Supplementary material of

High-resolution inverse modelling of European CH_4 emissions using novel FLEXPART-COSMO TM5 4DVAR inverse modelling system

Peter Bergamaschi et al.

S1 Calculation of total emissions of country regions from gridded emission data

For use in TM5-4DVAR emission inventory data are usually pre-processed to a horizontal resolution of 1° × 1° (which is also the resolution of the zoom regions in TM5-4DVAR). In order to extract total emissions of country regions, two different approaches can be applied. In the first approach the gridded emissions are directly sampled with a corresponding country mask:

$$E_{\mathrm{tot},k} = \sum_{i,j} e_{ij} mask_{i.j.k} \qquad \text{(S1)}$$

where e_{ij} are the emissions of the grid cell with longitude index i and latitude index j and $mask_{i.j.k}$ is the mask for country k, providing the land fraction of the corresponding 1° × 1° grid cell attributed to country k (with $0 \leq mask_{i.j.k} \leq 1$). In case of coastal grid cells, however, this approach leads to a loss of emissions attributed to the country, if emissions are mainly over land. In contrast, the second approach attributes emissions to the land surface only (assuming that emissions over the sea can be neglected) by normalizing the emissions to the land fraction, $land_{i.j}$ ($0 \leq land_{i.j} \leq 1$) of the corresponding 1° × 1° grid cell:

$$E_{\mathrm{tot},k} = \sum_{i,j} e_{ij} \frac{mask_{i.j.k}}{land_{i.j}} \;\; \forall\, i,j \; with \; land_{i.j} \neq 0 \qquad \text{(S2)}$$

For the TM5-4DVAR CH_4 inversions discussed in this paper, generally Eq. (S2) is applied to compute the total emissions of the country regions.

Owing to the much higher spatial resolution of the COSMO-7 grid, a different approach is applied in FLEXVAR to extract the total emissions of country regions from the gridded data at horizontal resolution of 7 km × 7 km. Country masks have been generated using the "Natural Earth dataset" (https://www.naturalearthdata.com/), attributing each 7 km × 7 km COSMO-7 grid cell to a certain country (or sea). Total emissions are then computed by application of Eq. (S1), however in case of FLEXVAR the $mask_{i.j.k}$ of country k can have only values of 0 or 1, i.e., neglecting any sub-grid scale variability within the individual 7 km × 7 km grid cells. If the underlying inventories of the gridded data sets have high resolution (close to the 7 km × 7 km) as in the case of EDGARv6.0 or TNO-VERIFYv3.0, the described sampling approach should provide reasonable estimates of the total emissions of the country regions. However, this is not the case for the GCP-CH_4 inventory (emission data set E3) provided at resolution of 1° × 1°. In this case, the FLEXVAR sampling leads to a loss of emissions at the coastal grid cells, similar as described above for TM5-4DVAR when applying Eq. (S1). In fact, sampling of the E3 inventory in TM5-4DVAR using Eq. (S1) yields very similar total emission as the sampling of this inventory (pre-processed to the COSMO-7 resolution of 7 km × 7 km as described in section 3.2) in FLEXVAR, with differences in calculated total emissions of <1% for Germany, France and BENELUX. Therefore, we calculate correction factors ("grid scaling factors") to be applied for the total emissions extracted in FLEXVAR from the E3 inventory, using the ratio of total emissions computed for TM5-4DVAR applying Eq. (S2) and the total emissions computed applying Eq. (S1). The calculated correction factors are compiled in Table S2. These correction factors are applied both for the prior and posterior emissions of those FLEXVAR (and FLExKF) inversions which use E3 as prior. We note that also emission data sets E1 and E2 use the natural emissions from the GCP-CH_4 inventory (at horizontal resolution of 1° × 1°). In this case, however, no corrections are applied given the overall only very small contribution of the natural sources.

Since the COSMO-7 domain does not cover the upper northern part of the UK, a correction factor has been determined to estimate the total emissions of the country region "UK+Ireland" from the total emissions extracted from the corresponding grid cells within the COSMO-7 domain. This "domain scaling factor" has been calculated as ratio of total emissions of "UK+Ireland" computed for TM5-4DVAR (covering the whole area of UK+Ireland) for the emission data set E3 and the total emissions for that region computed for FLEXVAR. The domain scaling factor derived in this way is 1.057 and is applied to the prior and posterior emissions of the total emissions of UK+Ireland extracted from the gridded FLEXVAR data. In case of FLExKF this value is slightly higher (1.074), since for technical reasons FLExKF uses a slightly smaller grid (with 384 (longitude) × 336 (latitude) grid cells) compared to FLEXVAR (393 (longitude) × 338 (latitude)) and therefore covers a slightly smaller fraction of the UK.

S2 Estimates of uncertainties of total emissions reported to UNFCCC

In addition to the CH_4 emissions reported to UNFCCC, the EU countries (including UK) estimate the uncertainties of emissions for the major source categories (compiled in the national inventory reports). However, they do not report estimates of the uncertainties of total CH_4 emissions. Therefore, we apply here the approach used in Bergamaschi et al. [2015] to derive estimates of total uncertainties from the reported uncertainties per category. This approach evaluates the uncertainties of the 6 major emission categories (1) solid fuels, (2) oil and gas, (3) enteric fermentation, (4) manure management, (5) solid waste, and (6) waste water. For the aggregation of uncertainties, we assume that the uncertainties of these major emission categories are uncorrelated, but that uncertainties of sub-categories are fully correlated (as some countries provide uncertainty estimates only for sub-categories, e.g., in case of enteric fermentation separately for different animals). Furthermore, we assume fully correlated uncertainties when aggregating the uncertainties of major emission categories from different countries (for the evaluation of uncertainties of emissions from groups of different countries (e.g., BENELUX)). The estimated uncertainties are compiled in Table S3. The major source categories represent 87-95% of the total emissions for Germany, France, BENELUX, and UK+Ireland. The derived relative uncertainties of the emissions from major source categories are applied to the reported total emissions. In addition to the uncertainty introduced by the described assumptions of correlations, two further limitations should be noted: (1) in several cases different countries provide very different uncertainty estimates for the same category (in particular for fugitive emissions from fossil fuels, for the waste sector and for manure management), which suggests that there are probably large uncertainties of the reported uncertainties per category. (2) Our estimate of the total uncertainty does not include any estimate of the uncertainty related to potentially unaccounted emission sources.

Table S1: Additional global and European monitoring stations used in TM5-4DVAR inversions. "alt" is the surface altitude (m above sea level), "s.h." is the sampling height (m) above ground, "ST" specifies the sampling type ("I": in situ measurements; "D": discrete air sample measurements).

ID	station name	data provider	lat	lon	alt	s. h.	ST
ALT	Alert, Nunavut, Canada	NOAA	82.45	-62.52	185	5	D
ZEP	Zeppelin Observatory	ICOS	78.91	11.89	474	15	I
		NOAA	78.91	11.89	474	5	D
SUM	Summit, Greenland	NOAA	72.60	-38.42	3210	5	D
BRW	Barrow, Alaska, USA	NOAA	71.32	-156.60	11	5	D
PAL	Pallas	ICOS	67.97	24.12	565	12	I
		NOAA	67.97	24.12	565	5	D
SVB	Svartberget	ICOS	64.26	19.77	235	150	I
ICE	Heimay, Vestmannaeyjar, Iceland	NOAA	63.40	-20.29	118	4	D
SMR	SMEAR II-ICOS Hyytiala	ICOS	61.85	24.29	181	125	I
NOR	Norunda	ICOS	60.09	17.48	46	100	I
UTO	ICOS Uto	ICOS	59.78	21.37	8	57	I
CBA	Cold Bay, Alaska, USA	NOAA	55.21	-162.72	21	36	D
SHM	Shemya Island, Alaska, USA	NOAA	52.72	174.10	23	5	D
NWR	Niwot Ridge, Colorado, USA	NOAA	40.05	-105.58	3523	3	D
AZR	Terceira Island, Azores, Portugal	NOAA	38.77	-27.38	19	5	D
WLG	Mt. Waliguan, P. Rep. of China	NOAA	36.29	100.90	3810	5	D
LMP	Lampedusa, Italy	NOAA	35.52	12.62	45	5	D
BMW	Tudor Hill, Bermuda, UK	NOAA	32.26	-64.88	30	21	D
IZO	Tenerife, Canary Islands, Spain	NOAA	28.31	-16.50	2373	5	D
MID	Sand Island, Midway, USA	NOAA	28.21	-177.38	11	4	D
ASK	Assekrem, Algeria	NOAA	23.26	5.63	2710	5	D
MLO	Mauna Loa, Hawai, USA	NOAA	19.53	-155.58	3397	5	D
KUM	Cape Kumukahi, Hawaii, USA	NOAA	19.52	-154.82	3	5	D
GMI	Mariana Islands, Guam	NOAA	13.43	144.78	0	5	D
RPB	Ragged Point, Barbados	NOAA	13.17	-59.43	15	5	D
CHR	Christmas Island, Rep. of Kiribati	NOAA	1.70	-157.17	0	5	D
SEY	Mahe Island, Seychelles	NOAA	-4.68	55.53	2	5	D
ASC	Ascension Island, UK	NOAA	-7.97	-14.40	85	5	D
SMO	Tutuila, American Samoa, USA	NOAA	-14.25	-170.56	42	5	D
EIC	Easter Island, Chile	NOAA	-27.15	-109.45	47	22	D
CGO	Cape Grim, Tasmania, Australia	NOAA	-40.68	144.68	94	70	D
CRZ	Crozet Island, France	NOAA	-46.43	51.85	197	5	D
USH	Ushuaia, Argentina	NOAA	-54.85	-68.31	12	20	D
PSA	Palmer Station, Antarctica, USA	NOAA	-64.92	-64.00	10	5	D
SYO	Syowa Station, Antarctica, Japan	NOAA	-69.00	39.58	14	5	D
SPO	South Pole, Antarctica, USA	NOAA	-89.98	-24.80	2810	5	D

Table S2: Grid scaling factors for different country regions applied to the prior and posterior emissions of the FLEXVAR and FLExKF inversions which use emission data set E3 as prior (as described in section S1).

country region	grid scaling factors
Germany	1.030
France	1.089
BENELUX	1.095
UK+Ireland	1.170

Table S3: Emissions reported to UNFCCC and estimated uncertainties (2 sigma).

		Germany	France	BENELUX	UK+Ireland
solid fuels (1B1)					
emissions 2018	Gg CH_4 yr^{-1}	66.6	0.7	1.8	19.5
relative uncertainty	%	24.1	11.2	53.8	21.5
Oil and natural gas (1B2)					
emissions 2018	Gg CH_4 yr^{-1}	195.3	42.7	40.0	195.9
relative uncertainty	%	16.6	100.5	37.5	20.2
Enteric fermentation (3A)					
emissions 2018	Gg CH_4 yr^{-1}	959.2	1372.8	511.0	1346.0
relative uncertainty	%	16.7	15.8	19.8	14.6
Manure management (3B)					
emissions 2018	Gg CH_4 yr^{-1}	236.6	154.2	202.2	223.4
relative uncertainty	%	16.4	30.4	85.6	11.5
Solid waste (5A)					
emissions 2018	Gg CH_4 yr^{-1}	302.2	507.8	129.4	608.9
relative uncertainty	%	50.0	175.6	29.9	48.4
Waste water(5D)					
emissions 2018	Gg CH_4 yr^{-1}	20.4	92.6	18.7	68.4
relative uncertainty	%	26.4	105.9	56.2	27.1
total					
total major categories	Gg CH_4 yr^{-1}	1780.3	2170.8	903.0	2461.9
total all reported emissions	Gg CH_4 yr^{-1}	2054.4	2275.4	1011.2	2597.7
relative uncertainty	%	12.7	42.6	22.7	14.5

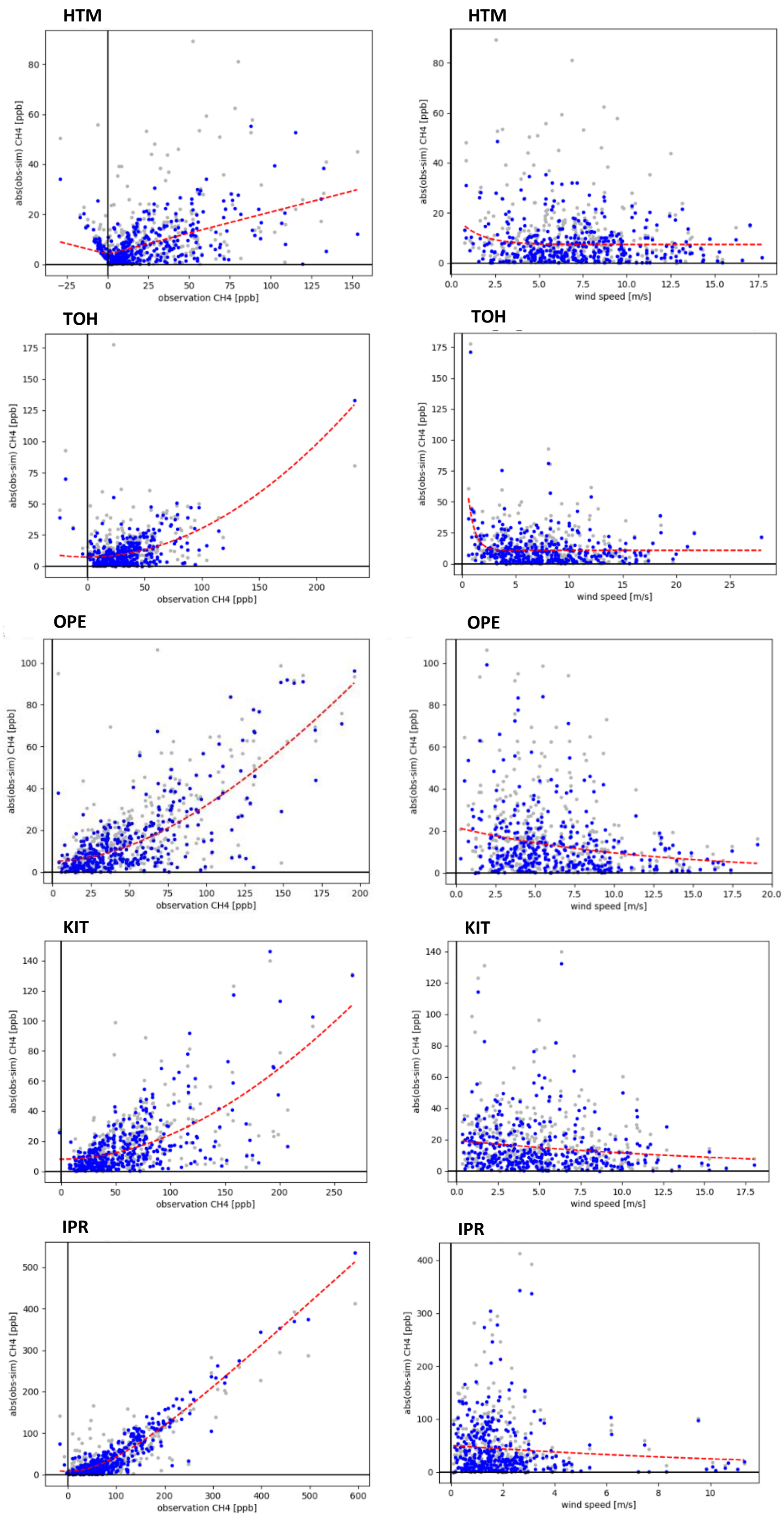

Figure S1: Parameterization of model representation error (described in section 2.2.3) shown for 5 example stations. Left: Parameterized as function of the absolute observed CH_4 enhancement ("OBS"); right: parameterized as function of wind speed ("METEO"). Blue data and red fit functions from second outer loop iteration. Grey data from first outer loop iteration.

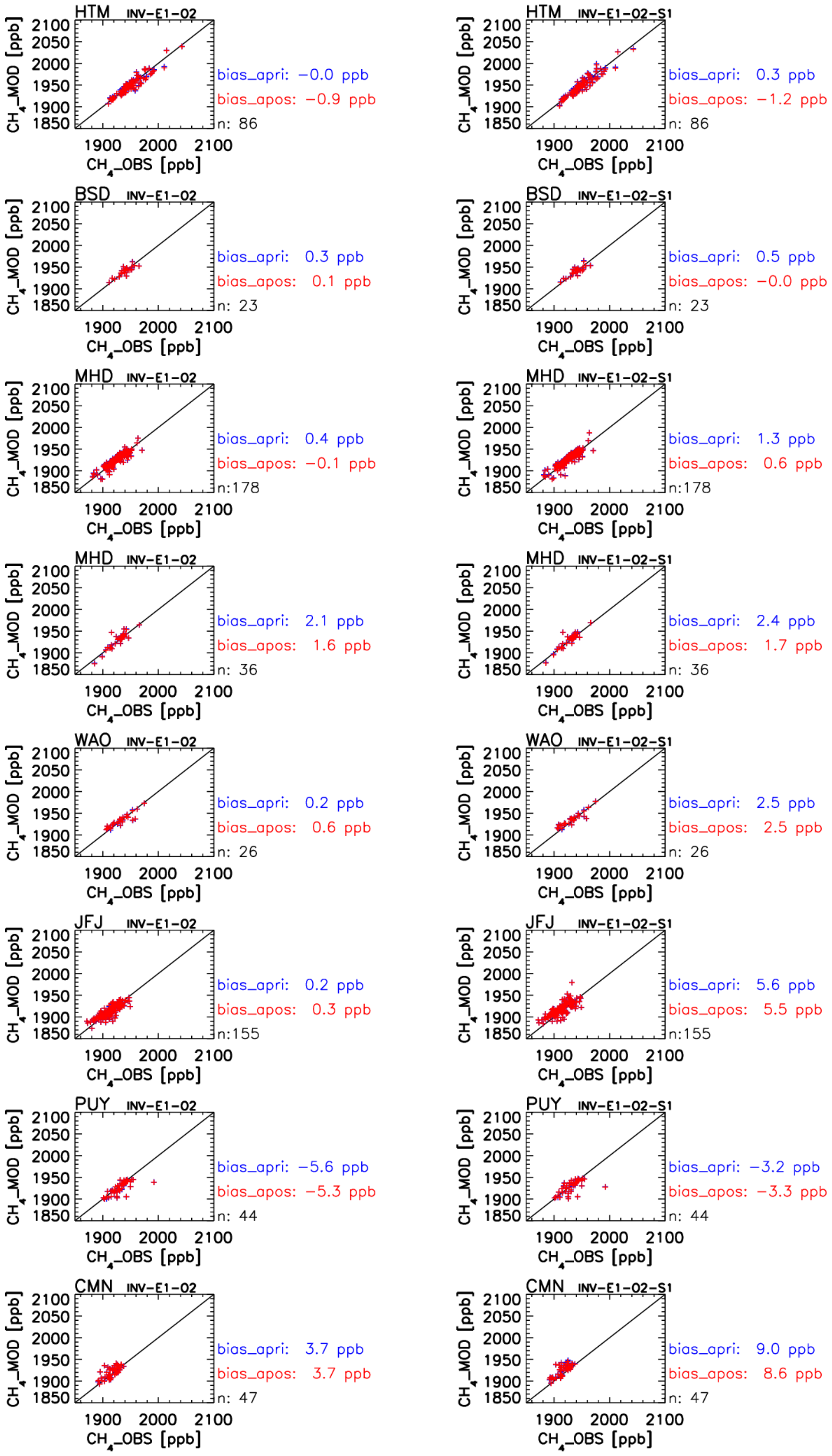

Figure S2: Model simulations vs. observations for 'background conditions' (defined as contribution of European emission < 5 ppb). Only stations, for which number of data with 'background conditions' is greater than 20, are shown. Left column: INV-E1-O2 ("Rödenbeck baselines"); right column: INV-E1-O2-S1 ("particle position baselines").

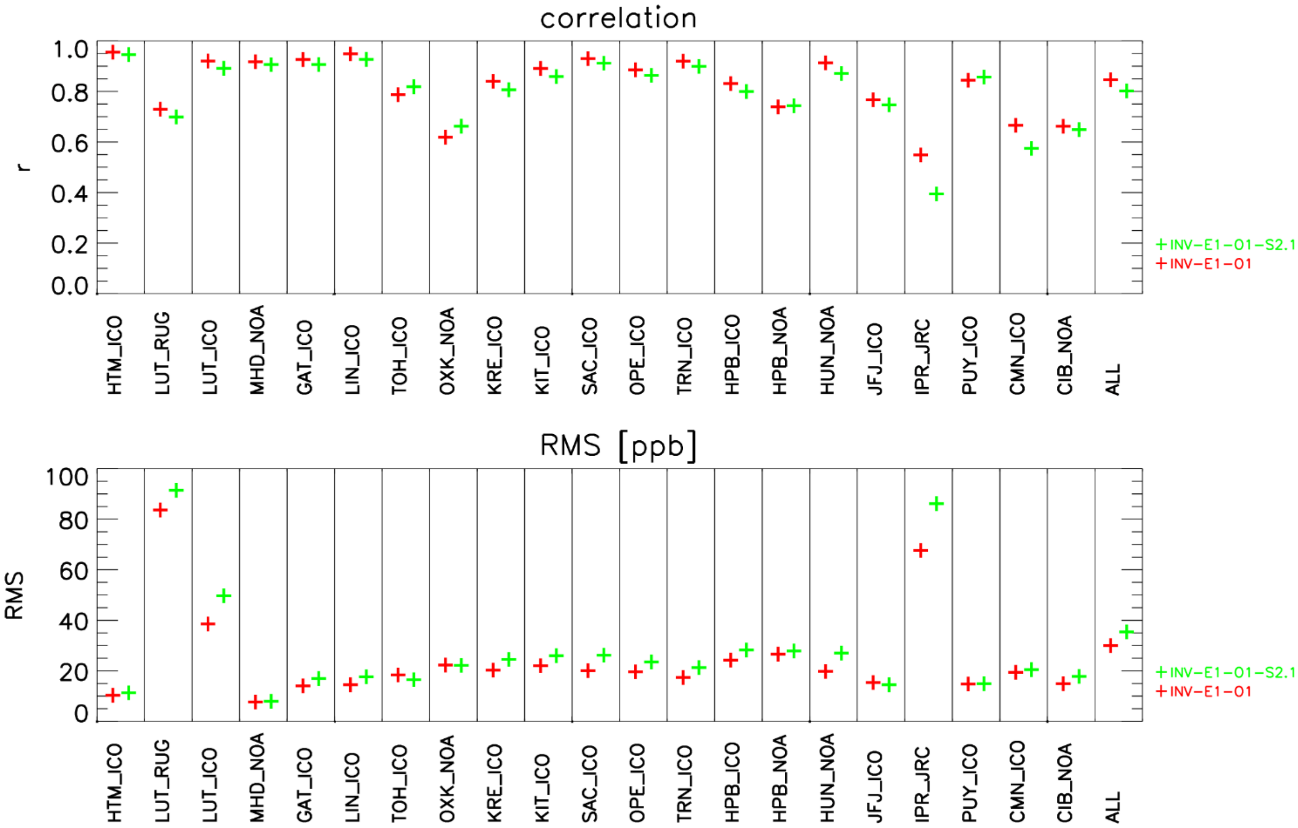

Figure S3: Station statistics for base inversion INV-E1-O1 (with the "METEO" parameterization of model representation error) and sensitivity inversion INV-E1-O1-S2.1 (with "OBS" model representation error). Upper panel: correlation coefficient r; lower panel: root mean square difference between observations and model simulations. The last column ("ALL") shows the average correlation coefficient and root mean square difference, respectively, from all stations.

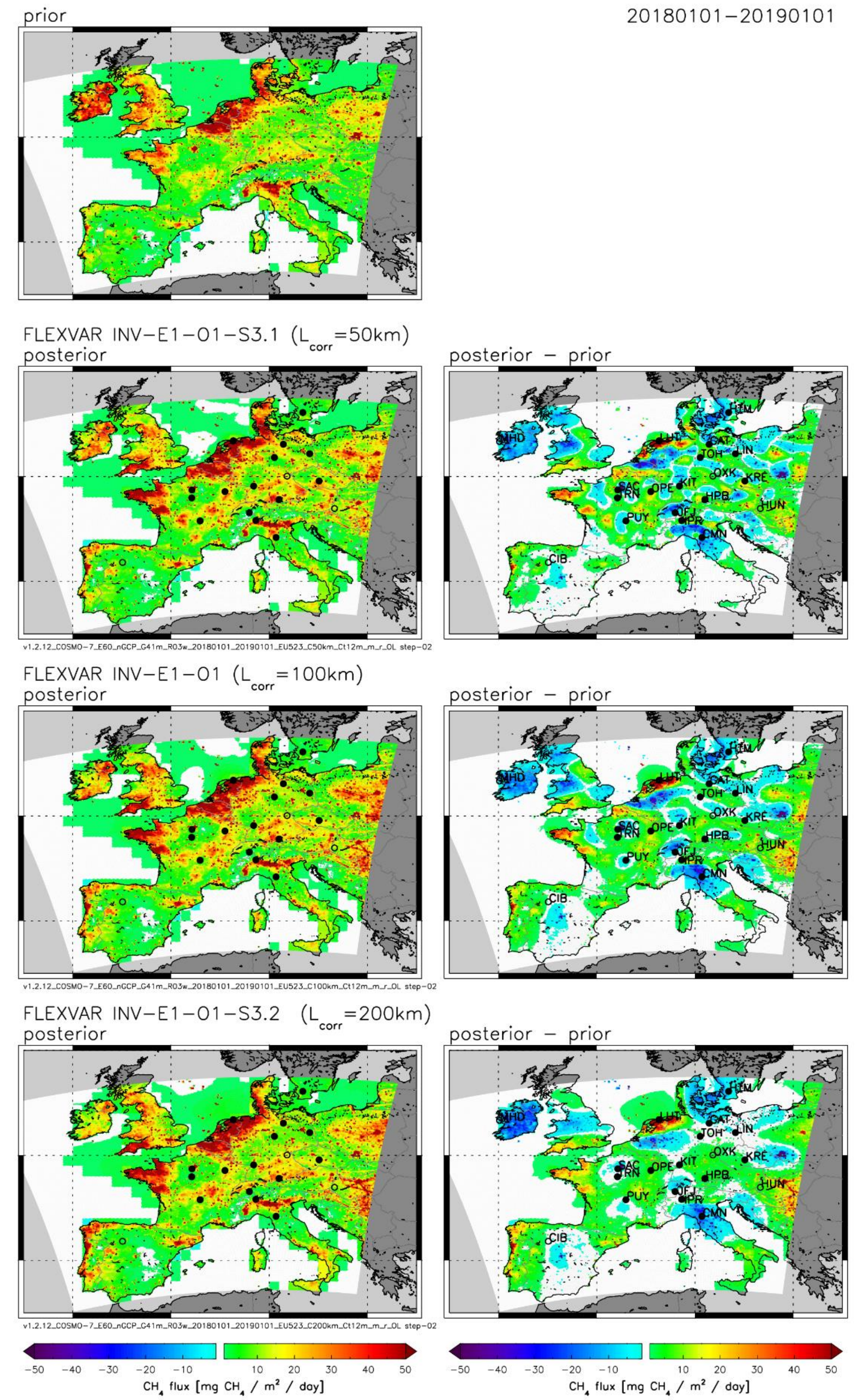

Figure S4: Sensitivity of FLEXVAR inversions to applied spatial correlation length scale of prior emissions.

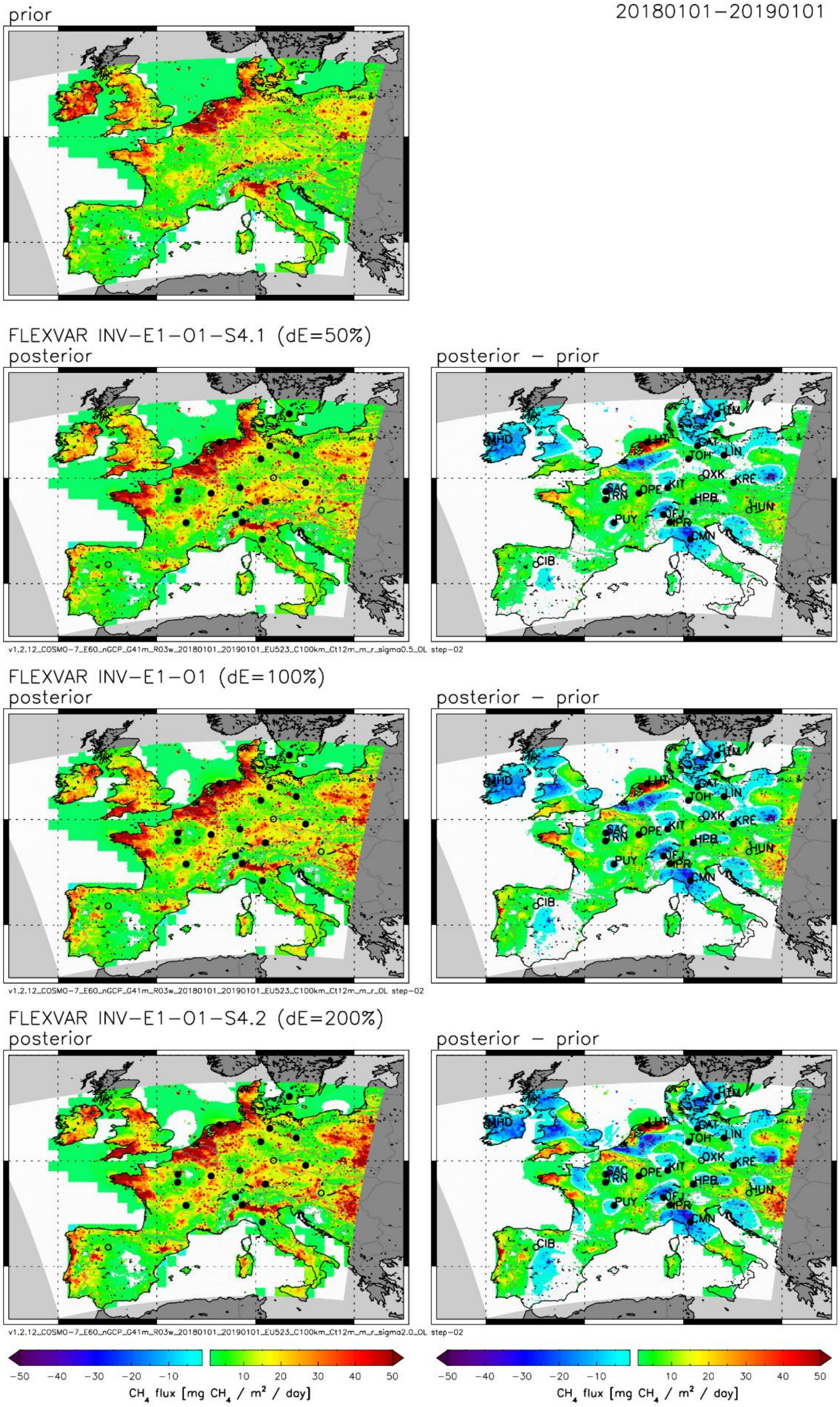

Figure S5: Sensitivity of FLEXVAR inversions to assumed uncertainties of prior emissions per grid cell and month.

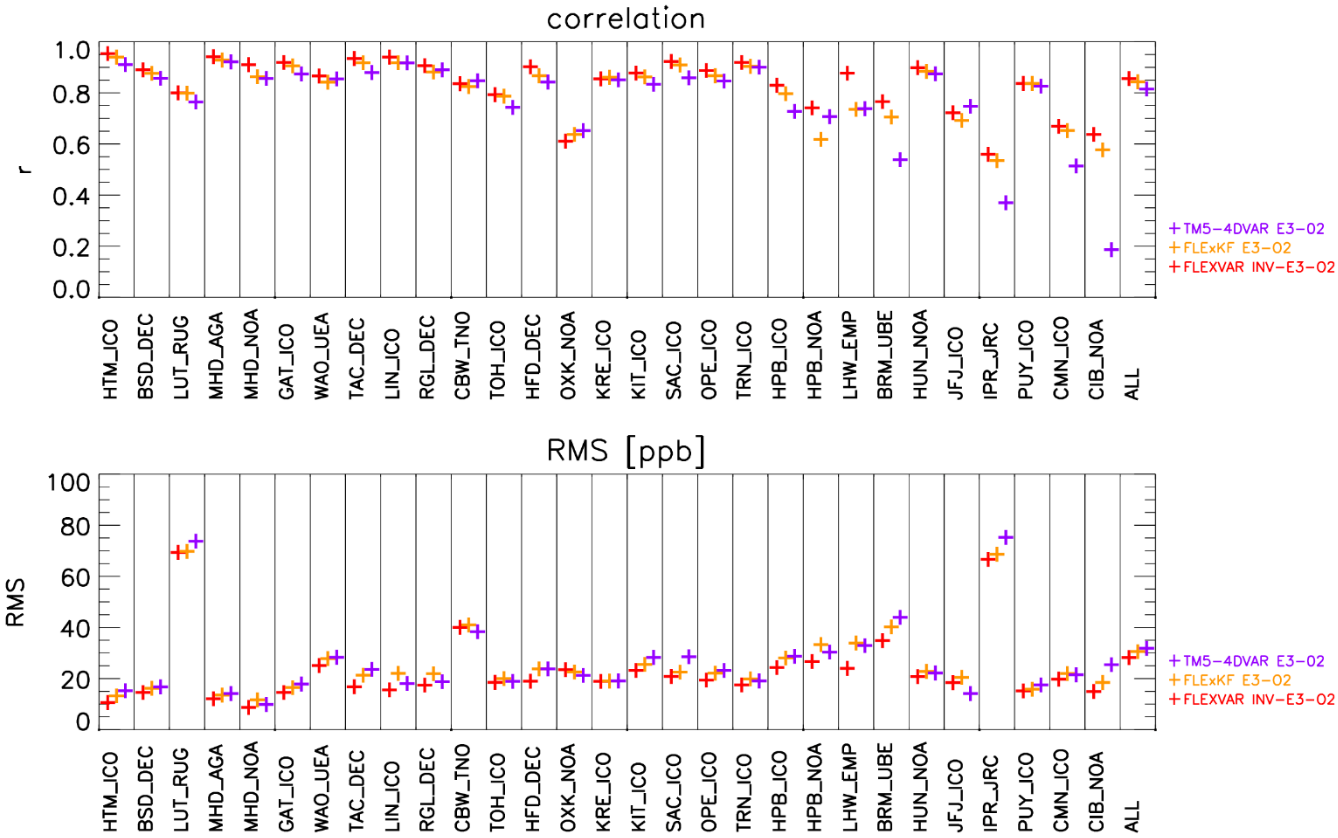

Figure S6: Model comparison: Station statistics for FLEXVAR, FLExKF and TM5-4DVAR for inversions using the same emission data set E3 as prior and observation data set O2. Upper panel: correlation coefficient r; lower panel: root mean square difference between observations and model simulations. The last column ("ALL") shows the average correlation coefficient and root mean square difference, respectively, from all stations.

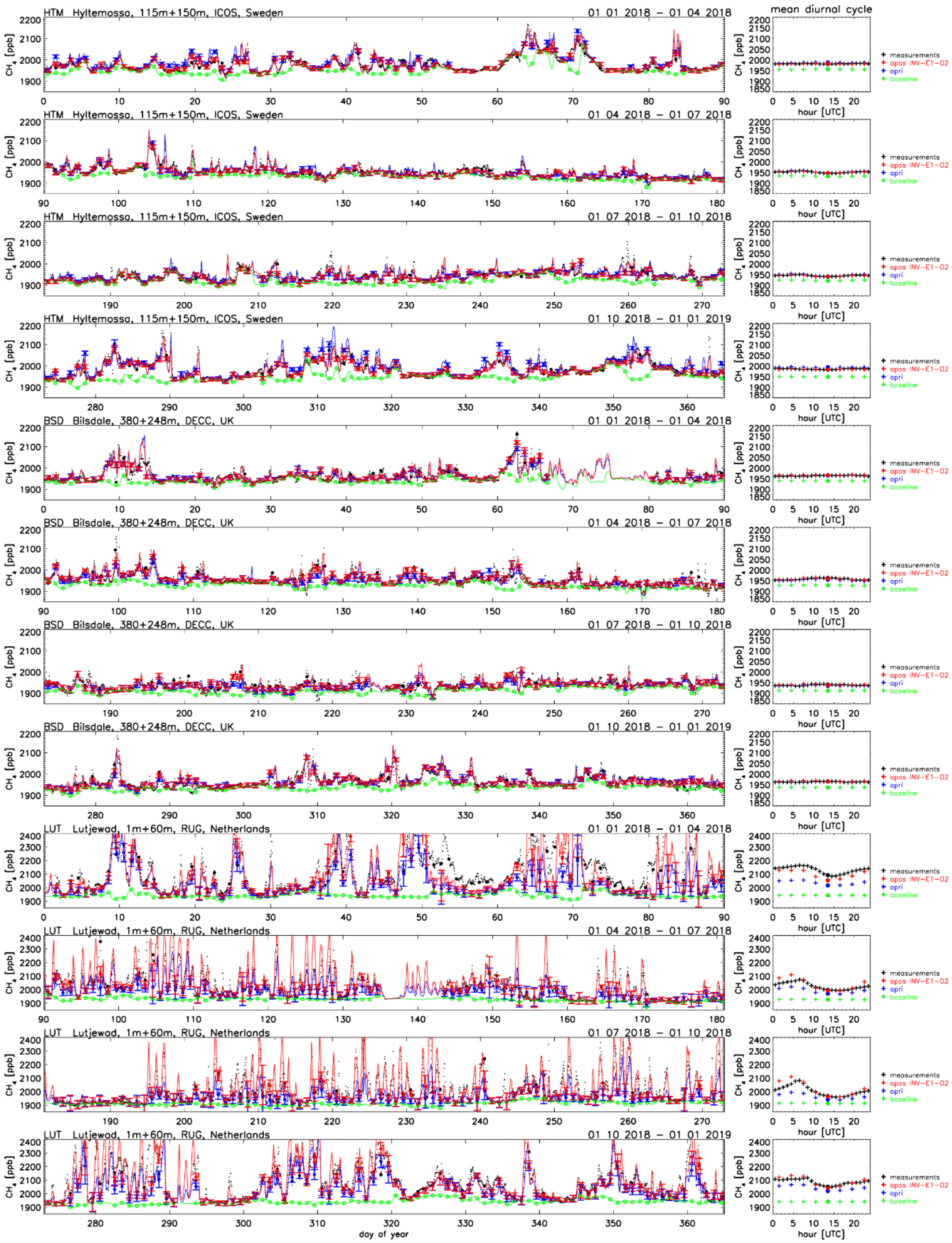

Figure S7: Simulated and observed CH_4 mole fractions at monitoring stations (inversion INV-E1-O2). The main panels (left) show the time series for 3-months periods. Green curves show the baselines, blue the prior simulations, and red the posterior simulations. Small black dots: Hourly-averaged observations. Solid black circles: assimilated observations. Colored circles show the corresponding assimilated values and error bars the applied model representation error. The smaller figures on the right side show the average diurnal cycle for each 3-months period.

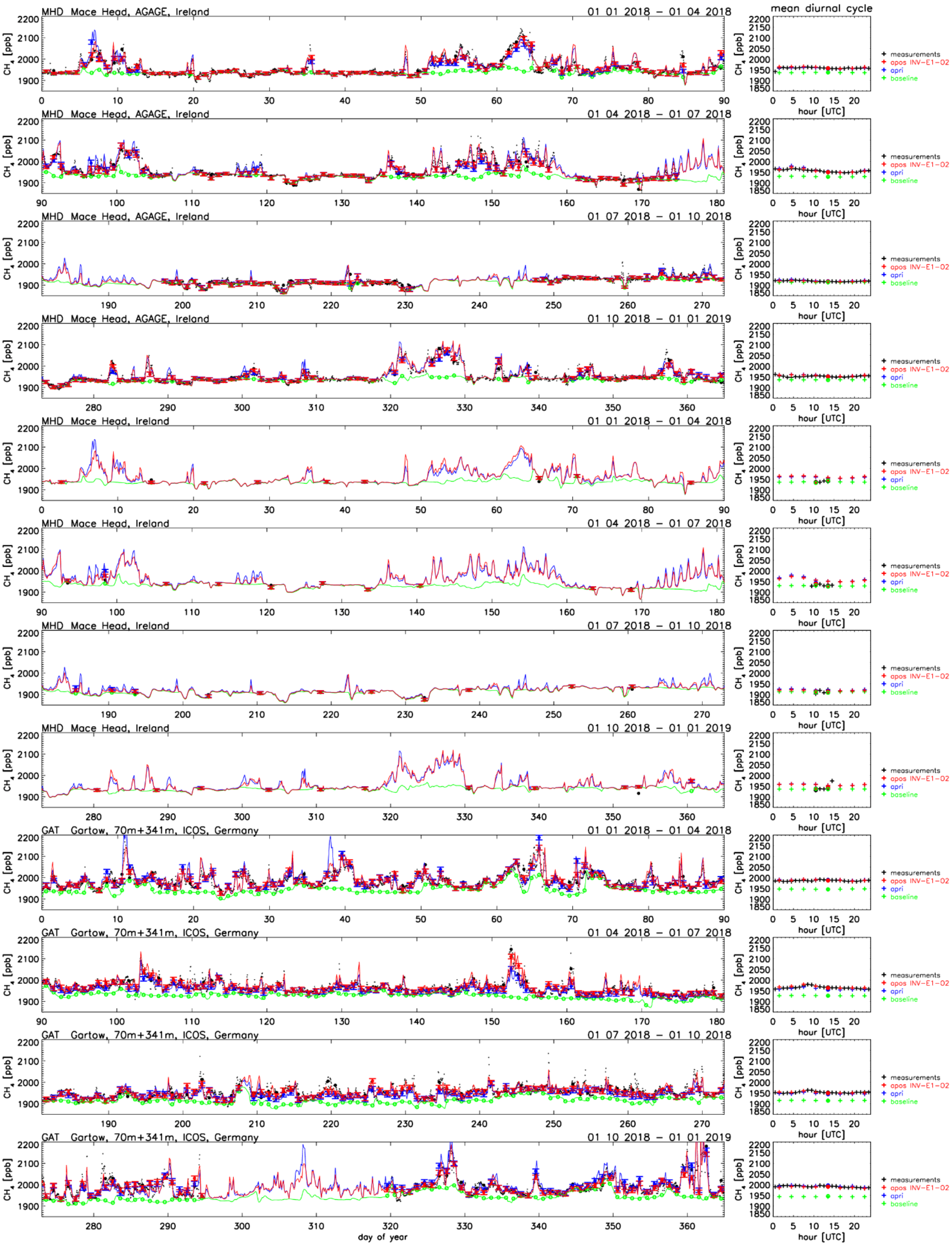

Figure S7: (cont.)

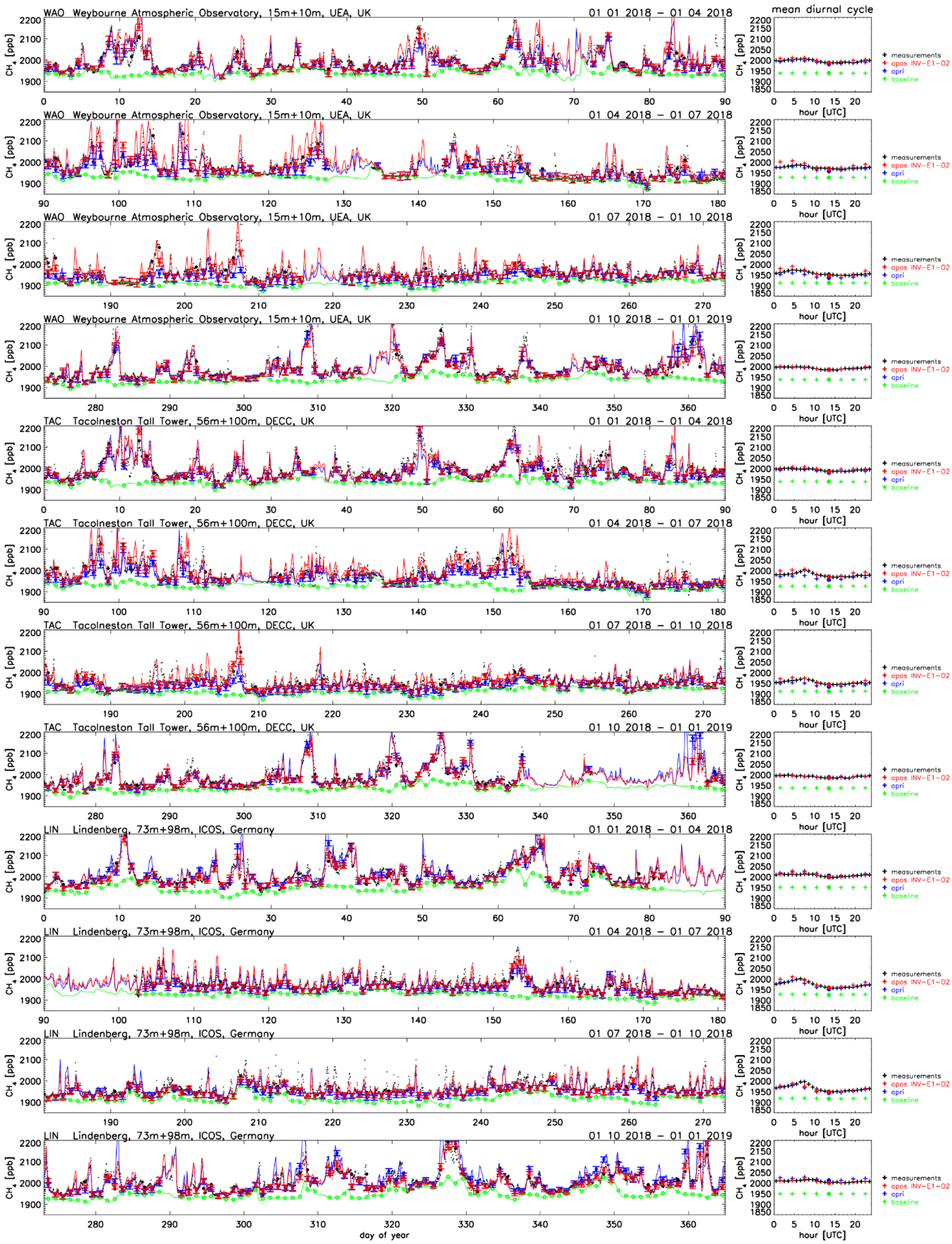

Figure S7: (cont.)

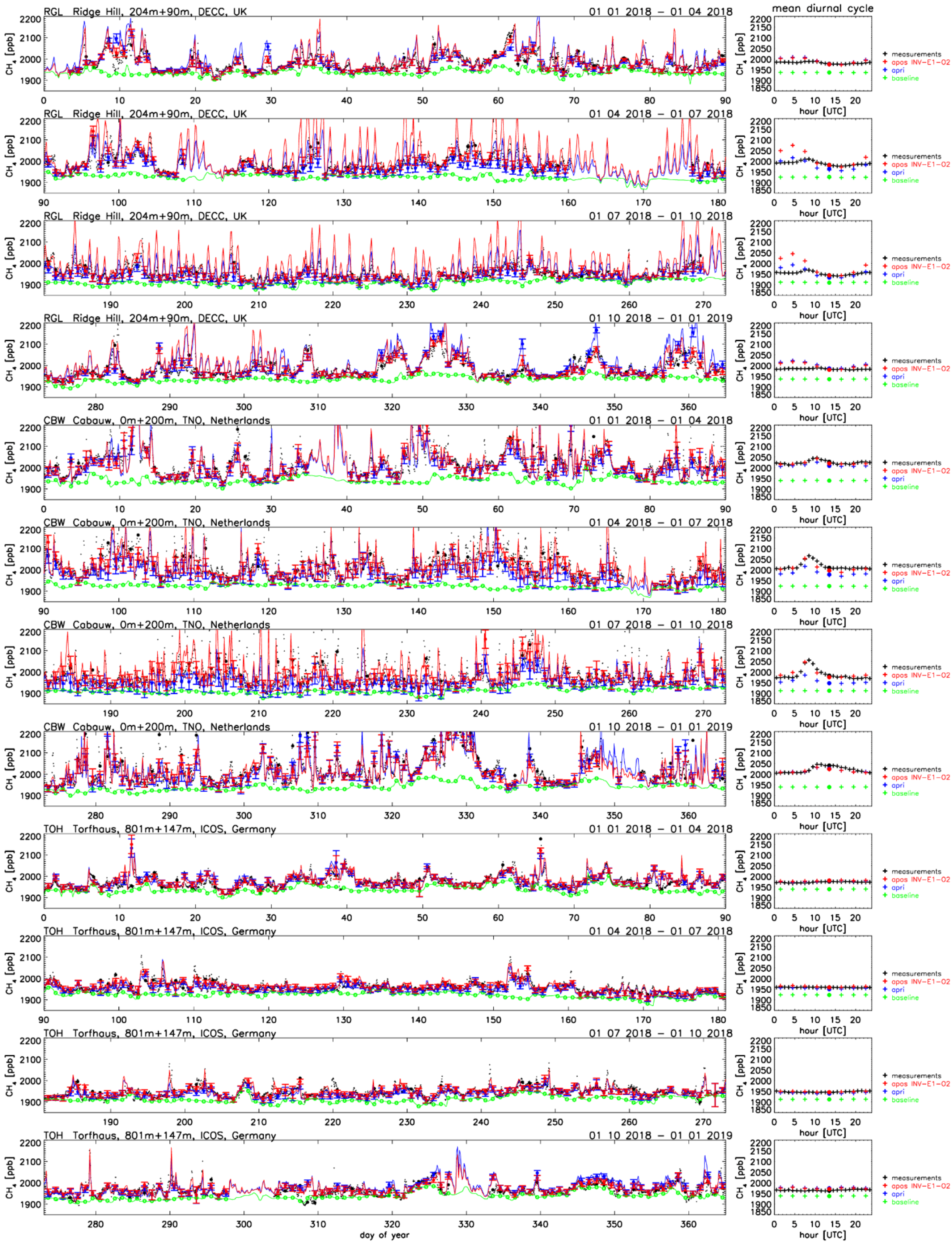

Figure S7: (cont.)

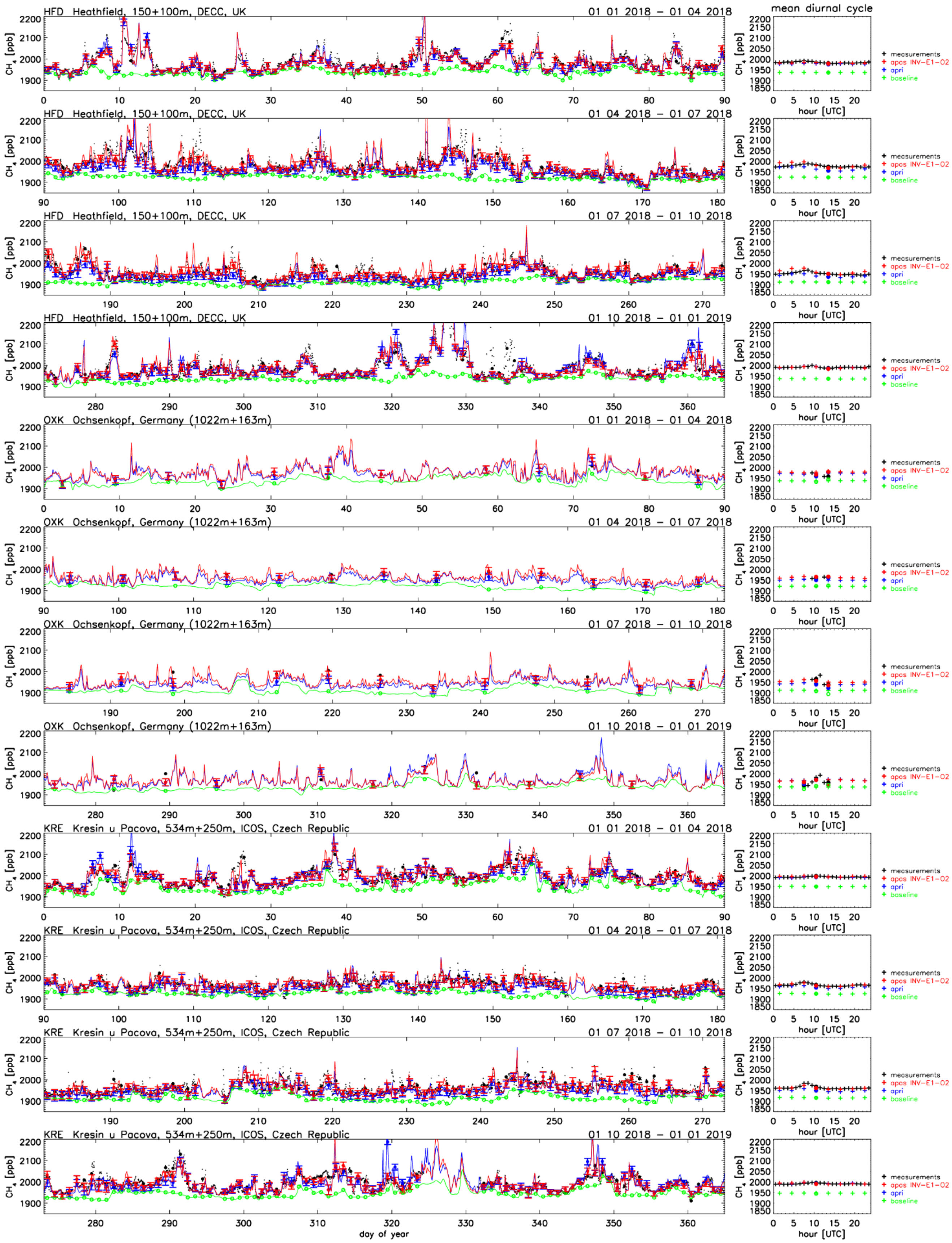

Figure S7: (cont.)

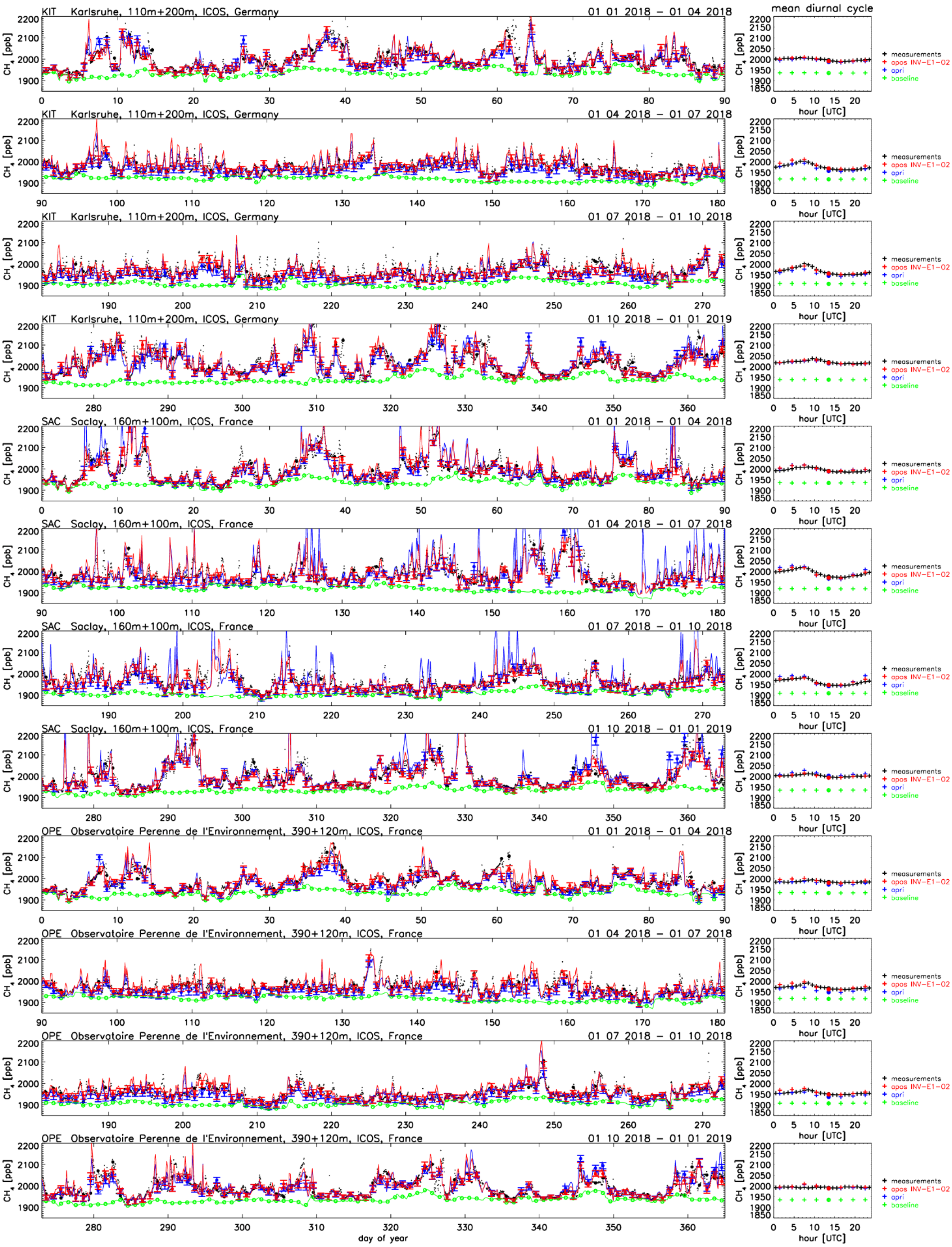

Figure S7: (cont.)

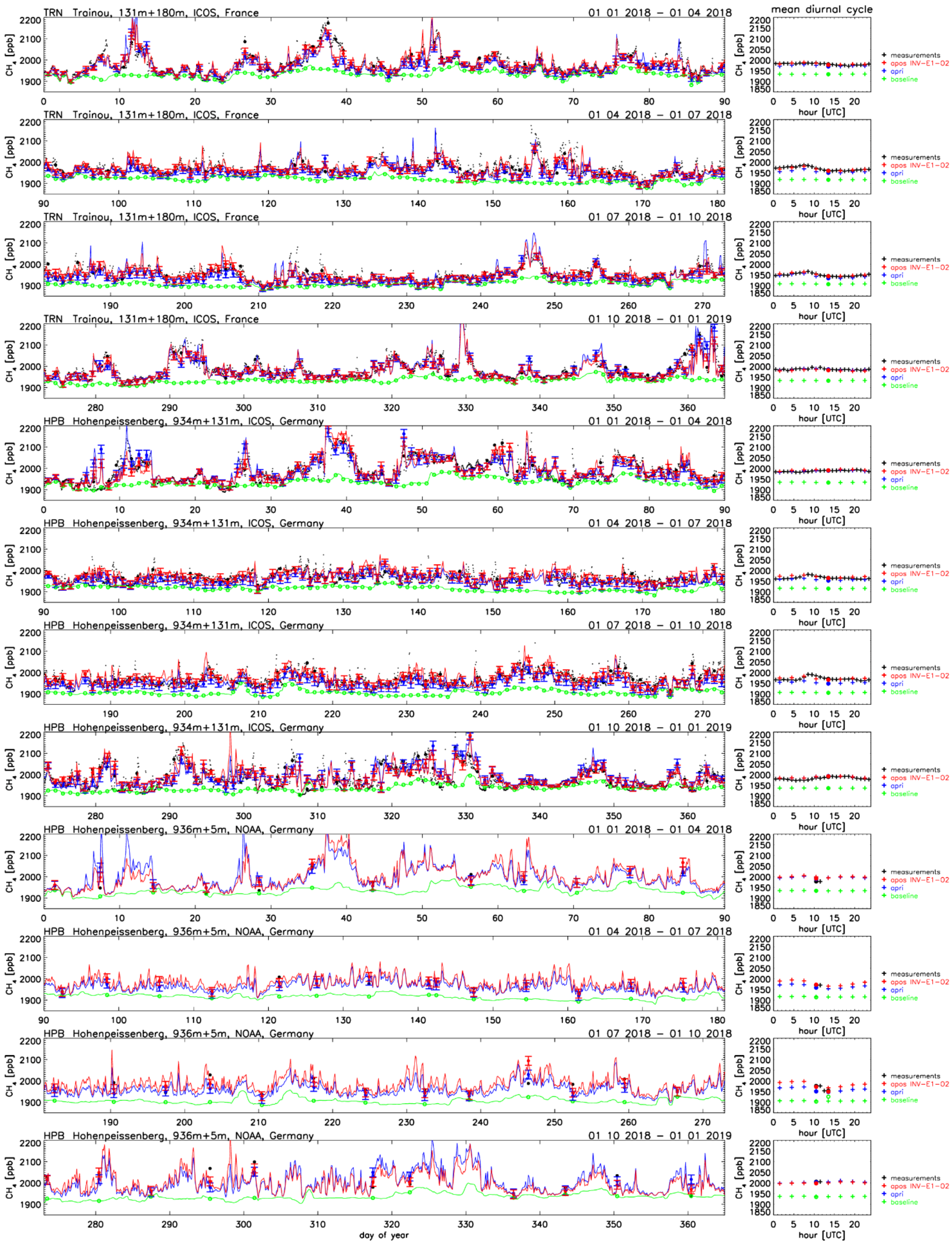

Figure S7: (cont.)

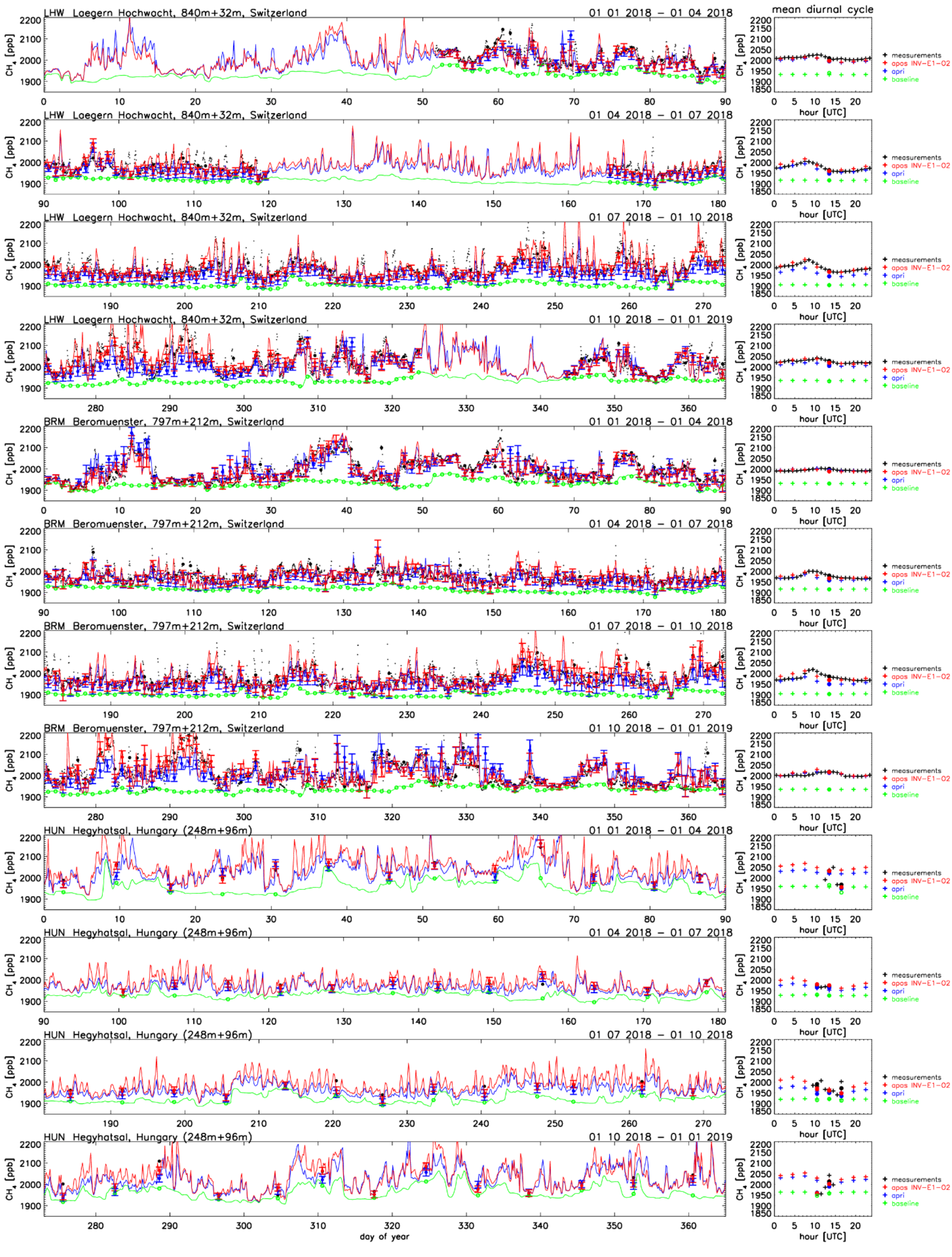

Figure S7: (cont.)

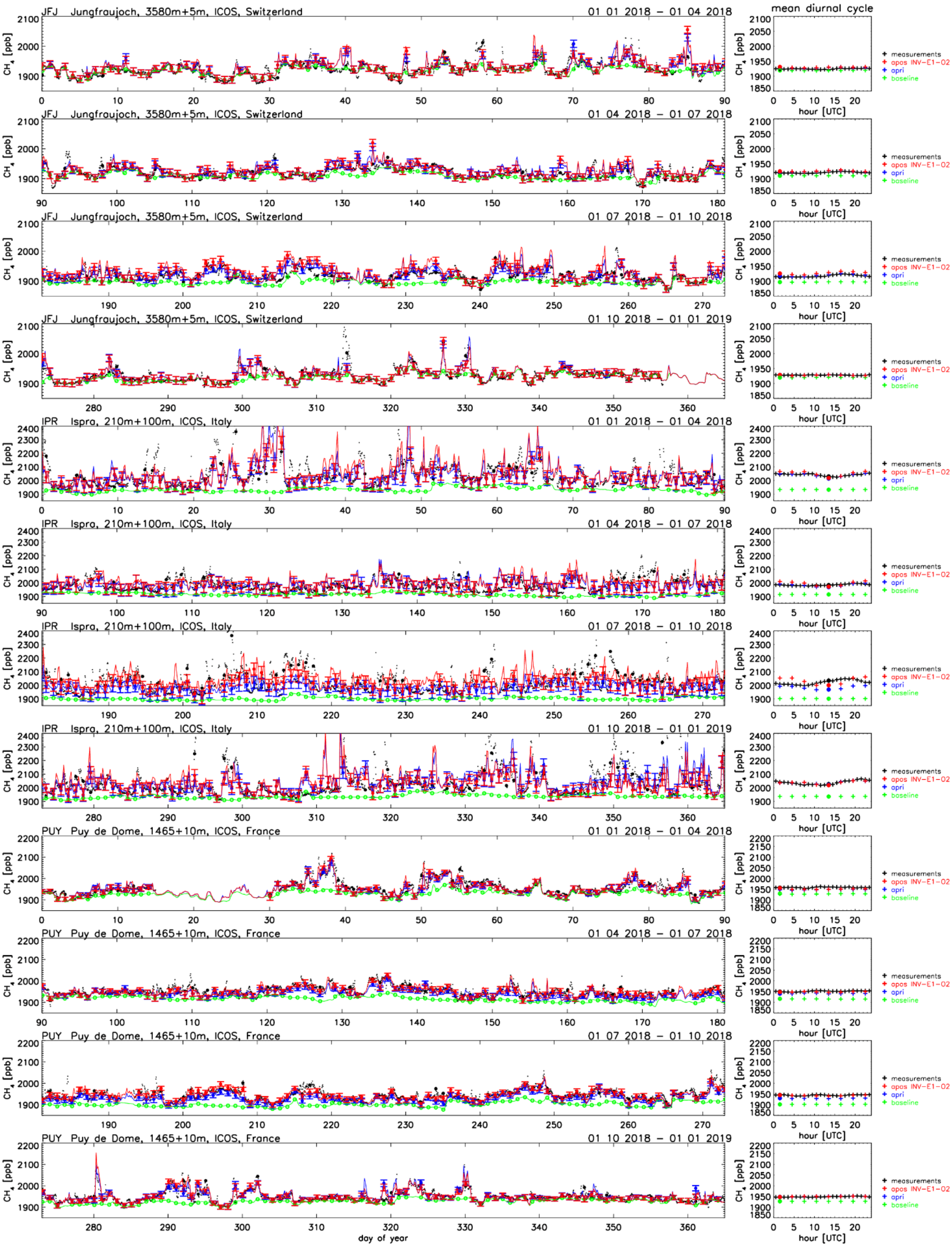

Figure S7: (cont.)

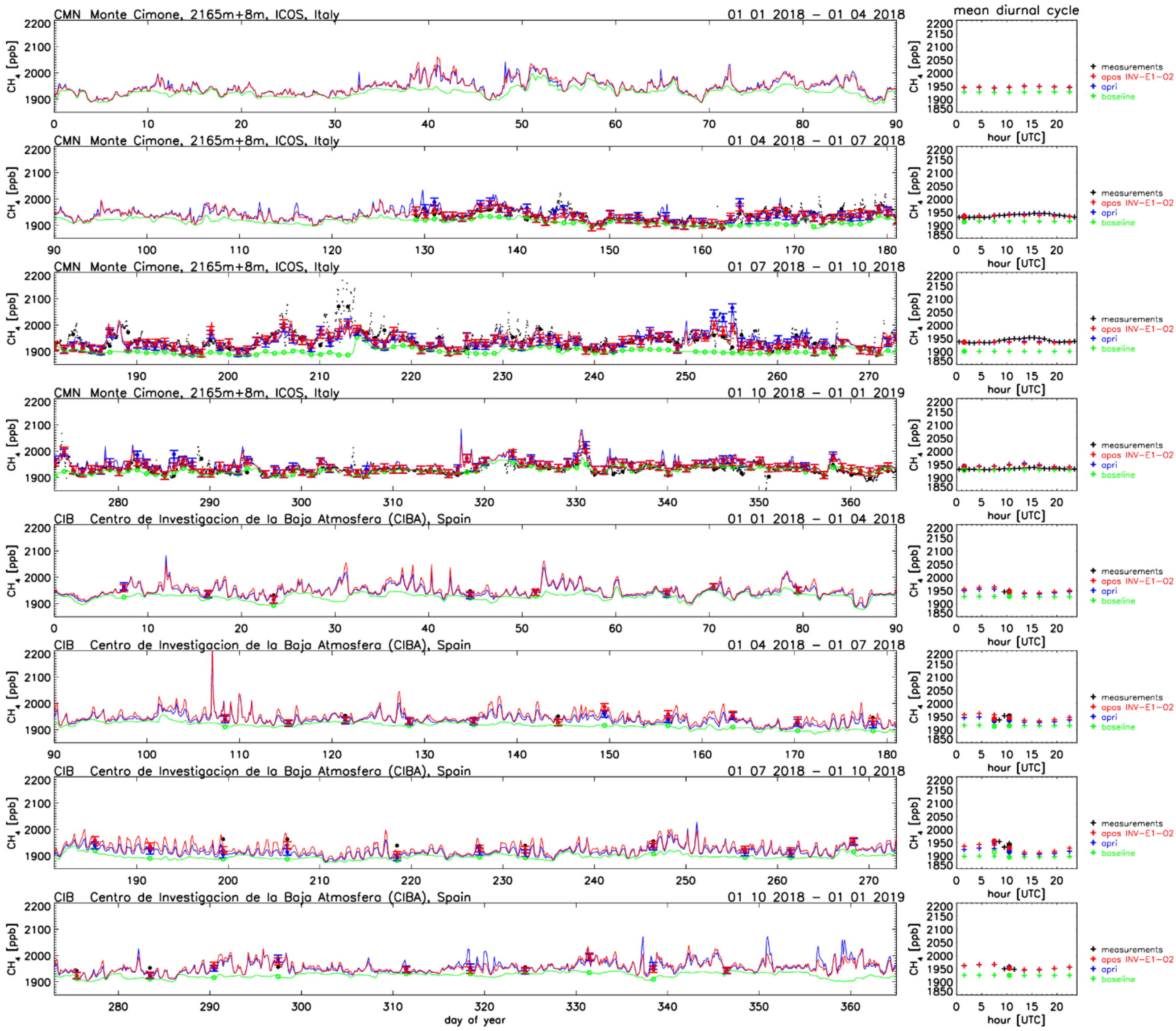

Figure S7: (cont.)