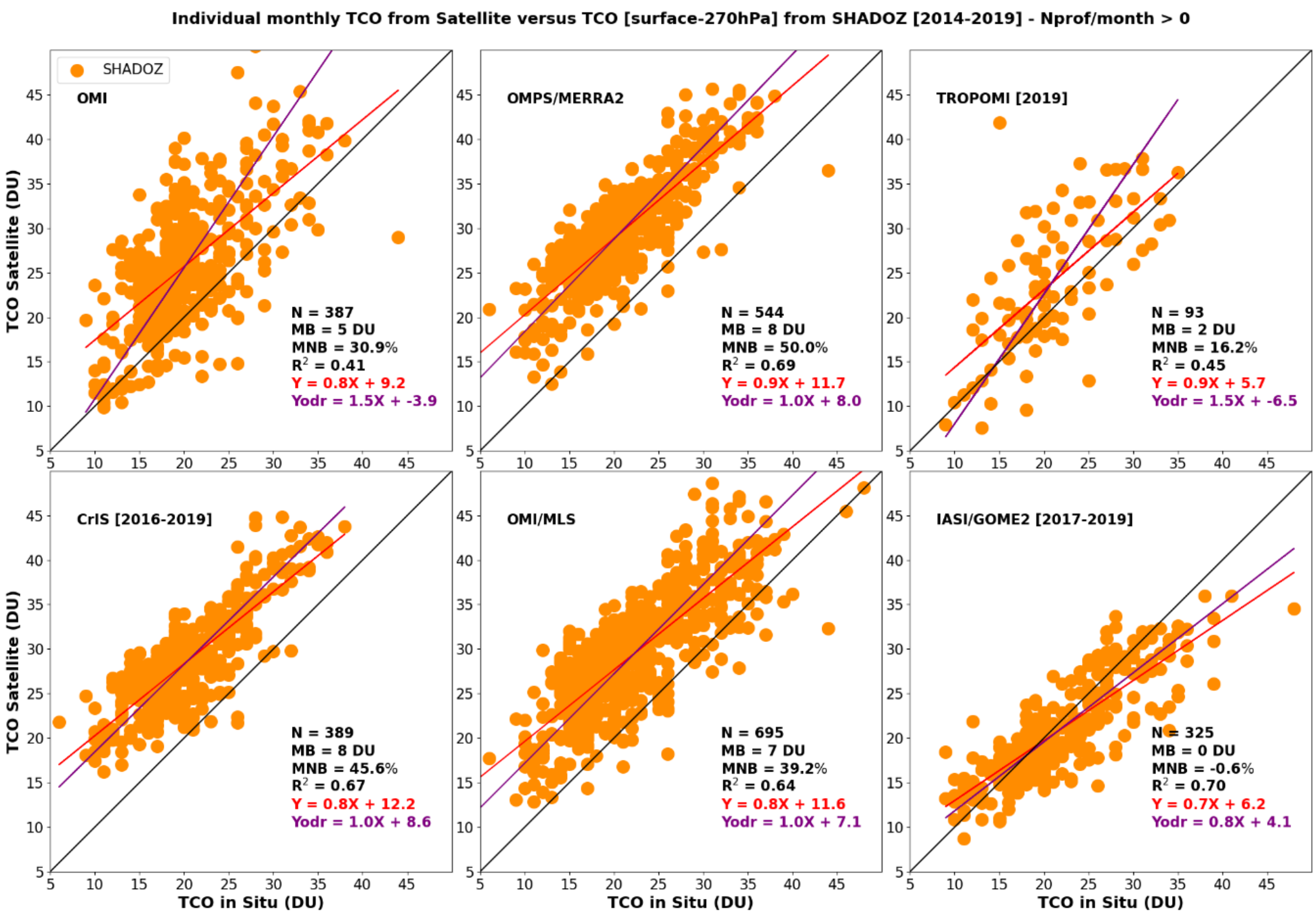
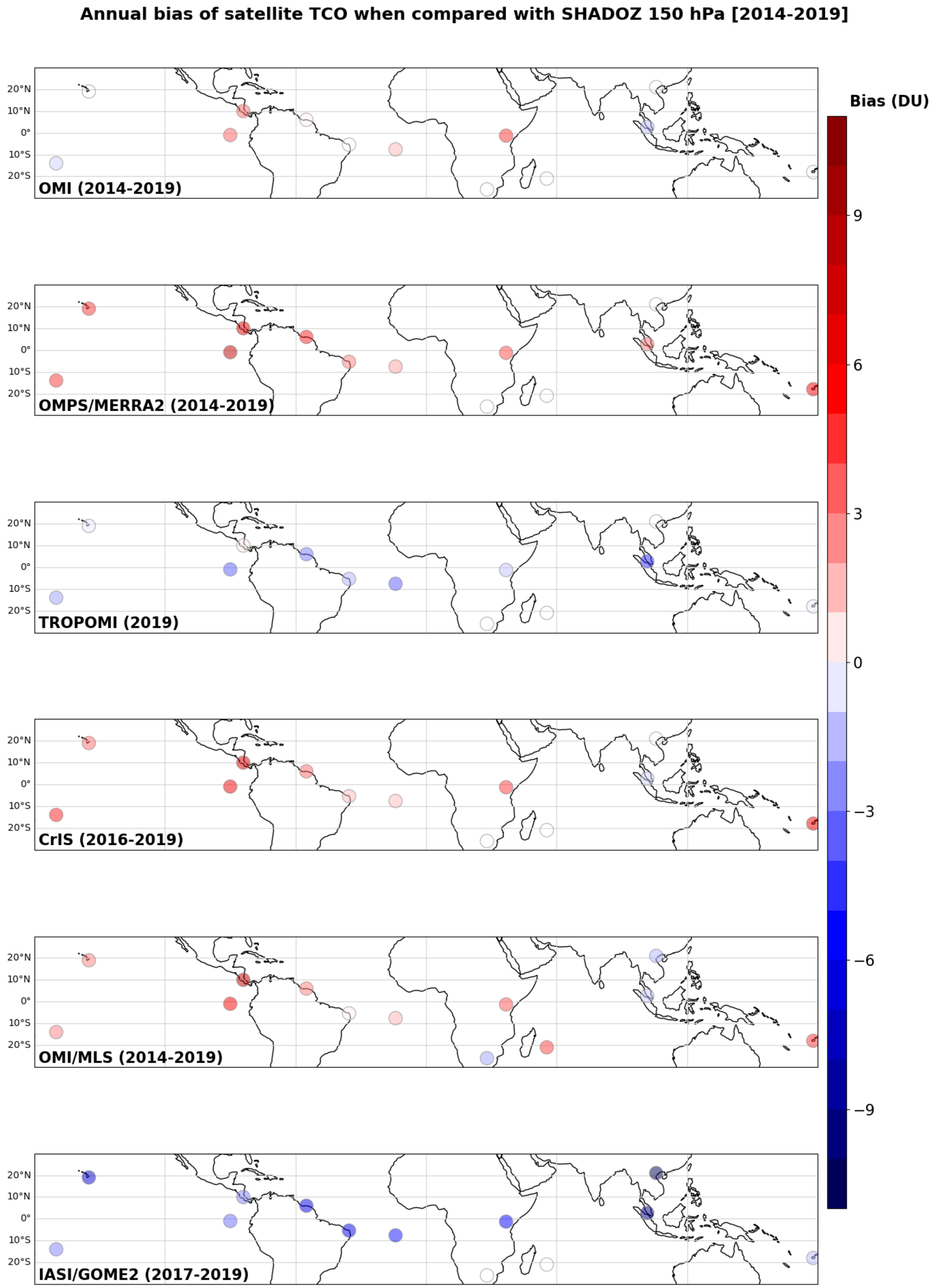


Figure S14. Similar to Figure 5 but the in situ TTCO is derived from integrating the column up to 270 hPa instead of 100 hPa.



Annual bias of satellite TCO when compared with SHADOZ 150 hPa [2014-2019]
Bias (DU)
9
6
3
0
−3
−6
−9
20°N
10°N
0°
10°S
20°S
OMI (2014-2019)
OMPS/MERRA2 (2014-2019)
TROPOMI (2019)
CrIS (2016-2019)
OMI/MLS (2014-2019)
IASI/GOME2 (2017-2019)

Figure S15. Same as Figure S13 but against SHADOZ only (TTCO up to 150 hPa).

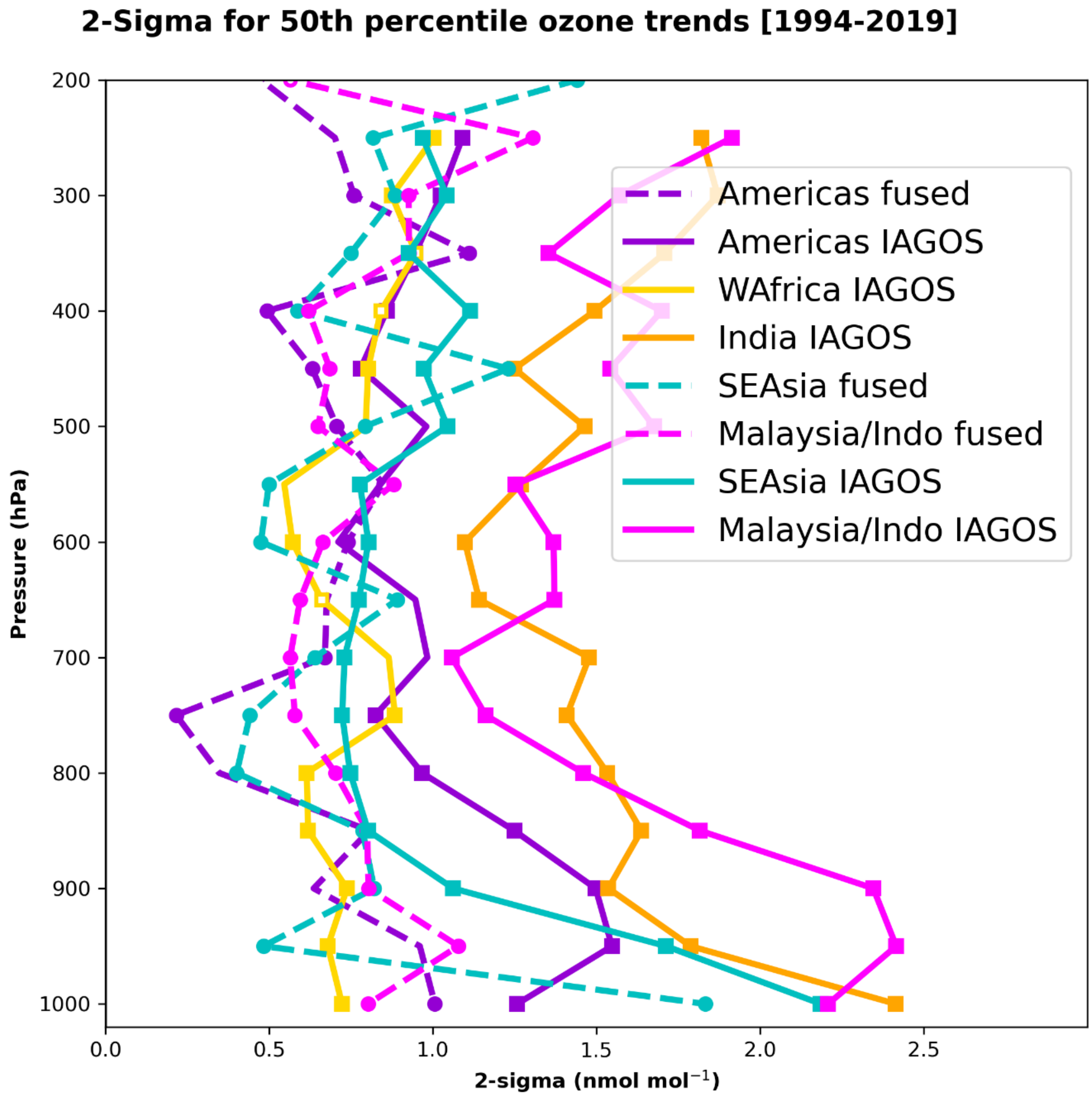


Figure S16. Vertical profiles of the 2-sigma that characterize ozone trends (nmol mol^{-1} which is equivalent to nmol mol^{-1}) between 1994 and 2019 with 50 hPa resolution. The 2-sigma metric is calculated for the 5 IAGOS regions in the tropics:
Americas, Malaysia/Indonesia, Western Africa, India, Southeast Asia. We assessed the 2-sigma associated to IAGOS trends (squares) and the fused trends (IAGOS + SHADOZ, circles) in the 3 regions where SHADOZ data are available.

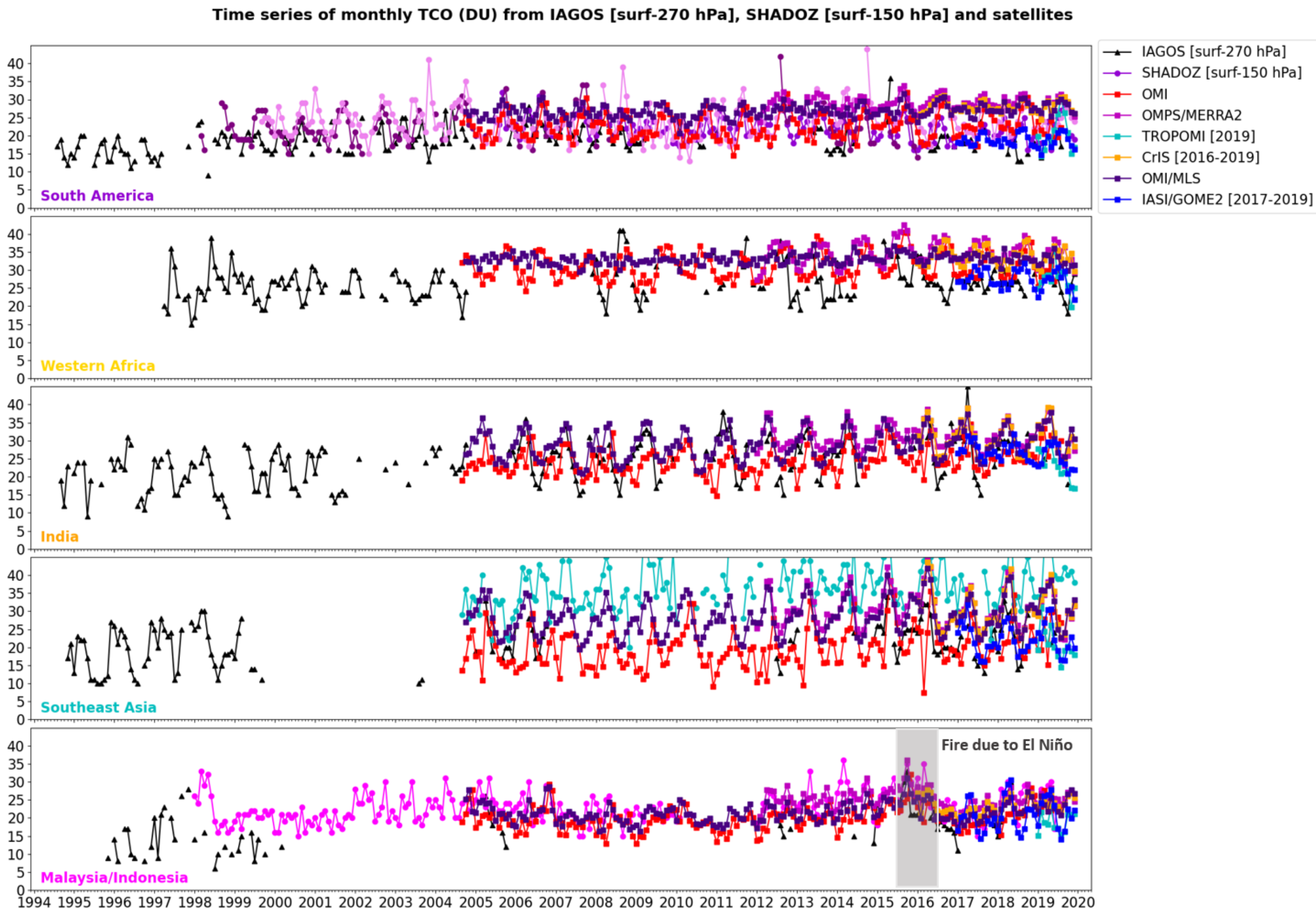


Figure S17. Time series of the monthly tropical tropospheric column ozone (TTCO) from IAGOS ozone profiles (TTCO: surface-270 hPa, black line with triangle markers), from SHADOZ ozone profiles (TTCO: surface-150hPa, colored lines with circle markers), and satellite data (colored line with square markers) extracted above the IAGOS regions. These monthly columns are not used to assess the trends reported in Table 1.

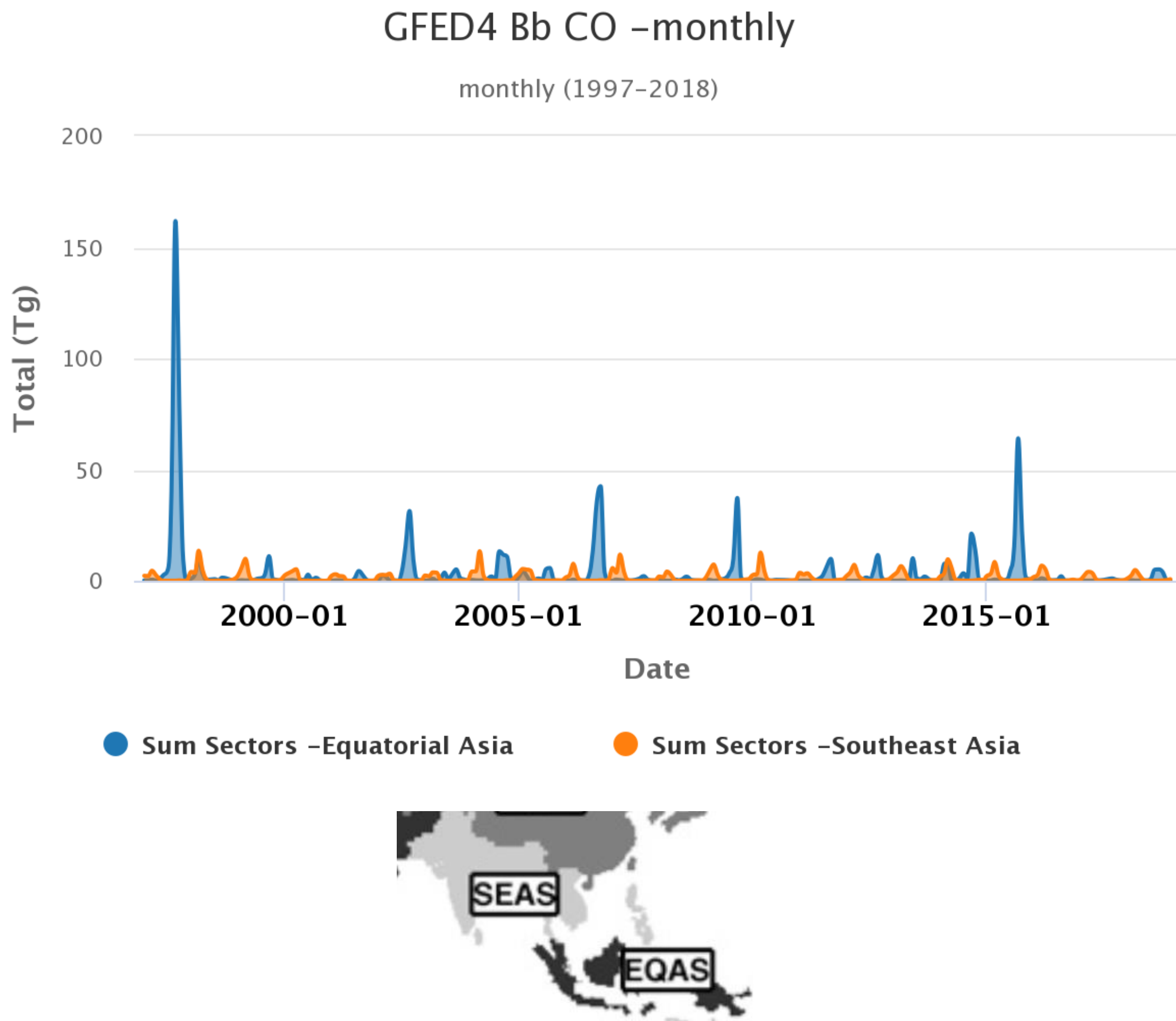
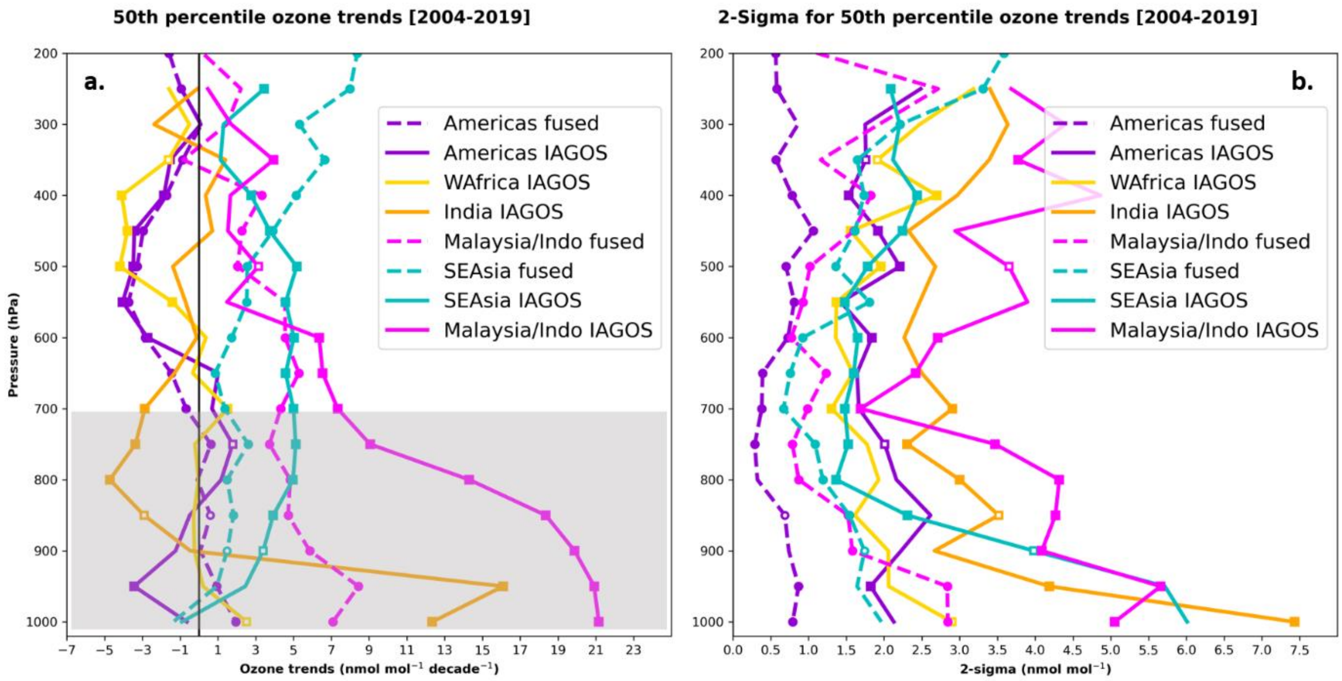


Figure S18. Time series of monthly emissions of CO in Tg due to biomass burning over two GFED source regions: Equatorial Asia (EQAS) and Southeast Asia (SEAS). *Source: ECCAD (https://eccad.aeris-data.fr/)*



Figures S19. Vertical profiles of ozone trends (nmol mol^{-1} decade^{-1}) (panel a.) and the associated uncertainties (2-sigma, panel b.) between 2004 and 2019 with 50 hPa resolution. Trends are calculated for the 5 IAGOS regions in the tropics:
Americas, Western Africa, India, Southeast Asia and Malaysia/Indonesia. SHADOZ data are available for 3 out of the 5 IAGOS regions and fused trends (IAGOS + SHADOZ) were assessed.
Squares (IAGOS trends) or circles (fused trends) indicate trends with *p* values less than 0.05. Open squares or circles indicate trends with *p* values between 0.05 and 0.1. The zero-trend value is indicated with a vertical black bar. The vertical range below 700 hPa is colored in grey to indicate that the fused trends are based on several sites and airports influenced by different local air masses.

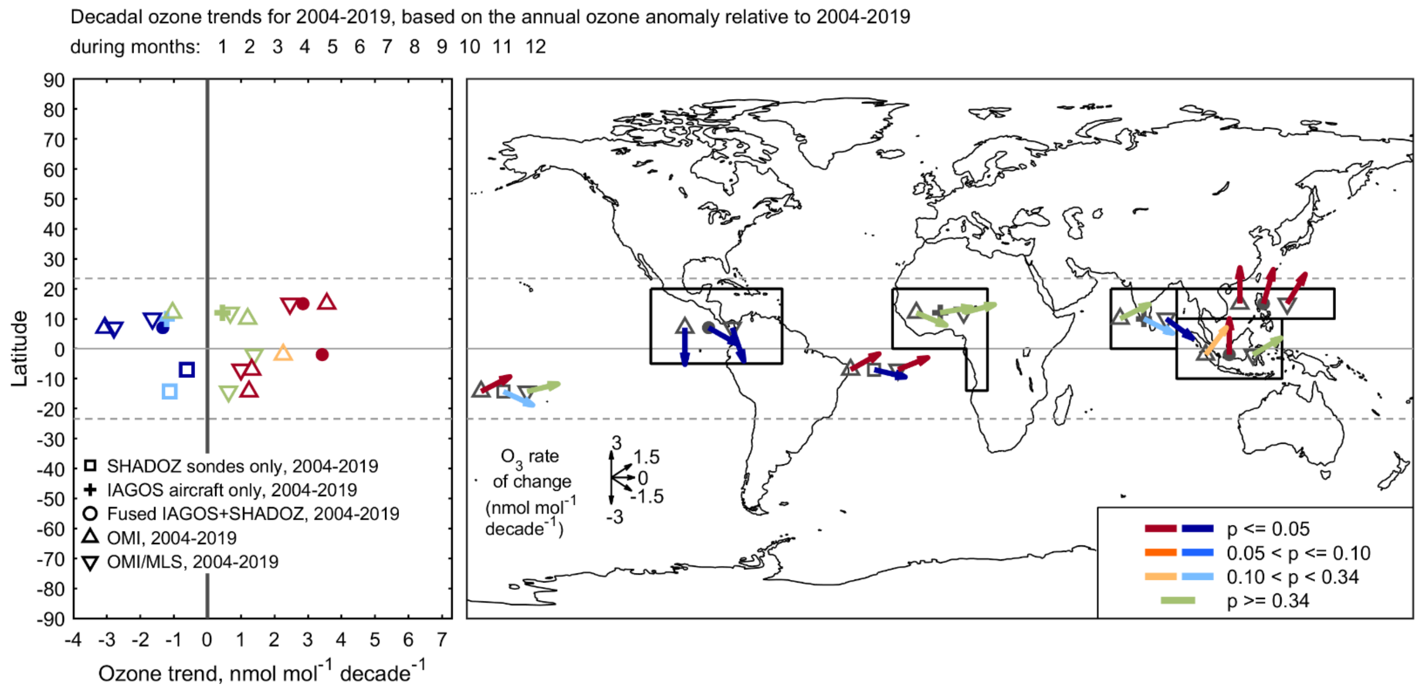


Figure S20. As in Figure 7 but the satellite sample sizes have been greatly reduced so that they only coincide with the specific months and grid-cells sampled by the IAGOS aircraft. Trends of tropical tropospheric column ozone (TTCO) in nmol mol^{-1} decade^{-1} between 2004 and 2019 from IAGOS (crosses), SHADOZ (squares), IAGOS fused with SHADOZ (circles), OMI (triangles up) and OMI/MLS (triangles down) above the five continental IAGOS regions (Americas, Africa, India, Southeast Asia and Malaysia/Indonesia) and two oceanic SHADOZ regions (Samoa and Natal + Ascension Island). The left panel shows the trends of ozone as a function of latitude. The right panel shows the trends of ozone on the map with the black rectangles demarcating the five IAGOS regions. On the map, the longitude of the crosses, circles, triangles and squares are arbitrary and the latitude is the mean latitude of the black rectangles or relative to the SHADOZ sites. The direction of the arrows shows the magnitude of the trends and the colors indicate the *p*-value. The TTCO trends from in situ data are calculated from the monthly TTCO between the surface and 100 hPa, except over India where IAGOS profiles are available between the surface and around 200 hPa. The TTCO trends from OMI and OMI/MLS are calculated from the monthly TTCO defined between the surface and around 102-105 hPa (Figure S1).