# Supplementary material to Emissions of Carbon Tetrachloride ($CCl_4$) from Europe

## Dispersion model

We run the Lagrangian particle dispersion model FLEXPART v-9.02 (Stohl et al., 1998, 2005; http://www.flexpart.eu) releasing every three hours, from all measurement sites, 40 000 particles followed backward in time for 20 days. This in order to calculate the emission sensitivity footprint also called source-receptor-relationship (SRR). The SRR describes the relationship between the contribution of potential sources at the receptor and the change in mixing ratios at the measurement site. Fig.1S shows the SRR for the three single stations, highlighting how the two continental stations (CMN and JFJ) are influenced by air masses originating in central Europe, whereas MHD is predominately influenced by Atlantic/Arctic air masses.

Fig. 2S shows the yearly (2012) emissions sensitivity produced using the three measurement sites. We observe a good SRR in the whole European Geographic Domain (EGD), with the exception of a small region in the Aegean area.

All the simulations are driven by European Centre for Medium-range Weather Forecast wind fields using 3-hourly ERA-Interim reanalyses (analysis fields at 00:00, 06:00, 12:00 and 18:00 UTC, and 3-h forecasts at 03:00, 09:00, 15:00 and 21:00 UTC were used) with 1°x1° horizontal resolution and 91 vertical levels.

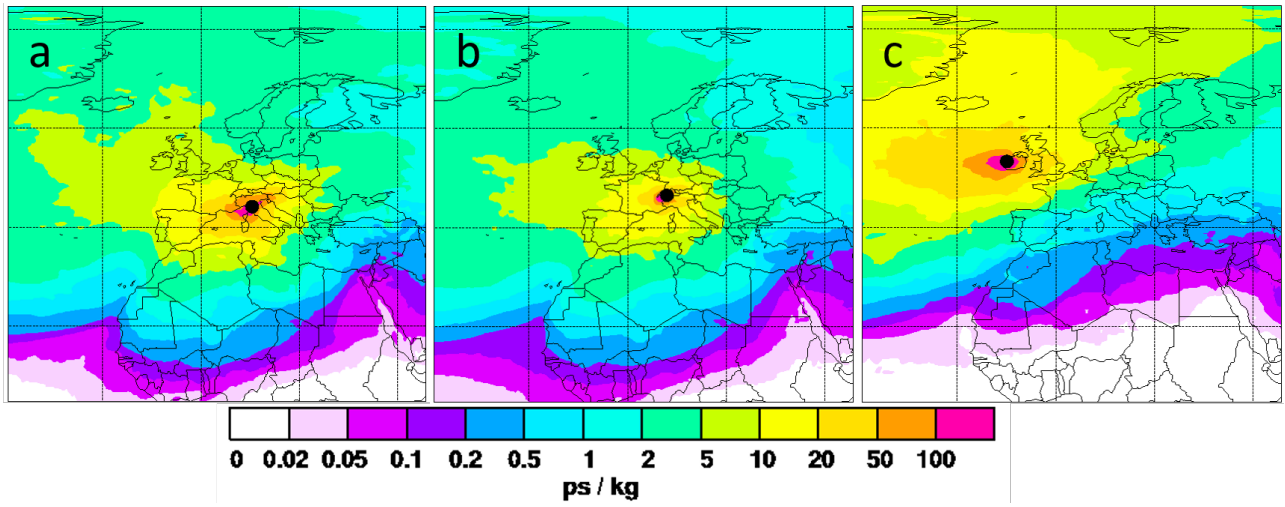


**Figure 1S. Single station SRR maps expressed in picoseconds per kilogram (ps $kg^{-1}$) obtained from FLEXPART 20 days backward calculations averaged over year 2012. Measurement sites are marked with black *dots*.**

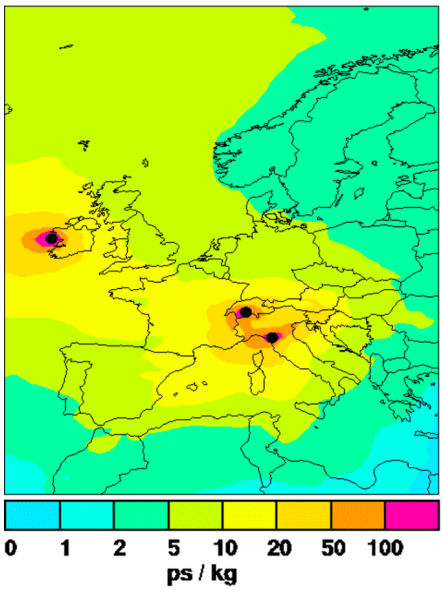


**Figure 2S. As in Figure 1S, but for the three stations.**

## Inversion method

To estimate the emissions of CTC from the EGD we used the inversion method, based on a Bayesian optimization technique, described by Stohl et al. (2009, 2010), where all mathematical details can be found. The emission distribution and intensity found by the inversion represent the best fit between observation data and model simulation. Using a limited number of stations not all regions are well constrained by the observations, making the problem ill-conditioned and unstable. Therefore, to get the solution to our problem, we used an a priori gridded field of emission distribution and the associated uncertainty (Stohl et al., 2009; 2010).
The cost function to be minimized is:

$$1)\quad J = (M\tilde{x} - \tilde{y})^T diag(\sigma_0^{-2})(M\tilde{x} - \tilde{y}) + \tilde{x}^T diag(\sigma_x^{-2})\tilde{x}$$

Where the matrix $M$ contains the model sensitivity, in our case all simulations produced by 40.000 particles run in backward mode for 20 days; the term $\tilde{x}$ represents the difference between the *a posteriori* and *a priori* emission vectors; $\tilde{y}$ is the difference between the observations and a priori simulated mixing ratios, $\sigma_0^{-2}$ is the vector of the standard error of observations, and $\sigma_x^{-2}$ is the *a priori* standard error vector.
Overall, the Bayesian inversion minimizes the cost-function reducing the model-observation misfit, represented by the first term on the right side of equation 1, optimizing the deviation of the solution from a priori emissions and its uncertainty, expressed by the second term of equation 1.

## Uncertainty evaluation

We associate for every grid cell an uncertainty value, $\sigma_x^j$

$$2)\quad \sigma_{x\_priori}^j = p * \max\,(k * x_j; l * x_{surf})$$

Where $p$ is an appropriate uncertainty scaling factor; $x_j$ the *a priori* emission value in grid $j$; $x_{surf}$ the average land surface emissions flux; $k$ and $l$ are scaling factors set at 0.5 and 1, respectively (Keller et al., 2011; Fang et al., 2014; Maione et al., 2014). The last term on the right side of equation 2 allows associating large uncertainty values even to low emission grid cells. We tested several uncertainty scaling factors $p$ in order to optimise the agreement between modelled and observed mixing ratios. The increase of the uncertainty scaling factor $p$ yields a higher variability of the *a posteriori* flux from the single grid cells, leading to a decreasing root mean square (RMS) and increasing correlation coefficients between modelled and observed mixing ratios in all the three stations. However, for $p$ values larger than 6, new hotspots emerge in the *a posteriori* emission field with unrealistically large emissions from low sensitivity regions. We used $p$ = 2, a value giving higher correlation coefficients and lower RMS values. Noteworthy, differences in the EGD emissions lower than 5 % are obtained using $p$ values ranging between 1.5 and 4. The minimisation of the cost function reduces the a priori sigma value $\sigma^j_{x\_priori}$ giving, for each inverted grid cell an uncertainty value $\sigma^j_{x\_posteriori}$. For the whole domain we obtained an average uncertainty $\sigma^j_{x\_posteriori}$≅30%, with a smaller uncertainty (≈20 %) in high sensitivity boxes close to the receptors (e.g., FR and UK) and a larger uncertainty (≈80 %) in low sensitivity regions far away from the receptors (e.g., Scandinavian region).

**E-PRTR database**

The *a priori* emission field used in this study makes use of the European Pollutant Release and Transfer Register (E-PRTR) inventory. E-PRTR is the Europe-wide register that provides data from industrial facilities in European Union Member States and in Iceland, Liechtenstein, Norway, Serbia and Switzerland. It replaced and improved upon the previous European Pollutant Emission Register (EPER).
The register contains data reported annually by more than 30,000 industrial facilities covering 65 economic activities across Europe. For each facility, information are provided concerning the amounts of pollutant releases to air, water and land as well as off-site transfers of waste and of pollutants in wastewater from a list of 91 key pollutants including CTC. The "industrial scale production of basic organic chemical" is the main CTC declared source in the database, being responsible for the 93.9 % of total European emissions. Figure 3S shows the percentage contribution to CTC emissions from each industrial facility, averaged over 2007-2013.
As reported in the paper main text (paragraph "Emission hot spots"), the inversion results estimate a CTC emission flux much larger than that declared in the E-PRTR. For major detail, we report In Fig. 4S the percent ratio between emissions reported in the E-PRTR and our estimates for each macro-area in the domain during 2007-2013. The E-PRTR reported emissions from the EGD represent on average, over the considered period, 4 % of the emissions obtained through the inversion. Lower discrepancies are found for the BE-NE-LUX and FR macro areas where the declared emissions reach the 43 % and 21% of inversion estimation, respectively.

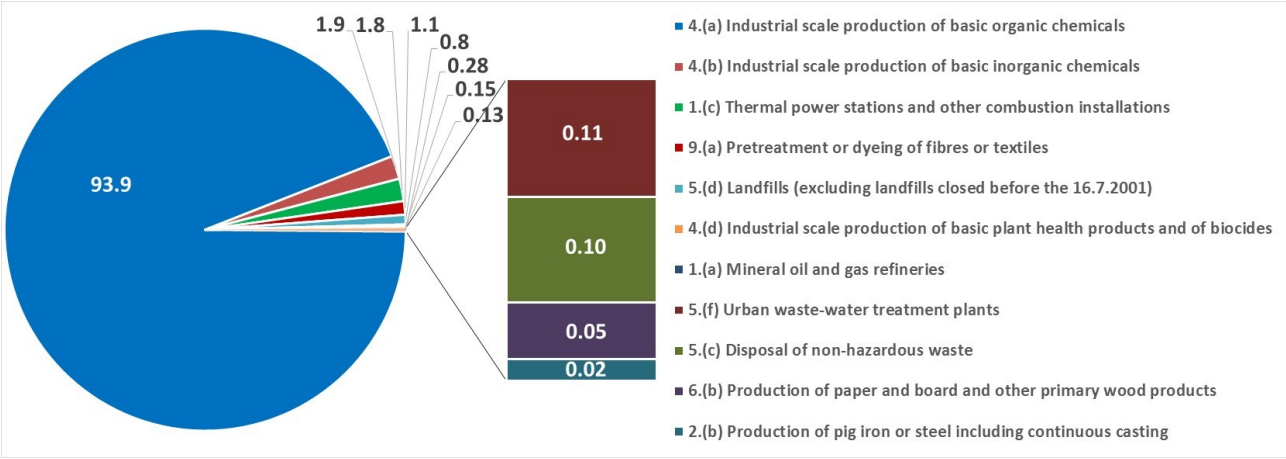


**Figure 3S. Average percentage contribution of different source sectors to the total CTC emissions reported in the E-PRTR (2007 -2013).**

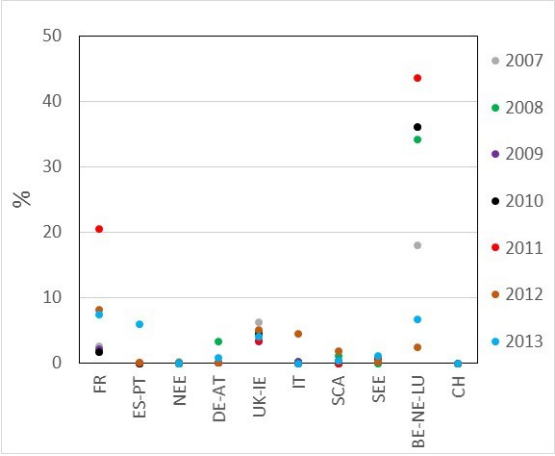


**Figure 4S. Percent ratio between emissions as in the E-PRTR and the inversion results for each macro-area in the domain during 2007-2013.**

## Subsets of data

Because of the limited numbers and localisation of the receptors, the simulations cannot produce a homogeneous sensitivity over the study domain. In order to assess to what extent our results are sensitive to the receptors used, we run the inversions removing one station at a time

The EGD emissions obtained with different subsets of observation data are consistent with those obtained using the full set. The larger difference, 26 %, is registered when removing MHD. Removing JFJ and CMN, produced a similar percentage difference of -10 % and -9 %, respectively, as a consequence of the similar footprint of the two receptors. This result indicates the stability of the inversion system even when using a subset of data and reinforce the benefit of the increased sensitivity over domain when using an increasing number of receptors.

## Model performance at the stations

With the aim of evaluating the model performance and the station specific errors, we compared the observed and modelled time series at the three stations, taking into account different statistical parameters, in a similar way as described in Stohl et al. (2009), Maione et al. (2014) and Graziosi et al. (2015). The results of this comparison, carried out for the year 2012, are reported in Table 1S.

$1 - E_b/E_a$ is the relative error reduction, where $E_a$ and $E_b$ are the *a priori* and *a posteriori* RMS errors. The values achieved at the stations used in this study are in a range between 16 % and 23 %, in spite of the different station characteristics.

The Pearson correlation coefficients described in the following show a better performance for MHD because of the poorer model performance in the mountain area. However, as stated in Mahowald et al. (1997), using receptors closer to the main source regions would improve the model performance to acquire source information.

$r_a^2$ is the squared Pearson correlation coefficients between the time series obtained at receptor using the *a priori* emission field and the observed time series, and $r_b^2$ between the *a posteriori* and observed time series. These coefficients are used to evaluate the proximity of the modelled emission field to the real one. The obtained $r_b^2$ values higher than $r_a^2$ are an indication of the improvement of the *a posteriori* emission field with respect to the *a priori*.

Analogously to $r_a^2$ and $r_b^2$, the squared Pearson correlation coefficients $r_{ba}^2$ (and $r_{bb}^2$) between the modelled *a priori* (and *a posteriori*) and the measured baseline mixing ratios at the three stations indicate the capability of the system to reproduce the variability and trends of the baseline.

Transport events from the source regions to the receptors generate the variability in the observed enhancements above the baseline. The correlation analyses between the observed and simulated *a priori* ($r_{ea}^2$) and the *a posteriori* ($r_{eb}^2$) polluted mixing ratios describe the system capability to reproduce concentrations above the background. Higher correlation values are obtained at the remote station of MHD. Despite the relatively low $r_{ea}^2$ and $r_{eb}^2$ values at CMN and JFJ, data from these two mountain stations improve the inversions on the regional scale, thanks to the station sensitivity to the main source regions. For the same reason, the two mountain stations also present higher standard deviation (SD) of the observed mixing ratios.

**Table 1S. Station parameters. Mean, mean CTC mixing ratios; SD, standard deviation of the observed mixing ratios; N, number of observations; $E_a$, RMS *a priori* error; $E_b$, RMS *a posteriori* error; $1-E_a/E_b$, relative error reduction; $r^2_a$ and $r^2_b$, squared Pearson correlation coefficients between the observations and the *a priori* ($r^2_a$) and *a posteriori* ($r^2_b$) simulated time series; $r^2_{ba}$ (and $r^2_{bb}$) is the squared Pearson correlation coefficients between the *a priori* (and *a posteriori*) baseline and the measured concentrations; $r^2_{ea}$ (and $r^2_{eb}$) is the squared Pearson correlation coefficients between the *a priori* (and *a posteriori*) enhancements above the baseline and the measured concentrations.**

| Station | Mean (ppt) | SD (ppt) | N | $E_a$ (ppt) | $E_b$ (ppt) | $1-E_b/E_a$ | $r^2_a$ | $r^2_b$ | $r^2_{ba}$ | $r^2_{bb}$ | $r^2_{ea}$ | $r^2_{eb}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CMN | 85.7 | 1.1 | 2039 | 0.84 | 0.65 | 0.23 | 0.48 | 0.58 | 0.51 | 0.55 | 0.23 | 0.32 |
| JFJ | 84.7 | 0.8 | 2124 | 1.15 | 0.94 | 0.19 | 0.35 | 0.44 | 0.24 | 0.26 | 0.19 | 0.22 |
| MHD | 84.9 | 0.3 | 2833 | 0.64 | 0.50 | 0.21 | 0.70 | 0.79 | 0.73 | 0.75 | 0.54 | 0.63 |

## Relevance of the inclusion of the E-PRTR database in the a priori emission field

As reported in the main text, the best model performances have been obtained using for the *a priori* the European emission values estimated by Xiao et al. (2010) combined with the industrial emissions declared in the E-PRTR database. We refer to such *a priori* field as the *reference a priori* field.

To investigate the relevance of the inclusion of the E-PRTR database in the *a priori* emission field, we compared the squared Pearson correlation coefficients between observations and the *a priori*

emission field ($r_a^2$) and *a posteriori* ($r_b^2$) model results, with those obtained using an alternative *a priori* field that does not include the emission fluxes reported by E-PRTR.

Figure 5S shows $r_a^2$ and $r_b^2$ values of all stations derived using both the *reference* and the *alternative a priori* emission fields ($r^2_{a\_reference}$ $r^2_{b\_reference}$; *and* $r^2_{a\_alternetive}$ and $r^2_{b\_alternative}$, respectively). The comparison between the correlation results obtained using the two different *a priori* emission fields highlighted how the inclusion of the E-PRTR information improved the performance of the inversion. This despite the finding, reported in the main text, that the industry declared emissions are systemically lower than the emissions obtained by the inversion.

Due to the inversion model skills, smaller differences are achieved between the two *a posteriori* correlation parameters $r_b^2$ comparing the two *a priori* correlation values $r_a^2$ (fig 5S). An EGD emission difference of only 5 % between the two different *a posteriori* fields is achieved. This thanks to the inversions capability of generating similar a posteriori emissions starting from different *a priori* values.

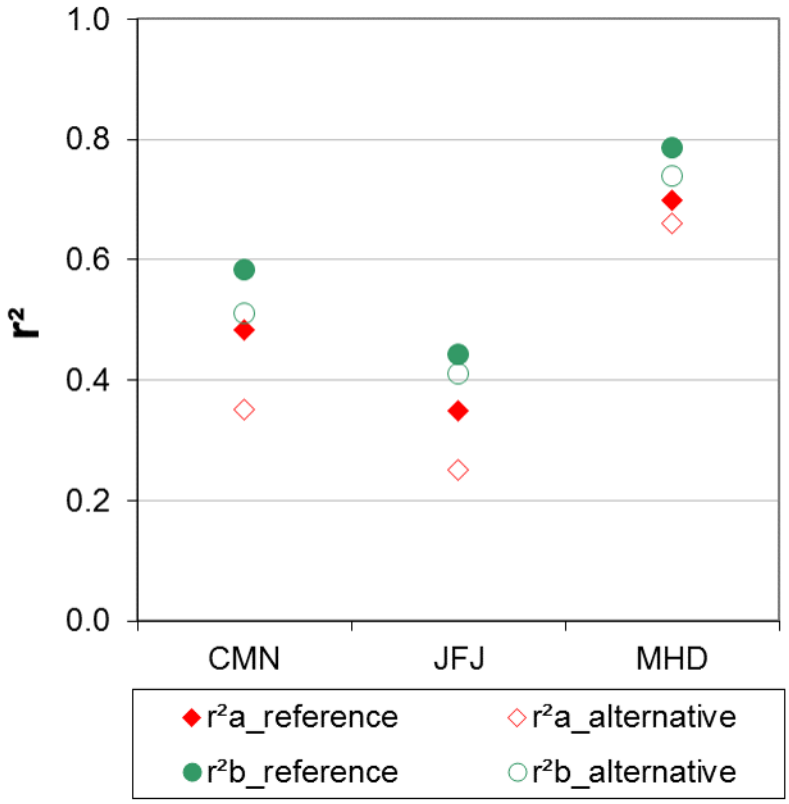


**Figure 5S. Inversion model performance using the "reference" and "alternative" *a priori* emission field at the three stations.**

## References

Fang X. K., Thompson R. L., Saito T., Yokouchi Y., Kim J., Li S., Kim K. R., Park S., Graziosi F., Stohl, A.: Sulfur hexafluoride (SF6) emissions in East Asia determined by inverse modelling, Atmos. Chem. Phys., 14, 4779-4791, 2014.

Graziosi, F., Arduini, J., Furlani, F., Giostra, U., Kuijpers, L. J. M., Montzka, S. A., Miller, B. R., O'Doherty, S. J., Stohl, A., Bonasoni, P., and Maione, M.: European emissions of HCFC-22 based on eleven years of high frequency atmospheric measurements and a Bayesian inversion method, Atmos. Environ., 112, 196, 2015.

Keller, C. A., Brunner, D., Henne, S., Vollmer, M. K., O'Doherty, S., and Reimann, S.: Evidence for under-reported western European emissions of the potent greenhouse gas HFC-23, Geophys. Res. Lett., 38, L15808, doi:10.1029/2011gl047976, 2011.

Mahowald, N. M., Prinn, R. G., and Rasch, P. J.: Deducing $CCl_3F$ emissions using an inverse method and chemical transport models with assimilated winds, J. Geophys. Res., 102, 28153–28168, 1997.

Maione M., Graziosi, F., Arduini, J., Furlani, F., Giostra, U., Blake, D.R., Bonasoni, P., Fang, X., Montzka, S.A., O'Doherty, S.J., Reimann, S., Stohl, A., and Vollmer, M.K.: Estimates of European emissions of methyl chloroform using a Bayesian inversion method, Atmos. Chem. and Phys. 14, 9755-9770, doi:10.5194/acp-14-9755-2014, 2014.

Stohl, A., Hittenberger, M., and Wotawa, G.: Validation of the Lagrangian particle dispersion model FLEXPART against large scale tracer experiment data, Atmos. Environ., 32, 4245–4264, 1998.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmos. Chem. Phys., 5(9), 2461–2474, 2005.

Stohl, A., Seibert, P., Arduini,J., Eckhardt, S., Fraser, P., Greally, B. R., Lunder, C., Maione, M., Mühle, J., O'Doherty ,S., Prinn, R. G., Reimann, S., Saito, T., Schmidbauer, N., Simmonds, P. G., Vollmer, M. K., Weiss, R. F., and Yokouchi, Y.: An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons, Atmos. Chem. and Phys. 9, 1597-1620, 2009.

Xiao, X., Prinn, R. G., Fraser, P. J., Weiss, R. F., Simmonds, P. G., O'Doherty, S., Miller, B. R., Salameh, P. K., Harth, C. M., Krummel, P. B., Golombek, A., Porter, L. W., Butler, J. H., Elkins, J. W., Dutton, G. S., Hall, B. D., Steele, L. P., Wang, R. H. J., and Cunnold, D. M.: Atmospheric three-dimensional inverse modelling of regional industrial emissions and global oceanic uptake of carbon tetrachloride, Atmos. Chem. Phys., 10, 10421-10434, doi: 10.5194/acp-10-10421-2010, 2010.