Supplement to Technical note: An interactive dashboard to facilitate quality control of in-situ atmospheric composition measurements

Yuri Brugnara[1], Martin Steinbacher[1], Simone Baffelli[2], Lukas Emmenegger[1]

[1]Laboratory for Air Pollution / Environmental Technology, Empa, Dübendorf, 8600, Switzerland
[2]Scientific IT, Empa, Dübendorf, 8600, Switzerland

Table S1: List of GAW stations supported by GAW-qc

GAW ID	Name	Latitude	Longitude	Elevation (m)
ALT	Alert	82.4991	-62.3415	185
AMY	Anmyeon-do	36.5386	126.3300	42
ARH	Arrival Heights	-77.8301	166.6605	187
BCK	Behchoko	62.7979	-115.9183	184
BHD	Baring Head	-41.4082	174.8708	85
BKT	Bukit Kototabang	-0.2019	100.3181	864
BMW	Tudor Hill (Bermuda)	32.2700	-64.8800	30
BOS	Boulder Table Mountain (CO)	40.1250	-105.2370	1689
BRA	Bratt's Lake	50.2016	-104.7113	580
BRW	Barrow (AK)	71.3230	-156.6115	11
CBY	Cambridge Bay	69.1284	-105.0577	25
CGO	Kennaook / Cape Grim	-40.6822	144.6883	94
CGR	Capo Granitola	37.5711	12.6597	5
CHC	Chacaltaya	-16.2000	-68.1000	5340
CHL	Churchill	58.7379	-93.8206	16
CMN	Monte Cimone	44.1934	10.7010	2165
CPA	Cholpon-Ata	42.6369	77.0675	1613
CPS	Chapais	49.8216	-74.9746	383
CPT	Cape Point	-34.3535	18.4897	230
CVO	Cape Verde Atmospheric Observatory	16.8640	-24.8675	10
DCC	Concordia, Dôme C	-75.0999	123.3335	3233
DIG	Diabla Gora / Puszcza Borecka	54.1500	22.0667	157
EGB	Egbert	44.2310	-79.7838	255
ESP	Estevan Point	49.3829	-126.5441	7

EST	Esther	51.6706	-110.2063	707
ETL	East Trout Lake	54.3537	-104.9869	500
EUK	Eureka	80.0500	-86.4167	610
FKL	Finokalia	35.3378	25.6694	150
FSD	Fraserdale	49.8400	-81.5167	210
GAT	Gartow	53.0655	11.4427	69
GLH	Giordan Lighthouse	36.0722	14.2184	167
GSN	Gosan	33.2938	126.1628	71
HBA	Halley	-75.5715	-25.5039	30
HEL	Helgoland	54.1804	7.8832	43
HFD	Heathfield	50.9767	0.2306	160
HKG	Hok Tsui / Cape d Aguilar	22.2094	114.2535	53
HKO	King's Park	22.3119	114.1729	65
HPB	Hohenpeissenberg	47.8015	11.0096	985
HUN	Hegyhátsál háttérszennyettség-mérő állomás	46.9559	16.6521	248
INU	Inuvik	68.3178	-133.5342	107
IRB	Iskrba	45.5612	14.8580	540
IZO	Izaña (Tenerife)	28.3090	-16.4994	2373
JCZ	Jarczew	51.8167	21.9833	180
JFJ	Jungfraujoch	46.5475	7.9851	3580
JUE	Jülich	50.9102	6.4096	98
KIT	Karlsruhe	49.1000	8.4380	111
KMW	Kollumerwaard	53.3333	6.2667	N/A
KOS	Kosetice Observatory	49.5833	15.0833	534
KPS	Kecskemét K-puszta háttérszennyettség-mérő állomás	46.9667	19.5833	125
LAU	Lauder	-45.0380	169.6840	370
LEB	Leba	54.7500	17.5333	2
LIN	Lindenberg	52.2167	14.1167	112
LLB	Lac La Biche (Alberta)	54.9538	-112.4666	548
LMP	Lampedusa	35.5182	12.6305	45
MKN	Mt. Kenya	-0.0622	37.2972	3678
MLO	Mauna Loa (HI)	19.5362	-155.5762	3397
MNM	Minamitorishima	24.2883	153.9833	7
NGL	Neuglobsow	53.1428	13.0333	62
NMY	Neumayer	-70.6660	-8.2660	42
NWR	Niwot Ridge - T-van (CO)	40.0500	-105.5900	3523
OXK	Ochsenkopf	50.0301	11.8084	1185
PAL	Pallas	67.9736	24.1158	560
PAY	Payerne	46.8129	6.9435	490

PDI	Pha Din	21.5731	103.5157	1466
PUY	Puy de Dôme	45.7723	2.9658	1465
RCV	Rucava	56.1620	21.1732	18
RGL	Ridge Hill	51.9976	-2.5400	204
RIG	Rigi	47.0674	8.4633	1031
RUN	La Réunion	-21.0796	55.3841	2160
RYO	Ryori	39.0319	141.8222	260
SMO	Samoa (Cape Matatula)	-14.2475	-170.5645	77
SNB	SONNBLICK Observatory	47.0539	12.9589	3106
SNZ	Sniezka	50.7333	15.7333	1603
SPO	South Pole	-89.9969	-24.8000	2841
SSL	Schauinsland	47.9000	7.9167	1205
STE	Steinkimmen	53.0431	8.4588	29
SUM	Summit	72.5800	-38.4800	3238
SVT	Svratouch	49.7333	16.0333	737
SYO	Syowa	-69.0053	39.5811	29
TAC	Tacolneston Tall Tower	52.5177	1.1386	56
THD	Trinidad Head (CA)	41.0541	-124.1510	107
TIK	Tiksi	71.5862	128.9188	8
TLL	El Tololo	-30.1683	-70.8036	2154
TOH	Torfhaus	51.8088	10.5350	801
ULD	Ulleungdo	37.4800	130.9000	221
USH	Ushuaia	-54.8485	-68.3107	18
WES	Westerland	54.9231	8.3080	12
WLG	Mt. Waliguan	36.2875	100.8963	3810
WSA	Sable Island	43.9326	-60.0086	2
YON	Yonagunijima	24.4667	123.0106	30
ZEP	Zeppelin Mountain (Ny Ålesund)	78.9067	11.8893	475
ZRN	Zavodnje	46.2600	15.0000	770
ZSF	Zugspitze-Schneefernerhaus	47.4165	10.9796	2666
ZSN	Zoseni	57.1351	25.9056	183

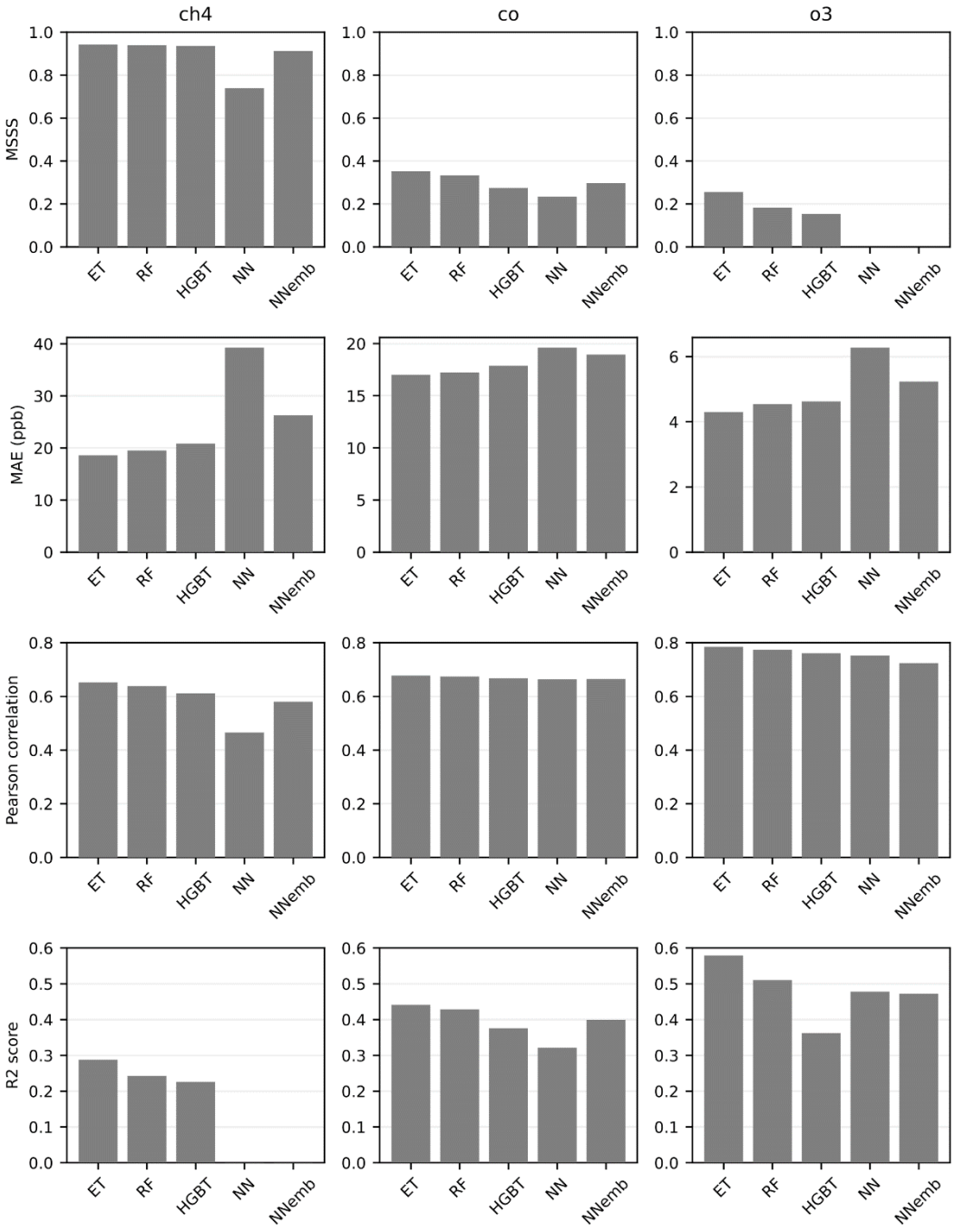

Figure S1: Performance of five tested regression models (ET=Extra-Trees; RF=Random Forest; HGBT=Histogram-based Gradient Boosting; NN=feedforward Neural Network with two hidden layers; NNemb=feedforward Neural Network with three hidden layers and stations as embedded variable) for the same sets of GAW stations (27 stations for CH_4, 24 stations for CO, 30 stations for O_3). All models were trained over the period 2020-2022 and validated using the year 2023. Negative scores are omitted.

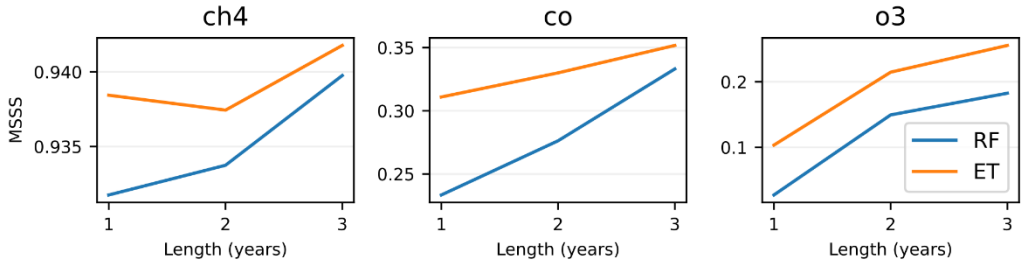

Figure S2: Average Mean Squared Skill Score (MSSS) at available GAW stations using training periods of different length (all ending on 31.12.2022) to predict the measurements of 2023 (RF = Random Forest; ET = Extra-Trees)

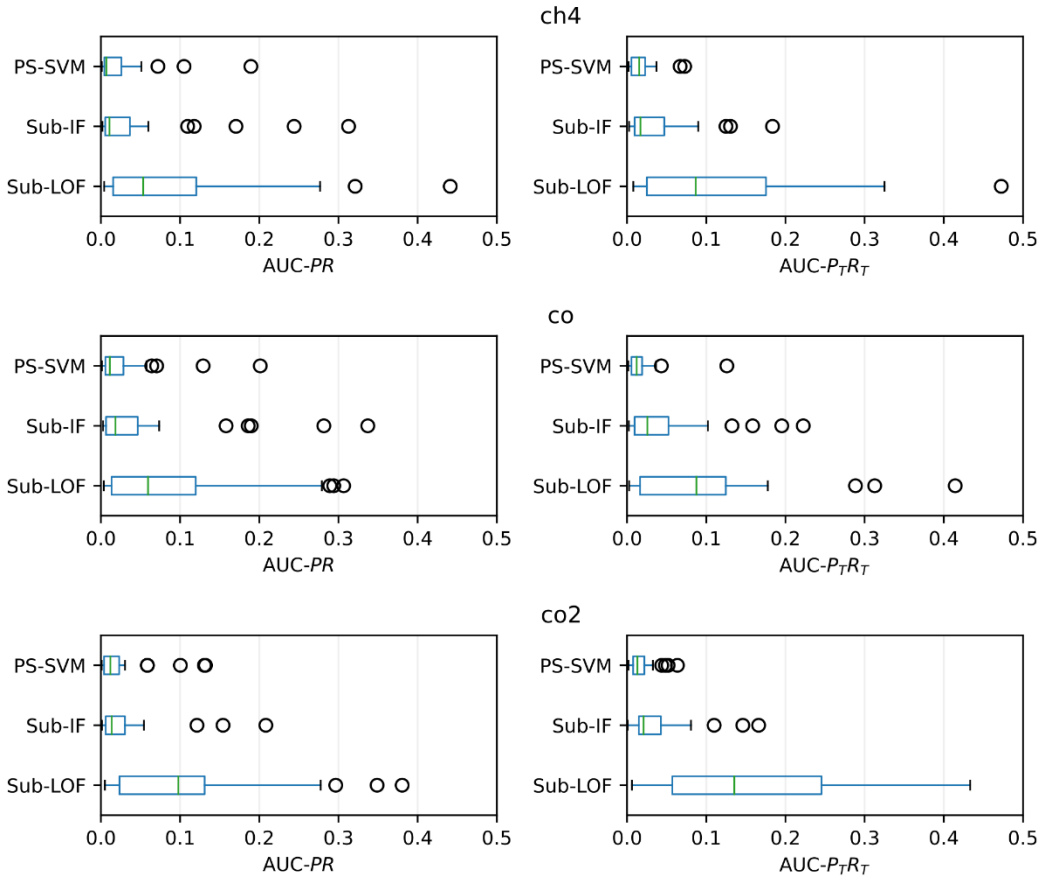

Figure S3: Distribution of the Area Under the Precision-Recall Curve (AUC-PR) score (left column: classic; right column: range-based) for three different outlier detection methods for sequences (for n_s=5 and k=100) based on ICOS stations that have at least 30 flags in Level 2 data between April 2021 and March 2024 (for towers the largest available inlet height is used). The green lines indicate the median of the tested stations, the boxes indicate the interquartile range (IQR), while the whiskers extend to up to 1.5 IQR from the edges of the boxes.

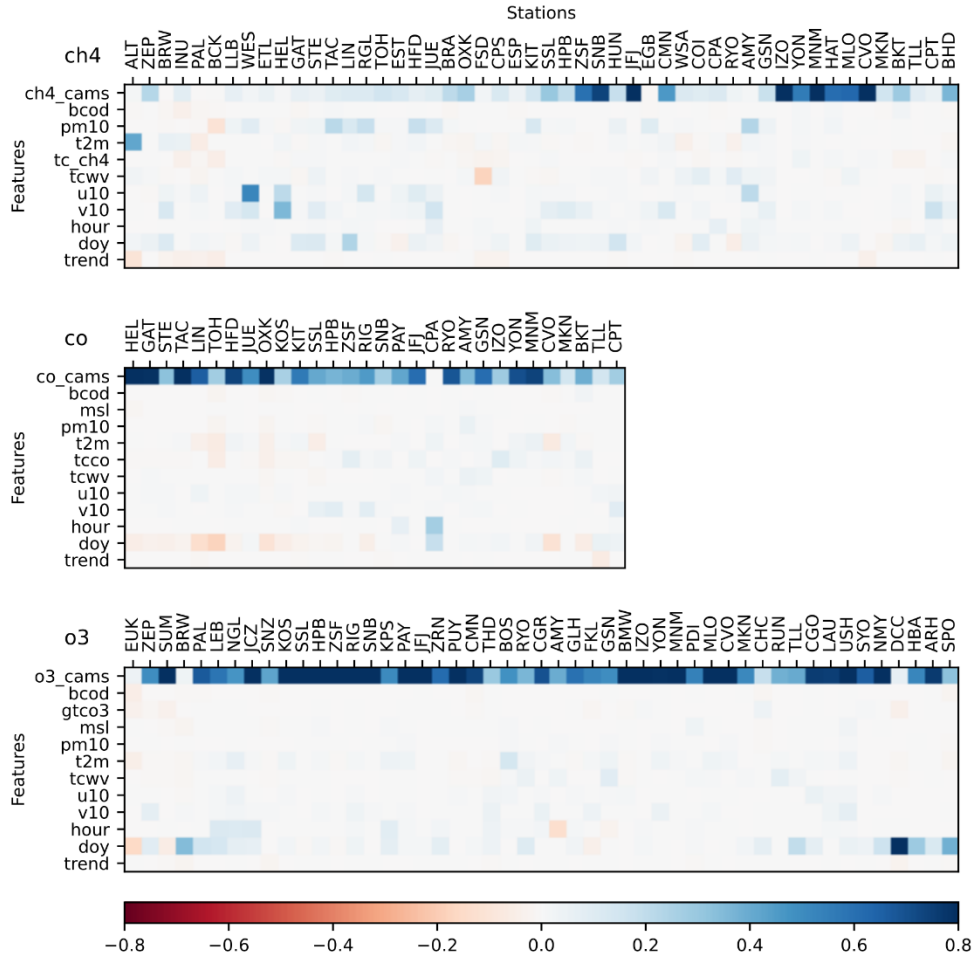

Figure S4: Minimum feature importance for the three available gas species at validation stations (identified by the GAW code and ordered by decreasing latitude), based on 2020-2023 data and the MSSS. The feature importance is calculated separately for each year and then the minimum of the yearly values is taken (at least 3 years required). For tower stations the highest sampling height is used. For station names and coordinates see Table S1.