S.1) Evaluation of meteorological conditions

The models' meteorological conditions, temperature (T), wind speed (Ws), and specific humidity (Qv), were evaluated against ERA5 and flight measurements for the Scientific Aviation (SA) flights. The evaluation was performed using direct comparison and three quantitative metrices. Table S1 shows the results of the comparison between simulated and measured meteorological parameters for the SA flights.

Table S1 - Difference between simulated and measured meteorological parameters temperature (T, wind speed (Ws) and specific humidity (Qv). Average values along the flight track are used to calculate the differences. Green color indicates that the simulation results and measurements are consistent within the thresholds, orange color indicates that simulation and measurements deviate more than the thresholds. The thresholds for the various variables are: temperature < 1.5□C, wind speed < 0.6 m s-1, and Qv between 25th and 75th percentile.

No	Date	Hour	Model - OBS					
			T (°C)		Ws (m/s)		Qv	
			MECO(3)	COSMO-GHG	MECO(3)	COSMO-GHG	MECO(3)	COSMO-GHG
1	03/10/2019	11:49 - 15:03	-2.1	-1.3	-1.0	-1.0	0.00178	0.00069
2	07/10/2019	11:05 - 14:41	2.5	-1.2	-2.1	-0.4	0.00005	0.00058
3	08/10/2019	06:53 - 11:38	-1.8	-1.2	1.4	0.4	-0.00051	-0.00036
4	09/10/2019	09:08 - 13:52	-4.2	-1.9		0.9		
5	10/10/2019	08:35 - 11:47	-2.9	-1.9	-0.9	1.5	-0.00159	0.00024
6	11/10/2019	09:07 - 13:17	-3.1	-2.6	0.3	0.2	0.00096	0.00017
7	12/10/2019	10:05 - 14:24	-3.9	-2.4	-1.7	0.3	0.00016	0.00061
8	15/10/2019	10:04 - 14:34	-3.7	-1.8	-1.4	-0.3	-0.00062	0.00044
9	17/10/2019	10:00 - 13:43	-0.6	-2.5	0.3	-0.1	-0.00246	0.00071
10	18/10/2019	11:26 - 14:36	-3.8	-2.2	-1.0	0.4	-0.00024	0.00173
11	21/10/2019	10:57 - 14:30	-5.6	-1.7	-1.7	-1.9	-0.00031	-0.00131

In addition, three quantitative indices, the Nash-Sutcliffe Efficiency (NSE), the Kling-Gupta Efficiency (KGE) and the Mean Absolute Relative Error (MARE) were used to evaluate the performance of the models compared to reanalysis data from ERA5 and flight measurements.

NSE is calculated as one minus the ratio of the error variance of the modeled time-series divided by the variance of the reference time-series, here we use the ERA5 time-series (Nash and Sutcliffe, 1970). It is defined as:

$$NSE = 1 - \frac{\sum_{t=1}^{T}(Q_m^t - Q_0^t)^2}{\sum_{t=1}^{T}(Q_o^t - \bar{Q}_o)^2}$$ Eq. S1

where $\bar{Q}_o$ is the mean of observed data, Q_m^t is model data at time t, and Q_o^t is observed data at time t.

KGE is a goodness-of-fit indicator and was developed based on a decomposition of NSE into its constitutive components (correlation, variability bias, and mean bias) (Gupta et al., 2009;Knoben et al., 2019). The KGE is calculated as follows:

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2}$$ Eq. S2

where σ_{obs} is the standard deviation in observations, σ_{sim} is the standard deviation in simulations, μ_{sim} is the simulation mean, μ_{obs} is the observation mean, and r is the linear correlation between observations and simulations.

29   MARE is a measure of absolute errors between simulation and observation time series normalized by
30   the observation series (Lee and Deming, 1998) and is defined as:

31    $$MARE = \frac{\sum_{i=1}^{N}|e_i - s_i|}{\sum_{i=1}^{N} e_i}$$ Eq. S3

32   where N is the length of the simulations and evaluation periods, e is the observed series, and s is the
33   model series.

34   NSE equal to 1 indicates perfect correspondence between models and ERA5 (green color); NSE equal
35   to 0 indicates that the model simulations have the same explanatory power as the mean of the ERA5;
36   and NSE < 0 indicates that the model is a worse predictor than the mean of the ERA5 (yellow color).
37   Similar to NSE, KGE equal to 1 indicates perfect agreement between models and ERA5 (green color)
38   a KGE score below 0 indicates poor model performance (yellow color). Unlike NSE and KGE, MARE
39   equal to 0 indicates a perfect fit to the observed data (ERA5, green color). There is no general threshold
40   to indicate poor model performance using MARE and we assess model performance as acceptable if
41   the MARE index lies below and equal to 0.2 (green color), otherwise model performance is assessed as
42   poor (yellow color). Table S2 summarizes the evaluation criteria and Table S3 and S4 shows the results
43   of model evaluation. It is important to note that the performance evaluation of the models in simulating
44   wind direction (Wdir) and cloud cover using quantitative metrics was not feasible. Therefore, we relied
45   on visual assessment of the simulation and reanalysis fields. For Wdir, a plus sign (dark green color)
46   was assigned if the models simulated Wdir similar to ERA5 within an angle <30 degrees, a circle sign
47   (green color) was assigned if the models simulated Wdir similar to ERA5 with an angle between 30 and
48   90 degrees (observed in only a few places), and a negative sign (orange color) was assigned if the
49   simulated Wdir significantly deviated more than 90 degrees from ERA5. Similarly for cloud cover, a
50   plus sign was assigned if both models and ERA5 indicated clear sky and/or if the models simulated
51   slightly different cloud patterns at 850 hPa compared to ERA5. A circle sign was assigned if the models
52   simulated slightly different cloud pattern at 850 hPa compared to ERA5 but overestimated cloud
53   fraction in the flight areas. Lastly, a minus sign was assigned if the models simulated different cloud
54   patterns during the flight times.

55   *Table S2 - Evaluation criteria for good and poor performance*

No	Metric	Good	Poor
1	NSE	NSE = 1	NSE < 0
2	KGE	KGE = 1	KGE = 0
3	MARE	MARE = 0	MARE < 0.2

Table S3 - Model performance in simulating temperature (T) and wind speed (Ws)

No	Date	Hour	Model vs ERA5											
---	---	---	MECO(3) T (°C)			COSMO-GHG T (°C)			MECO(3) Ws (m/s)			COSMO-GHG Ws (m/s)		
			NSE	KGE	MARE	NSE	KGE	MARE	NSE	KGE	MARE	NSE	KGE	MARE
1	01/10/2019	10:57 - 14:00	-5.79	0.85	0.14	0.69	0.81	0.03	-15.69	-1.82	0.52	-10.81	-1.24	0.48
2	02/10/2019	10:05 - 12:36	-4.16	0.75	0.14	-0.85	0.68	0.07	-18.13	-1.13	0.47	-12.04	-1.30	0.45
3	03/10/2019	08:10 - 11:25	-3.00	0.78	0.15	-1.06	0.88	0.11	-265.51	-2.13	1.98	-30.56	-2.85	0.55
4	03/10/2019	12:58 - 15:15	-83.85	-0.05	0.12	-39.61	-0.51	0.08	0.37	0.61	0.10	0.07	0.56	0.13
5	06/10/2019	07:34 - 11:22	0.46	0.83	0.06	0.00	0.86	0.09	-0.22	0.37	0.30	-0.19	0.59	0.36
6	07/10/2019	10:02 - 14:22	-36.37	0.42	0.57	-3.26	0.70	0.19	-6.24	-0.06	0.33	-2.15	0.32	0.22
7	09/10/2019	07:22 - 09:53	-9.48	0.51	0.24	-1.58	0.49	0.11	-12.28	-0.23	0.39	-11.03	-0.23	0.39
8	10/10/2019	08:15 - 11:43	0.79	0.68	0.10	0.82	0.78	0.10	0.48	0.75	0.25	-0.38	0.51	0.41
9	11/10/2019	08:48 - 11:52	0.56	0.93	0.04	0.69	0.66	0.03	-7.28	-0.34	0.49	-2.09	-0.12	0.23
10	12/10/2019	09:31 - 11:25	-1.97	0.84	0.08	-0.92	0.51	0.06	-9.14	0.30	0.48	-1.21	0.11	0.17
11	14/10/2019	09:15 - 12:21	-1.37	0.80	0.10	0.07	0.50	0.04	-3.68	-0.03	0.28	-2.68	0.55	0.29
12	17/10/2019	09:39 - 13:31	0.51	0.63	0.08	0.58	0.81	0.07	-17.77	-2.57	0.66	-16.54	-2.29	0.65
13	23/10/2019	09:21 - 13:45	-4.29	0.76	0.14	-1.36	0.06	0.07	-2.29	0.01	0.19	-3.83	-0.41	0.23
14	03/10/2019	11:49 - 15:03	-3.77	0.69	0.13	-1.65	0.80	0.09	-0.48	0.35	0.21	-0.53	0.54	0.22
15	07/10/2019	11:05 - 14:41	-13.85	0.27	0.72	0.52	0.87	0.13	-14.64	-0.73	0.29	-0.97	0.32	0.09
16	08/10/2019	06:53 - 11:38	0.71	0.72	0.11	0.47	0.83	0.17	-3.32	0.00	0.76	-0.19	0.55	0.40
17	09/10/2019	09:08 - 13:52	-4.35	0.60	0.22	0.39	0.63	0.06				-0.91	0.33	0.22
18	10/10/2019	08:35 - 11:47	0.72	0.69	0.12	0.84	0.77	0.10	0.62	0.76	0.32	0.50	0.66	0.33
19	11/10/2019	09:07 - 13:17	0.95	0.82	0.04	0.93	0.87	0.04	-6.94	-0.91	0.53	-7.59	-0.88	0.59
20	12/10/2019	10:05 - 14:24	-0.34	0.80	0.10	0.53	0.52	0.06	-3.85	0.13	0.35	-1.50	0.12	0.24
21	15/10/2019	10:04 - 14:34	-2.22	-0.15	0.09	0.66	0.90	0.03	-2.87	0.23	0.36	0.24	0.67	0.14
22	17/10/2019	10:00 - 13:43	-0.34	-0.10	0.05	-3.17	0.60	0.10	-24.02	-1.83	1.17	-6.74	-0.03	0.83
23	18/10/2019	11:26 - 14:36	-5.40	0.79	0.15	-0.43	0.69	0.07	-9.49	-1.18	0.63	-21.56	-0.99	1.03
24	21/10/2019	10:57 - 14:30	-17.74	0.57	0.22	0.52	0.64	0.03	-0.87	-0.02	0.29	-1.55	0.06	0.37

Table S4 - Model performance in simulating wind direction (Wdir), cloud cover, and humidity (Qv)

No	Date	Hour	MECO(3)		COSMO-GHG		MECO(3) Qv			COSMO-GHG Qv		
---	---	---	Wdir	Cloud	Wdir	Cloud	NSE	KGE	MARE	NSE	KGE	MARE
1	01/10/2019	10:57 - 14:00	+	+	o	+	-13.96	-0.17	0.26	-13.29	-0.48	0.25
2	02/10/2019	10:05 - 12:36	+	+	+	+	-2.41	0.81	0.08	-8.65	0.38	0.13
3	03/10/2019	08:10 - 11:25	+	+	+	+	-3.34	-0.74	0.09	-1.92	-0.65	0.07
4	03/10/2019	12:58 - 15:15	o	+	o	+	-74.51	-0.07	0.19	-18.99	-0.93	0.09
5	06/10/2019	07:34 - 11:22	-	-	+	-	-2.31	0.24	0.09	-0.69	0.35	0.06
6	07/10/2019	10:02 - 14:22	-	-	+	o	-12.49	-0.15	0.08	-3.69	-0.04	0.04
7	09/10/2019	07:22 - 09:53	+	+	+	+	-40.59	-0.12	0.16	-4.14	-0.45	0.05
8	10/10/2019	08:15 - 11:43	+	-	+	+	-32.48	-0.04	0.36	-3.21	0.01	0.11
9	11/10/2019	08:48 - 11:52	o	o	o	o	-3.86	0.42	0.14	-1.25	0.49	0.08
10	12/10/2019	09:31 - 11:25	o	+	+	+	-10.29	-0.33	0.14	-9.35	-0.44	0.12
11	14/10/2019	09:15 - 12:21		+	+	+	-0.59	0.01	0.05	-6.04	-0.11	0.10
12	17/10/2019	09:39 - 13:31		+	o	+	-0.43	0.56	0.11	-0.76	0.34	0.13
13	23/10/2019	09:21 - 13:45					-0.54	0.14	0.08	-1.60	0.27	0.12
14	03/10/2019	11:49 - 15:03	-	+	+	+	-55.89	-0.96	0.19	-17.37	-2.27	0.09
15	07/10/2019	11:05 - 14:41	-	o	+	o	-6.16	-0.19	0.09	-25.63	0.41	0.18
16	08/10/2019	06:53 - 11:38	-	-	o	+	-11.74	0.09	0.12	-7.75	-0.03	0.09
17	09/10/2019	09:08 - 13:52	+	+	+	+				-1.51	0.13	0.09
18	10/10/2019	08:35 - 11:47	+	-	+	+	-41.45	-0.99	0.33	-2.63	-0.19	0.07
19	11/10/2019	09:07 - 13:17	o	o	o	o	-2.22	0.31	0.19	0.40	0.68	0.08
20	12/10/2019	10:05 - 14:24	o	+	+	+	-4.47	-0.49	0.16	-7.89	0.15	0.21
21	15/10/2019	10:04 - 14:34	+	+	o	+	0.41	0.61	0.11	-1.09	0.42	0.21
22	17/10/2019	10:00 - 13:43		+	-	+	-11.21	0.19	0.29	-3.00	0.43	0.16
23	18/10/2019	11:26 - 14:36	-	+	-	+	0.42	0.56	0.06	-10.82	0.45	0.31
24	21/10/2019	10:57 - 14:30					-0.04	0.45	0.09	-0.39	0.52	0.10

After we evaluated the models' performance in simulating the meteorological conditions during flights, we investigated the simulation-measurement comparison for the individual flights for similar parameters. We assign a plus sign (green) for good model performance, a circle sign (yellow) for acceptable model performance, and a minus sign (red) for poor models' performance. The assignment

of the model performance (red, yellow, and green) was based on the three quantitative indices (NSE, KGE, and MARE, see above). Green if at least 2 metrics show good agreement between model and ERA5, yellow if one metric shows good agreement, and red if none of the metrics shows good agreement. For Wdir and cloud cover, we simply used the evaluation results outlined in Table S4. Table S5 shows the summary of the evaluation of the models' outputs during the flight measurement dates.

Table S5 - Summary of the models' evaluation

| No	Date	Region	MECO(3) Parameter					COSMO-GHG Parameter				
			T	Cloud	Ws	W dir	Qv	T	Cloud	Ws	W dir	Qv
1	01/10/2019	6	o	+	-	+	-	+	+	o	o	-
2	02/10/2019	7	o	+	-	+	+	+	+	o	+	+
3	03/10/2019	4	o	+	-	+	o	o	+	o	+	o
4	03/10/2019	5a	-	+	+	o	-	-	+	+	o	o
5	06/10/2019	4	+	-	o	-	+	+	-	o	+	+
6	07/10/2019	5a	-	-	-	-	o	o	o	-	+	o
7	09/10/2019	7	-	+	-	+	o	+	+	-	+	o
8	10/10/2019	11	+	-	o	+	-	+	+	-	+	o
9	11/10/2019	8	+	o	-	o	o	+	o	o	o	o
10	12/10/2019	6	+	+	o	o	0	+	+	+	+	o
11	14/10/2019	8	o	+	-		o	+	+	o	+	-
12	17/10/2019	15-16	+	+	-		+	+	+	-	o	o
13	23/10/2019	6-7	o		o		o	o		+		o
14	03/10/2019	4.4, 4.5, 4.6	o	+	o	-	-	+	+	o	+	o
15	07/10/2019	8, 8.1	-	o	-	-	o	+	o	+	+	o
16	08/10/2019	5a	+	-	-	-	+	+	+	o	o	o
17	09/10/2019	7, 7.5	-	+		+		+	+	o	+	
18	10/10/2019	11	+	-	o	+	o	+	+	o	+	o
19	11/10/2019	4.2, 4.3, 4.4, 4.5, 4.6	+	o	-	o	-	+	o	-	o	+
20	12/10/2019	7, 6.5	o	+	o	o	o	+	+	o	+	-
21	15/10/2019	5a	-	+	o	+	+	+	+	+	o	-
22	17/10/2019	6.2-6.7	o	+	-		o	o	+	-	-	o
23	18/10/2019	7.2, 7.3, 7.4, 7.8	o	+	o	-	+	+	+	-	-	-
24	21/10/2019	7c, 7.1, 6.1, 7.8	-		o		+	+		o		+

		No data
		Good
		Acceptable
		Poor

S.2) Measured and simulated plume areas from the mass balance flights (zoomed in figures)
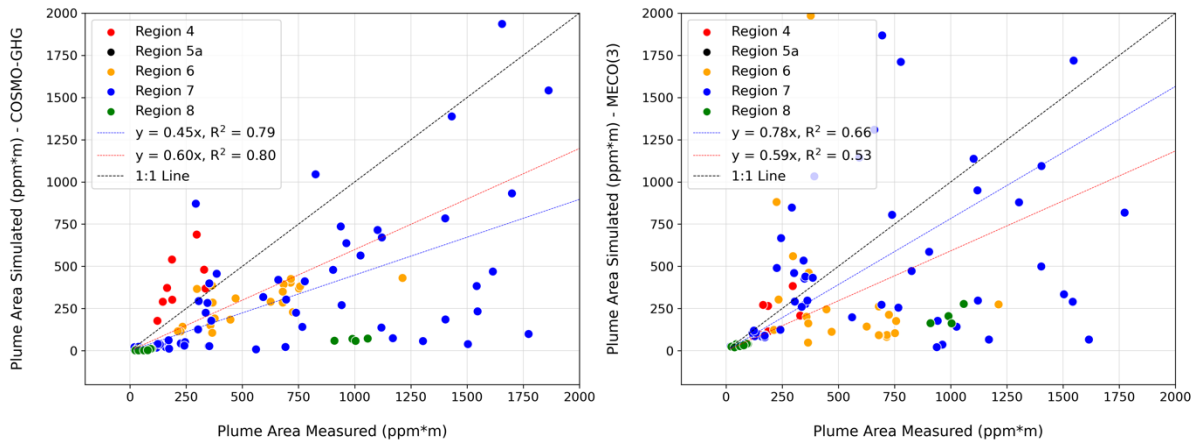
Figure S1 - Comparisons between plume areas calculated from measurements and simulations with the COSMO-GHG model (left) and MECO(3) model (right). Blue dashed lines show linear fits to all data and red dashed lines linear fits to the plumes from the clusters only, without the points from the larger regions.

S.3) Major farm and landfill methane emitters within the study areas from the E-PRTR/IED inventory
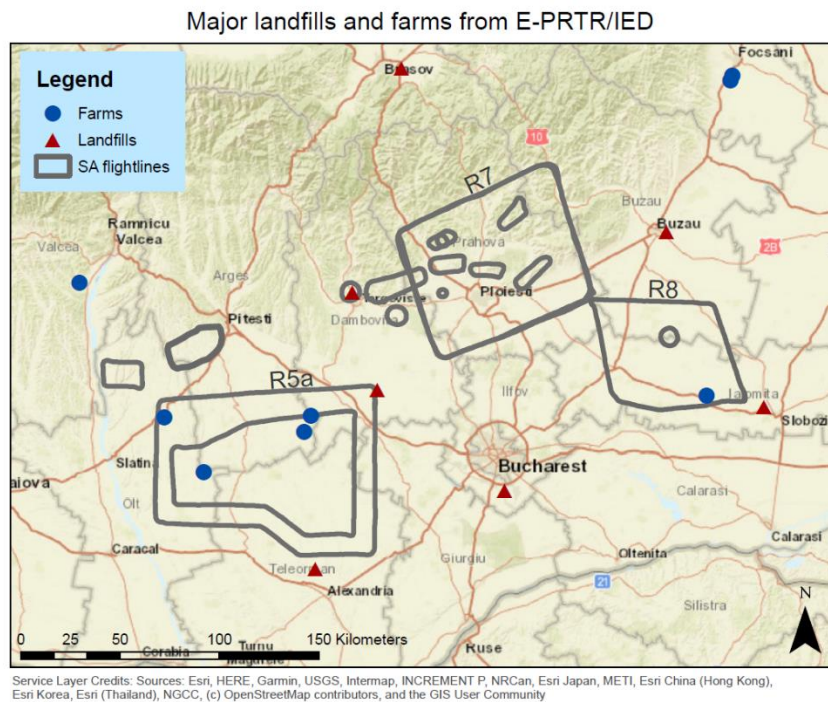
Figure S2 - Location of farms and landfills with significant methane emissions from the E-PRTR/IED inventory within the flightpaths of the mass balance flights.

S.4) Measured emission rates (ER) and estimates of the O&G related fraction of total CH4 emissions in target regions and clusters using the EDGAR inventory.

Table S6 - Measured emission rates (ER) and estimates of the O&G related fraction of total CH_4 emissions in target regions and clusters. "Non-O&G emissions (kg hr^{-1})" are extracted from the EDGAR inventory for the target regions and are used to derive ERs from the O&G industry in the area (column O&G emissions). The last column shows the emission factor (kg CH_4 hr^{-1} $site^{-1}$). Numbers in bold are used for upscaling to the national scale (see text for details).

Flight ID	Target region/cluster	# facilities	# wells	Total Emissions (kg hr^{-1})	Non-O&G emissions (kg hr^{-1})	O&G emissions (kg hr^{-1})	EF (kg h^{-1} $site^{-1}$)
SA01	R7	496	337	8517 ± 2097	1903	6614 ± 2097	13.3 ± 4.2
SA02	R7	504	343	8335 ± 1440	1958	6377 ± 1440	12.7 ± 2.9
SA03	R5a	827	654	4556 ± 2570	808	3748 ± 2570	4.6 ± 3.1
SA04	R5a-small	818	642	2920 ± 935	367	2553 ± 935	3.1 ± 1.1
SA05	R6C2C3C4	471	379	1729 ± 912	-	1729 ± 912	3.7 ± 1.9
SA06	R7C3C4	124	92	1481 ± 287	-	1481 ± 287	11.9 ± 2.3
SA07	R7C2	71	44	1395 ± 546	-	1395 ± 546	19.6 ± 7.7
SA08	R7VentArea	67	41	602 ± 209	-	602 ± 209	9.0 ± 3.1
SA09	R4C5	390	347	477 ± 106	-	477 ± 106	1.2 ± 0.3
SA10	R6C6	29	16	469 ± 170	56†	413 ± 170	14.2 ± 5.9
SA11	R7Vent	37	20	266 ± 113	-	266 ± 113	7.2 ± 3.1
SA12	R7C5	59	45	259 ± 47	-	259 ± 47	4.4 ± 0.8
SA13	R4C2C3	247	186	246 ± 89	-	246 ± 89	1.0 ± 0.4
SA14	R6C5	27	21	131 ± 85	-	131 ± 85	4.9 ± 3.1
SA16	R8C1	29	19	90 ± 49	-	90 ± 49	3.1 ± 1.7
SA17	R7C8	48	43	78 ± 101	-	78 ± 101	1.6 ± 2.1
SA18	R7C1Facility	8	5	13 ± 9	-	13 ± 9	1.6 ± 1.1
Weighted mean, everything		4358	3303	31667 ± 10039	*, †	26472 ± 9765	6.1 ± 2.2
No double counting		2516	1956		*, †, ††	12732 ± 4932	5.1 ± 2.0
Sum of clusters in R7		377	270			3828 ± 1199	10.2 ± 3.2
Only clusters with 100% fossil		1570	1238			6970 ± 2610	4.4 ± 1.7

* considering the absolute non-O&G emissions from the EDGAR inventory for the large regions and 100% O&G contribution for the clusters

† accounting for landfill within R6C6

†† excluding cluster quantifications in R7

S.5) Emission rate estimates for the ROMEO regions from EDGAR and TNO-CAMS inventories.

Table S7 - Emission rate estimates for non-O&G and O&G sources for the ROMEO regions using the EDGAR and TNO-CAMS inventories. The non-O&G emissions are aggregated emissions from all sectors other than O&G.

Region	EDGAR			TNO-CAMS		
	O&G emissions [kg/h]	Non-O&G emissions [kg/h]	Total emissions [kg/h]	O&G emissions [kg/h]	Non-O&G emissions [kg/h]	Total emissions [kg/h]
2	20	892	912	1001	799	1799
4	103	1365	1468	853	818	1671
5a	24	482	506	116	521	637
6	11	1307	1318	708	1094	1802
7	73	1461	1534	3112	1020	4133
8	7	951	959	177	420	597