S.1 Complementary description of TCCON data

In the Table S.1.1 a reader can find detailed description of the TCCON measurements sites with their respective codes (i.e. acronyms) used in this study.

Table S.1.1 Description of all TCCON sites ('Site' column) with coordinates ('Lat' stands for latitude and 'Lon' stands for longitude), with acronyms ('Code' column) that are used in our paper to denote every station. 'Region' column stands for the region a station belongs to (SH - South Hemisphere, NAM - North America). 'Land cover' stands for the corresponding land type where station is located (see for land cover classification in Table S.2.2)

Site	Lat	Lon	Code	Region	LC
Ascension	-7.92	-14.33	ASC	SH	CRO
Darwin	-12.43	130.89	DRW	SH	HER
Lauder	-45.05	169.68	LDR	SH	TRO
Manaus	-3.21	-60.6	MAN	SH	HWS
Reunion	-20.9	55.49	REU	SH	BDF
Wollongong	-34.41	150.88	WOL	SH	URB
Bialystok	53.23	23.02	BLY	Europe	CRO
Bremen	53.1	8.85	BRE	Europe	URB
Garmisch	47.48	11.06	GAR	Europe	TRO
Izana	28.3	-16.48	IZA	Europe	SHR
Karlsruhe	49.1	8.44	KAR	Europe	BEF
Ny Alesund	78.9	11.9	NYA	Europe	SNO
Orleans	47.97	2.11	ORL	Europe	CRO

Paris	48.85	2.36	PAR	Europe	URB
Sodankyla	67.37	26.63	SOD	Europe	NEF
Zugspitze	47.42	10.98	ZUG	Europe	TRO
Burgos	18.53	120.65	BUR	Asia	OVM
Rikubetsu	43.46	143.77	RIK	Asia	BDF
Saga	33.24	130.29	SAG	Asia	URB
Tsukuba	36.05	140.12	TSU	Asia	OVM
Anmyeondo	36.54	126.33	AMY	NAM	OVM
East Trout	54.35	-104.99	EAT	NAM	NEF
Edwards	34.96	-117.88	EDW	NAM	SAN
Eureka	80.05	-86.42	ERK	NAM	SPV
Four Corners	36.8	-108.48	FCO	NAM	SPV
Indianapolis	39.86	-86	IND	NAM	CRO
JPL	34.2	-118.18	JPL	NAM	URB
Lamont	36.5	-97.49	LAM	NAM	CRO
Park Falls	45.94	-90.27	PKF	NAM	NEF
Pasadena	34.14	-118.13	PSD	NAM	URB

10

Since numerous TCCON stations are located in the close vicinity to each other in Europe and North America, a reader can misjudge about a location of a station in these regions. To avoid such misjudging, we also present Fig. S.1.1 below where JPL and EDW stations can be distinguished on the map (left panel) as well as GAR and ZUG stations can be distinguished as well (right panel).

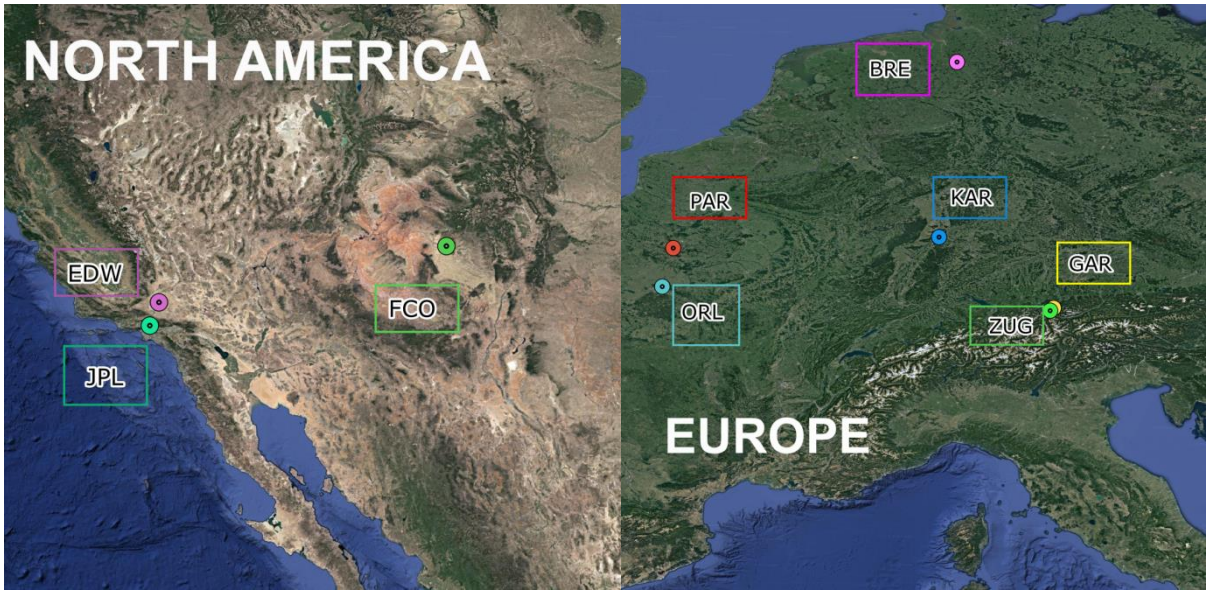

Figure S.1.1 Detailed map of locations of TCCON sites. Left panel - North America (EDW, JPL and FCO stations). Right panel - Europe (ZUG, GAR, PAR, ORL and BRE stations). The satellite map provided by Google maps is embedded using QuickMapServices plugin

Here, we provide the detailed description of TCCON-based AGR data (used for plotting Fig. 2, panel a). Table S.1.2 shows median AGR estimate (AGR median), standard deviation of AGR (AGR σ) and number of MGR values (MGR count) used for AGR calculation. We note that the number of TCCON observations have been constantly increasing since the onset of the network and a reader may question a dramatic reduction in total number of MGR by 2019. There are just 6 MGRs available since when we downloaded TCCON datasets, the latest available month was April 2019. This limitation constrained data abundance in 2019 and should be alleviated when all processed observations are uploaded in the open access. Due to sensitivity of TCCON observations to cloudy conditions, the abundance of monthly-averaged CO_2 estimates may also substantially vary from one site to another. To tackle this deficiency, we calculate the errorbars based on uncertainties taken using monthly-based weights. More specifically, we calculate the weight as the number of the months available for each TCCON site within 1 year. Then, the uncertainties of TCCON-based AGR are computed as the mean value of station-based standard deviation divided by the station-based weight (see TCCON errorbars on Fig. 2 shown by dashed gray lines). Such approach is rather reflecting the completeness of seasonal signal in one AGR value than reproducing complete error propagation that could stem from sub-monthly variability of TCCON observations. A reader can find the complete AGR statistics in the Table S.1.2 below.

Table S.1.2 Statistics for AGR calculation (used for plotting Fig. 2, panel A). Columns denote median AGR estimate (AGR median), standard deviation of AGR (AGR σ) and number of MGR values (MGR count) used for AGR calculation. Years 2004 and 2005 are missing due to dearth of observations available for MGR calculation.

AGR median	AGR σ	MGR count	Year
1.99	2.80	2	2006
2.40	2.99	2	2007
3.35	5.02	3	2008
1.71	2.40	5	2009
2.48	1.47	8	2010
2.00	1.24	13	2011
2.03	0.95	16	2012
2.79	0.87	19	2013
1.99	1.06	19	2014
2.08	1.20	19	2015
3.29	0.96	21	2016
2.59	0.87	23	2017
2.25	0.95	21	2018
2.23	0.79	6	2019

Table S.1.3 presents detailed statistics of the AGR sensitivity analysis (that is presented in the manuscript in Fig. 2).

Table S.1.3 Detailed statistics for threshold-dependent AGR analysis. GR stands for median AGR value calculated using TCCON data at global scales (GR-i format where i is the threshold for sub-monthly selection of MGR data). Similarly, CNT stands for the number of available AGR values based on threshold (i) of MGR calculation. This table supplements Fig. 2 of the manuscript.

AGR	AGR-2	AGR-3	AGR-5	AGR-10	AGR-15	AGR-20	AGR-30
2004							
2005							

2006	1.99	1.99	1.99	2.63	2.76	2.85	
2007	2.4	2.4	2.4	2.43	2.66	2.46	2.03
2008	3.35	3.35	3.35	3.35	3.35	2.53	3.54
2009	1.71	1.71	1.71	1.75	1.98	1.19	1.53
2010	2.48	2.48	2.48	2.42	2.28	2.61	2.69
2011	2	2	2	2.06	1.77	1.85	1.95
2012	1.98	2.08	1.99	2.04	1.89	2.15	1.58
2013	2.84	2.83	2.83	2.8	2.79	2.79	2.96
2014	1.99	1.99	2.01	1.88	1.95	1.92	2.01
2015	2.08	2.08	2.08	2.17	2.22	2.28	2.08
2016	3.29	3.29	3.28	3.27	3.29	3.33	3.44
2017	2.67	2.59	2.64	2.34	1.99	2.05	2.91
2018	2.18	2.24	2.27	2.23	2.13	1.95	1.83
2019	2.23	2.23	2.23	2.28	2.23	2.18	

Count	CNT-2	CNT-3	CNT-5	CNT-10	CNT-15	CNT-20	CNT-30
2004	0	0	0	0	0	0	0
2005	0	0	0	0	0	0	0
2006	2	2	2	2	2	2	0
2007	2	2	2	2	2	2	2
2008	3	3	3	3	3	3	1
2009	5	5	5	5	5	4	2
2010	8	8	8	8	7	7	5
2011	13	13	13	12	12	10	5
2012	16	15	15	14	12	12	8
2013	18	18	16	16	14	13	6
2014	19	19	18	17	15	15	10
2015	19	19	19	18	17	14	9

2016	21	21	21	19	16	15	10
2017	23	23	22	18	15	14	7
2018	20	20	20	17	14	13	6
2019	6	6	6	6	6	5	0

Figure S.1.2 present scatterplots of agreement between potential error spread across AGRs (driven by MGR abundance input) at annual scales with ENSO strength (panel a) and with total number of MGR (panel b)

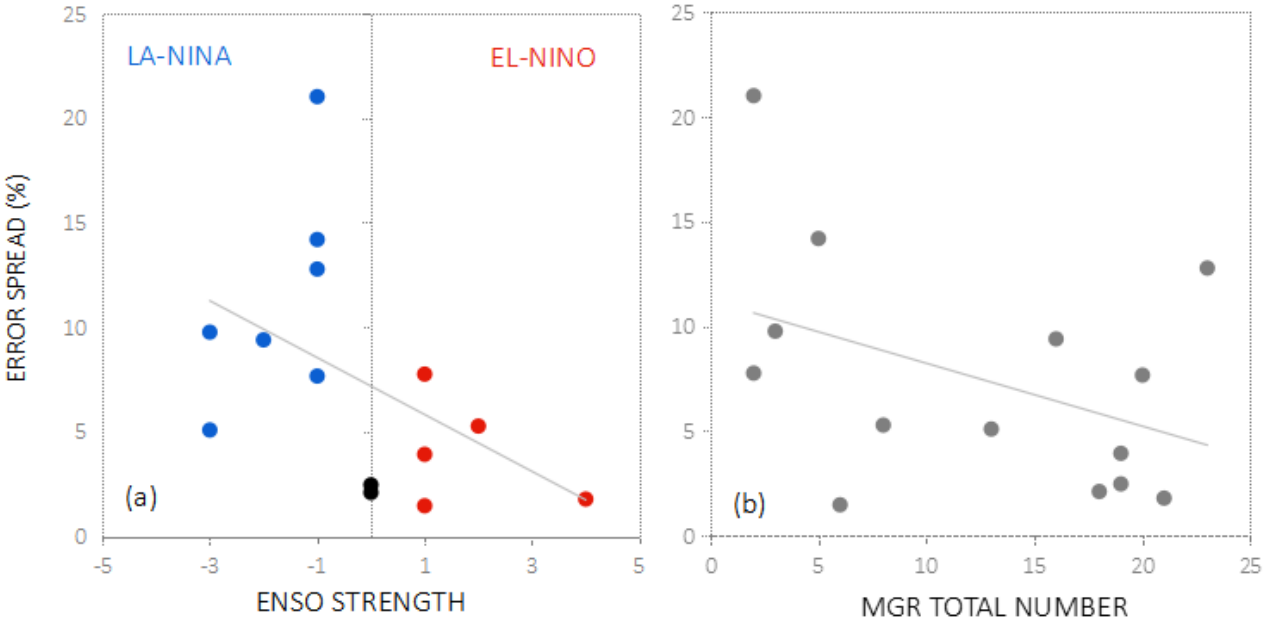

Figure S.1.2. Agreement between potential error spread across AGRs (driven by MGR abundance input) at annual scales with ENSO strength (panel a) and with total number of MGR (panel b). El-Nino events are shown by red and La-Nina events are shown by blue on the panel a.

S.2 Complementary description of land cover data

Here we shortly provide the details about land cover product that remained uncovered from the land cover description given in the main body of the manuscript (Sect. 2.1.5 titled 'Ancillary Datasets'). We present all land types in Table S.2.1 with the respective code (used for this study) and the number (assigned to the land type by the data provider). We note that most land

types are self-explanatory and require no further explication. Since our study deals with 3x2 degree resolution of the models, we have to optimize pixel-based raster format of land cover data. To this end, we simplify the land type classification by assigning 1 certain MODIS-based land type to each grid cell of 3x2 degree. This step is performed by attributing the predominant land type pixel to the sampling point of CO_2 model. In this way, every single grid is assigned with the label designating "land cover" according to MODIS-based classification (Table S.2.1). Therefore, any mixed types of land cover are not considered. This is an obvious but reasonable simplification given the comparably coarse land cover information used for determining land fluxes in both analyzed CO2 models. The geographical distribution of land types can be found in the Fig. S.2.1 below.

Table S.2.1 Description of all land types for land cover product. Column 'Number' stands for classification values originally provided in the data source. 'Code' is acronym used for denoting land type in our study.

Number	Land Type	**Code**
1	Broadleaf Evergreen Forest	**BEF**
2	Broadleaf Deciduous Forest	**BDF**
3	Needleleaf Evergreen Forest	**NEF**
4	Needleleaf Deciduous Forest	**NDF**
5	Mixed Forest	**MFR**
6	Tree Open	**TRO**
7	Shrub	**SHR**
8	Herbaceous	**HER**
9	Herbaceous with Sparse Tree/Shrub	**HWS**
10	Sparse vegetation	**SPV**
11	Cropland	**CRO**
12	Paddy field	**PDF**
13	Cropland / Other Vegetation Mosaic	**OVM**
14	Mangrove	**MGR**
15	Wetland	**WET**

16	Bare area,consolidated (gravel,rock)	**BAR**
17	Bare area,unconsolidated (sand)	**SAN**
18	Urban	**URB**
19	Snow / Ice	**SNO**
20	Water bodies	**WAT**

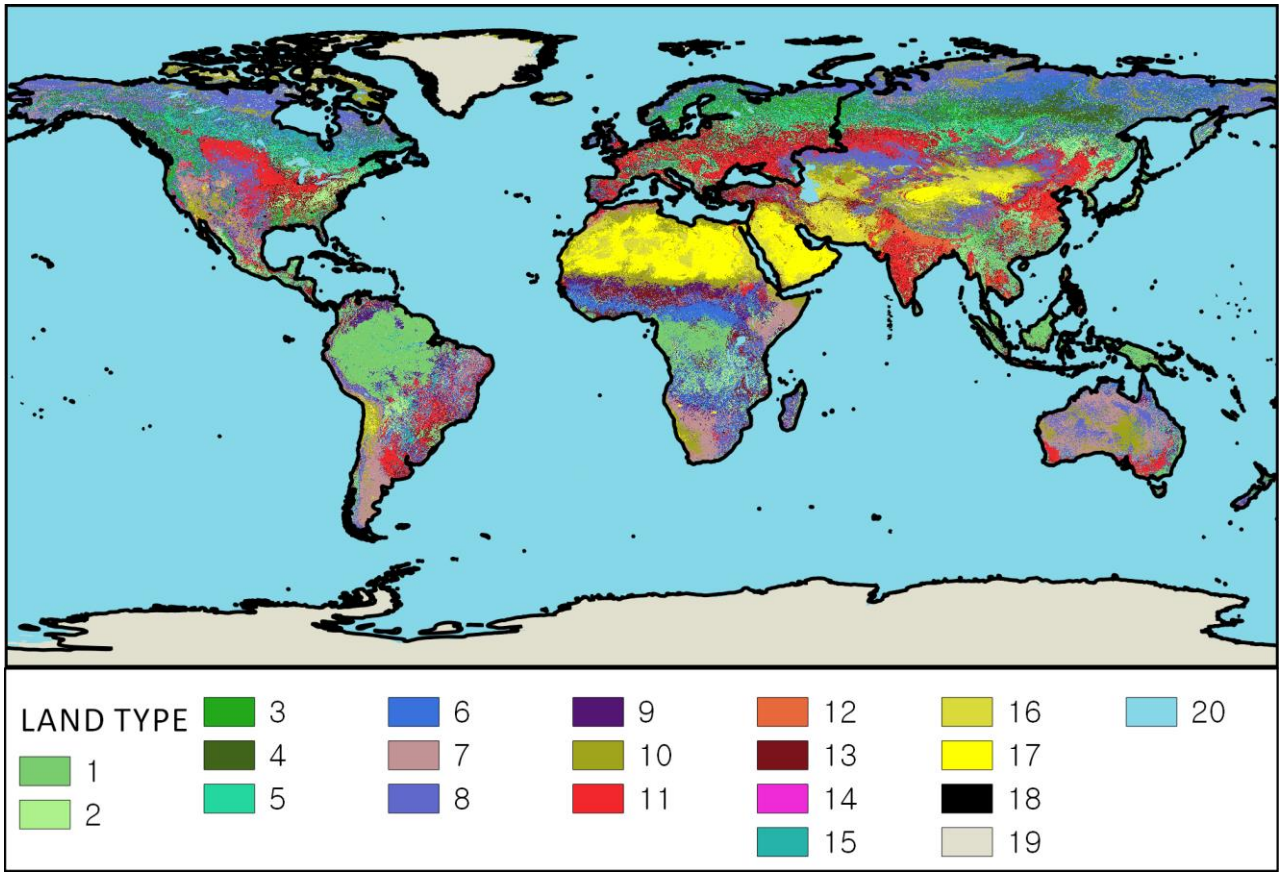

Figure S.2.1 Land cover types. Each number denotes the land type according to the data provider (see Table S.2.1). Note that contours of the oceans are shown by black and not represent urban areas (most urban areas are not seen from the scale used for this Figure).

This additional plot below (Fig. S.2.2) represents the regional classification that was used in the main body of the manuscript for making Table 2.

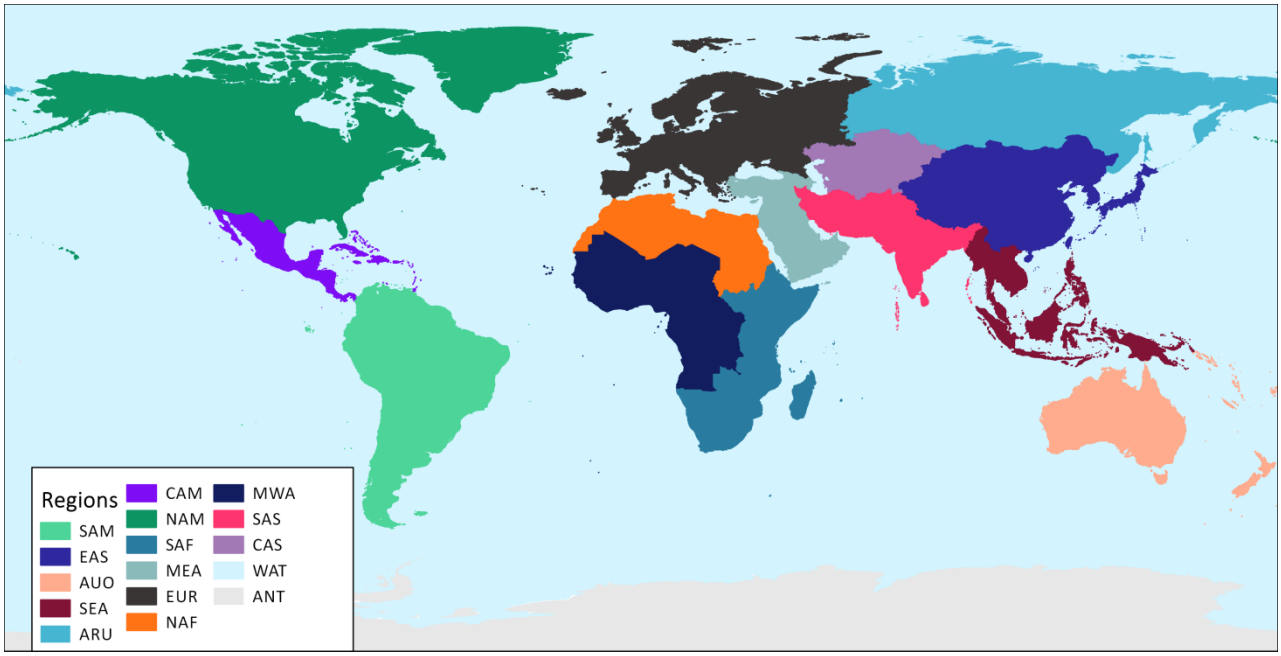

Figure S.2.2 Regional classification of geographic zones. SAM - South America, EAS - East Asia, AUO - Australia-Oceania, SEA - South-East Asia, ARU - Asian part of Russia, CAM - Central America, NAM - North America, SAF - South Africa, MEA - Middle East, EUR - Europe, NAF - North Africa, MWA - Middle-Western Africa, SAS - South Asia, CAS - Central Asia, WAT - Oceans, ANT - Antarctica.

S.3 Calculation of pressure-weighted XCO_2 using CarbonTracker simulations

Since CT provides CO2 simulations at a wide range of vertical levels (1-25). These vertical levels correspond to a range of atmospheric pressure levels that depend on the time and location as well. Approximate heights of the mid-levels range from 34.5 meters (1 level) to 80 000 meters (25 level). In order to ensure that integrated column-averaged CO_2 from CAMS (that is provided by data producer) can be compared with the integrated column-averaged CO_2 from CT (calculated by us), we should perform pressure-weighting procedure for CT-based estimate. To this end, we address to Equation 4 from Zhao et al., (2019). At first, we need to calculate the pressure-dependent multiplier (Δp) from Eq. (S1). To this end, for each available height level (in our case we take the range of 1-25 levels) we need to calculate calculate the difference between the pressure (P) of this layer (i) and pressure of a next vertical layer ($i+1$). Then, the resulted difference should be divided by the difference between a pressure on the bottom layer of the height range (P_s) and the pressure of the top layer of the same

height range (P_t). In our case, P_s corresponds to level 1 and P_t to level 25. We note that CT datasets contain spatio-temporal distribution of atmospheric pressure at monthly temporal domain provided for each vertical level of the model.

$$\Delta \mathrm{p(i)} = \frac{P(i) - P(i+1)}{P_s - P_t} \tag{S.1}$$

Then, we calculate pressure-weighted concentration of CO_2 in the atmospheric column between 1 and 25 levels (XCO_2). This step is done by using Eq. (S2) where n stands for total number of vertical layers and G_{CO2} is simulated CO_2 concentration (by CT) at level of i.

$$\mathrm{XCO_2} = \sum_{i=1}^{n} \Delta \mathrm{p(i)} \times \mathrm{G_{CO2}(i)} \tag{S.2}$$

S.4 Additional analysis and plots with results

Figure S.4.1 demonstrates the distribution of MGR across the seasons (from TCCON) and approves seasonally-independent behavior of TCCON-originated MGRs.

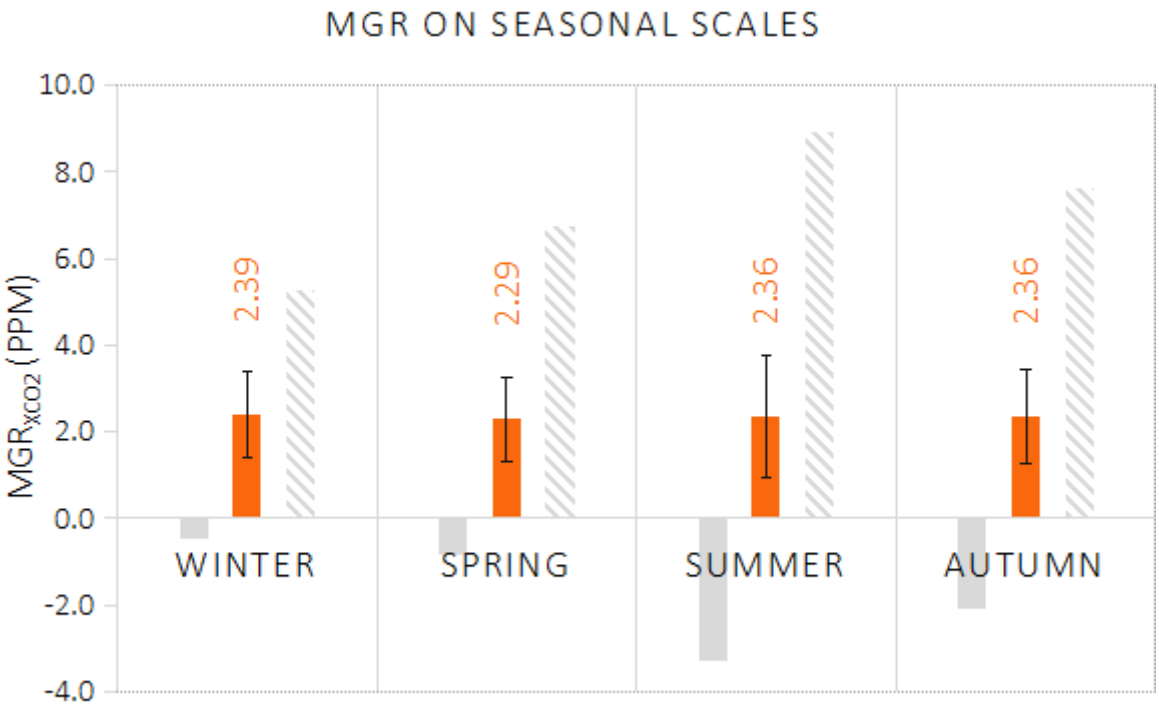

Figure S.4.1 Distribution of MGR depending on the seasons based on TCCON data. Orange color stands for median MGR, gray color stands for minimum (solid bar) and maximum of MGR (lined bar). Errorbars are calculated based on standard deviation of MGR during specific season around the globe in the entire period of available TCCON measurements (2004-2019)

Figure S.4.2 shows the analysis of MGR at the stations where TCCON-to-model agreement was low (Tsukuba, Ascension and Pasadena).

TSU
MGR XCO2 (ppm yr-1)
R2 = 0.11
R2 = 0.12
ASC
MGR XCO2 (ppm yr-1)
R2 = 0.44
R2 = 0.39
PSD
MGR XCO2 (ppm yr-1)
R2 = 0.38
R2 = 0.36

Figure S.4.2 Analysis of MGR at the stations where TCCON-to-model agreement was low (Tsukuba - TSU, Ascension - ASC and Pasadena - PSD).