Supporting Information

Toward resolving the mysterious budget discrepancy of ozone-depleting CCl_4: An analysis of top-down emissions from China

Sunyoung Park[1,2], Shanlan Li[2,3], Jens Mühle[4], Simon O'Doherty[5], Ray F. Weiss[4], Xuekun Fang[6], Stefan Reimann[7], Ronald G. Prinn[6]

[1]Department of Oceanography Kyungpook National University, Daegu 41566, Republic of Korea
[2]Kyungpook Institute of Oceanography Kyungpook National University, Daegu 41566, Republic of Korea
[3]Climate Research Division, National Institute of Meteorological Sciences, Seogwipo, Korea
[4]Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093, USA
[5]School of Chemistry, University of Bristol, Bristol, UK
[6]Center for Global Change Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[7]Empa, Laboratory for Air Pollution and Environmental Technology, Swiss Federal Laboratories for Materials Science and Technology, Überlandstrasse 129, 8600 Dübendorf, Switzerland

Correspondence to: Sunyoung Park (sparky@knu.ac.kr)

Trajectory residence time

Residence time trajectory analyses have been used extensively to identify the source locations and preferred transport pathways of atmospheric trace elements and particulate species (Ashbaugh et al., 1985). Residence times are calculated by the following equation:

$$\tau_{abk} = \sum_{k=1}^{n} \sum_{h=1}^{l} (S_{abkh} / v_{kh}) \qquad \text{S(1)}$$

τ_{abk} is the total residence time for all trajectories over grid cell *a, b*; S_{abkh} is the length of that portion of the h^{th} segment of the k^{th} trajectory over the grid cell *a, b*; and v_{kn} is the average speed of the air parcel as it travels along the h^{th} segment of the k^{th} trajectory.

The residence time analysis shown in Fig. 1S suggests that the major air masses arriving at Gosan station (GSN) vary seasonally, with predominant northwesterly and northeasterly continental outflows from fall through spring versus flows of clean air directly from the Pacific in summer and from northern Siberia in winter.

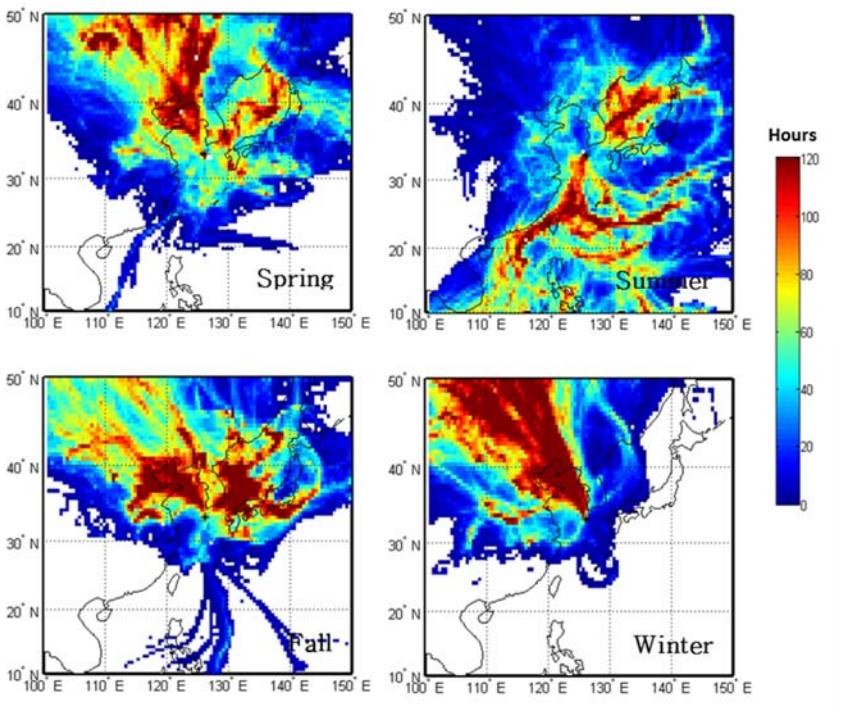

Fig. S1. Residence time analysis for 2008 using 6-day back-trajectories arriving at the Gosan station. Seasonal residence time distributions show a distinctive seasonally-varying wind pattern.

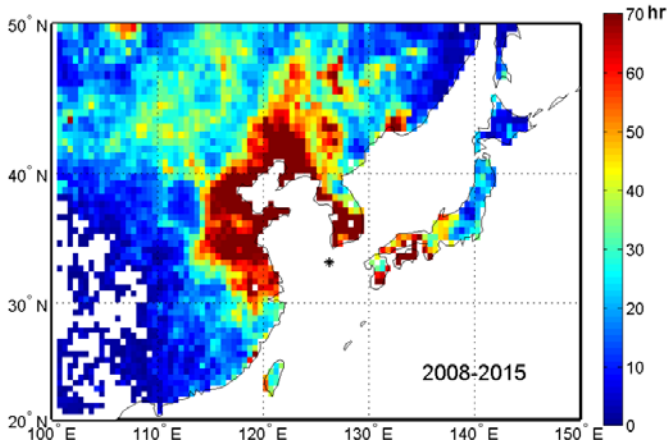

Fig. S2. Distribution of averaged residence times of air masses arriving at Gosan for the years 2008–2015. Residence times of over 24 h occurred over both northeastern continental regions and the central southern part of China. The asterisk denotes the Gosan measurement station.

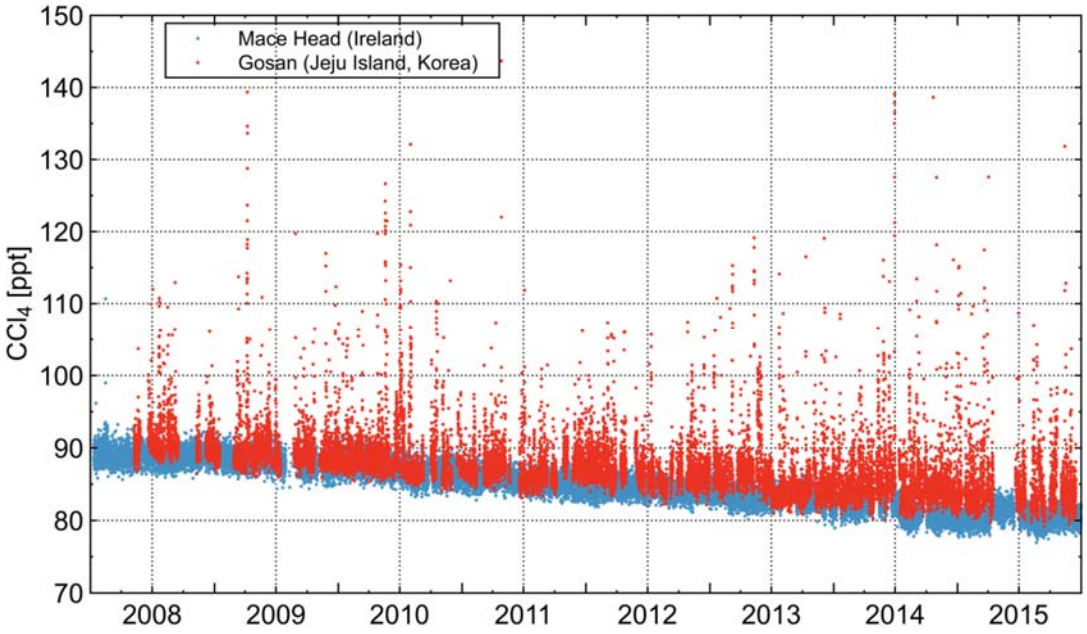

Fig. S3. The 8-year observation records for CCl_4 analyzed in this study are shown as red points (also shown in Fig. 1). For comparison, the corresponding observations taken at the Mace Head station (53°N, 10°W) in Ireland are represented by blue points. Note that the "background" concentrations from GSN agree well with the baseline values at Mace Head station, a background station in the Northern Hemisphere, and are declining at a similar rate to its global trend.

Trajectory Statistics

To identify the potential CCl_4 source regions, we applied a statistical analysis coupled with back trajectories to the time series of observed enhancements in CCl_4 concentrations from 2008 to 2015. The trajectory statistics method has often been applied to estimate the potential source areas of air pollutants (Reimann et al. 2004). The underlying assumption of the trajectory statistics method is that elevated concentrations at an observation site are proportionally related to both the average concentrations in a specific grid cell over which the observed air mass has passed and the residence time of the air mass over that grid cell. Thus, the method simply computes a residence-time-weighted mean concentration for each grid cell by superimposing the back-trajectory domain on the grid matrix. The formula is given by:

$$\overline{C}_{ab} = \frac{\sum_{i=1}^{n} (\tau_{abi} C_i)}{\sum_{i=1}^{n} \tau_{abi}} \qquad \text{S(2)}$$

where C_i is the enhanced concentration of CCl_4 at a given i^{th} time; τ_{abi} is the residence time of the trajectory arriving at Gosan at the i^{th} time, spent over the grid cell *a, b* (in 0.5°x 0.5°) within the atmospheric boundary layer; and $\overline{C}_{ab}$ represents the relative strength of the cell *a, b* as a potential source region of CCl_4 source. Back trajectories were calculated using the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model of the NOAA Air Resources Laboratory (ARL) using meteorological information from the Global Data Assimilation System (GDAS) model with 1°×1° grid cell. The HYSPLIT model was run using 6-day backward trajectories at 500-m altitude above the measurement site. The residence times were calculated by the method of Poirot and Wishinski (1986). To eliminate low confidence level areas, a point filter was applied that removed grid cells over which less than 12 trajectories had passed (Reimann et al. 2004).

This trajectory statistics method can also be applied to illustrate the potential location of each source factor determined from the PMF analysis. The formula is identical to Eq. S(2) in all respects except that it uses the normalized strength of each source factor. The enhanced concentrations from the j^{th} source contribute to the observation at the k^{th} time (which is denoted as "f_{jk}" of Eq. (1) in main text). Since the f_{jk} values from all eight sources cover a very wide range of concentrations, the f_{jk} values can be normalized against their time average for the j^{th} source in order not to bias the statistical significance of one source against the others. Therefore, the normalized time series of f_{jk} values were defined as

$$mjk = fjk / (\sum_{k=1}^{n} fjk) / n \qquad \text{S(3)}$$

Then, Eq. S(2) was modified to the following:

$$\overline{m}_{abj} = \frac{\sum_{k=1}^{n} (\tau_{abk} m_{jk})}{\sum_{k=1}^{n} \tau_{abk}} \qquad \text{S(4)}$$

where m_{jk} is the normalized strength of the j^{th} source at a given k^{th} time; τ_{abk} is the residence time of the trajectory arriving at Gosan at the k^{th} time, spent over the grid cell *a, b* (in 0.5°x 0.5°) within the atmospheric boundary layer; and $\overline{m}_{abj}$ represent a relative strength of the cell *a, b* as a potential source region of the j^{th} source.

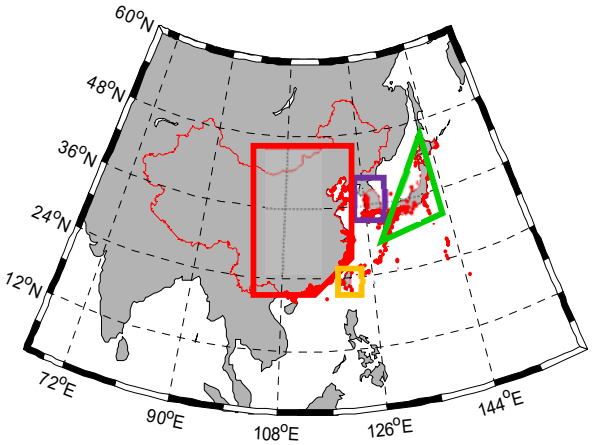

Fig. S4(a). Trajectory attribution: four country domains defined to separate country-specific pollution signals from the original observations. The Chinese domain is defined as a regional grid of 100–124° E and 21–45° N.

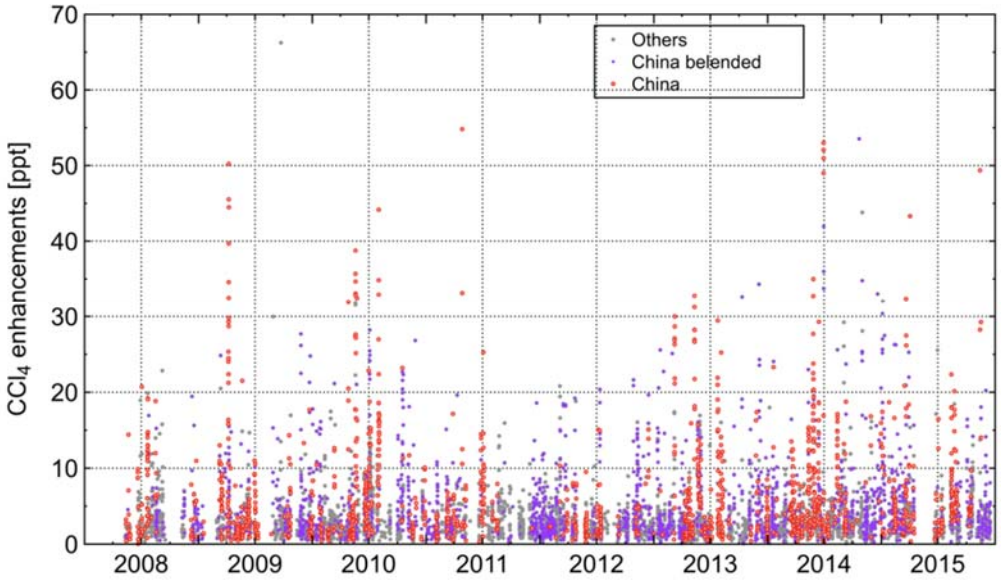

Fig. S4(b). CCl_4 pollution events in 2008–2015 classified according to origin. Air masses from China are shown in red. The purple dots represent blended air masses affected both by China and other countries. Together, these two groups explain about 75 % of the observed pollution data in 2008–2015. The remaining 25 % are shown as gray dots.

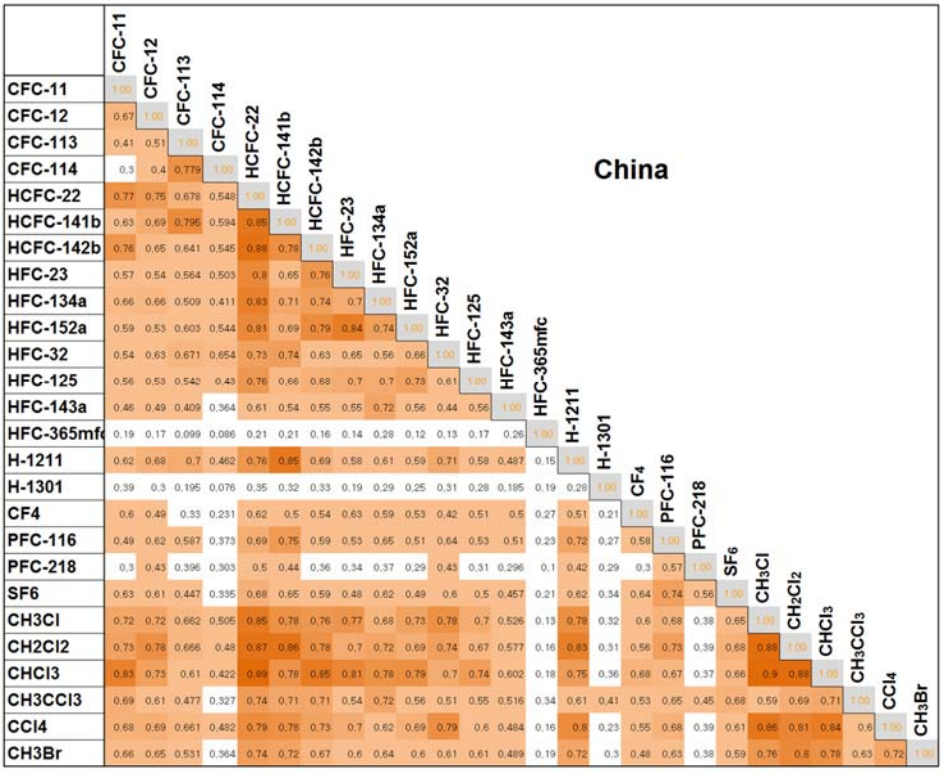

Fig. S5. The observed relationships of CCl_4 vs. 26 halocarbons for air masses originating from China. The CCl_4: HCFC-22 ratio (0.13 ppt/ppt) has one of the most significant correlation coefficient ($R^2 = 0.79$, $p < 0.01$) among the calculated 25-member correlation matrix.

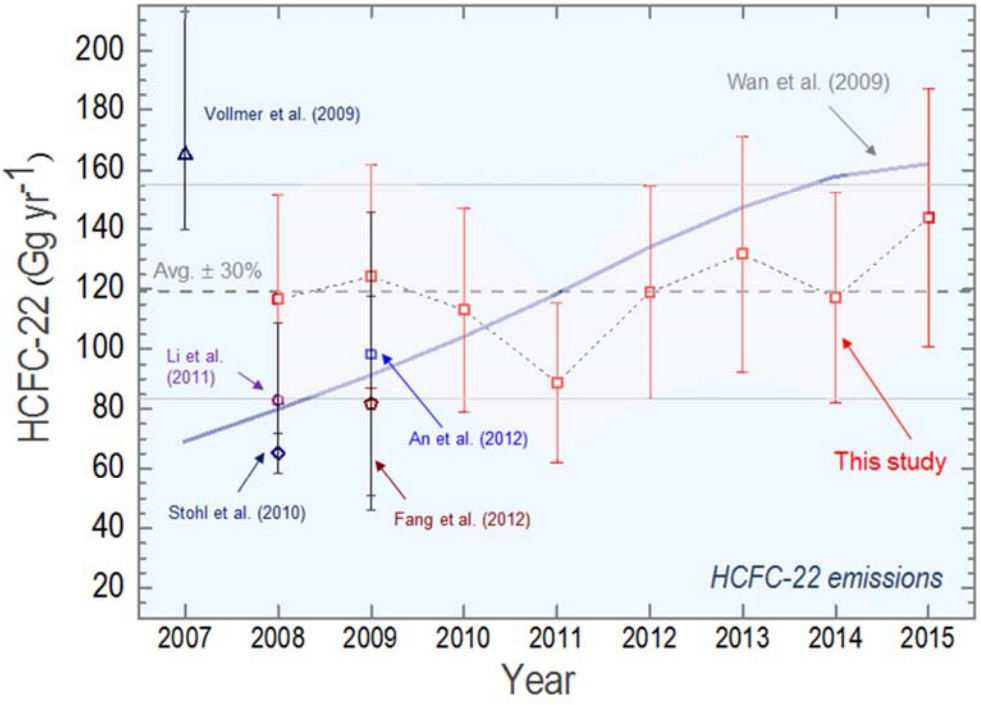

Fig. S6. Annual HCFC-22 emissions in China for 2008–2015 derived from atmospheric measurements data from Gosan station using an inverse technique based on a Lagrangian transport model analysis. The red error bars denote the estimation uncertainty of 30 %. The dashed and solid gray lines represent the average and its 30 % uncertainty ranges for the HCFC-22 emissions for 2008–2015. The estimates are very consistent overall with previous top-down studies and a bottom-up estimate.

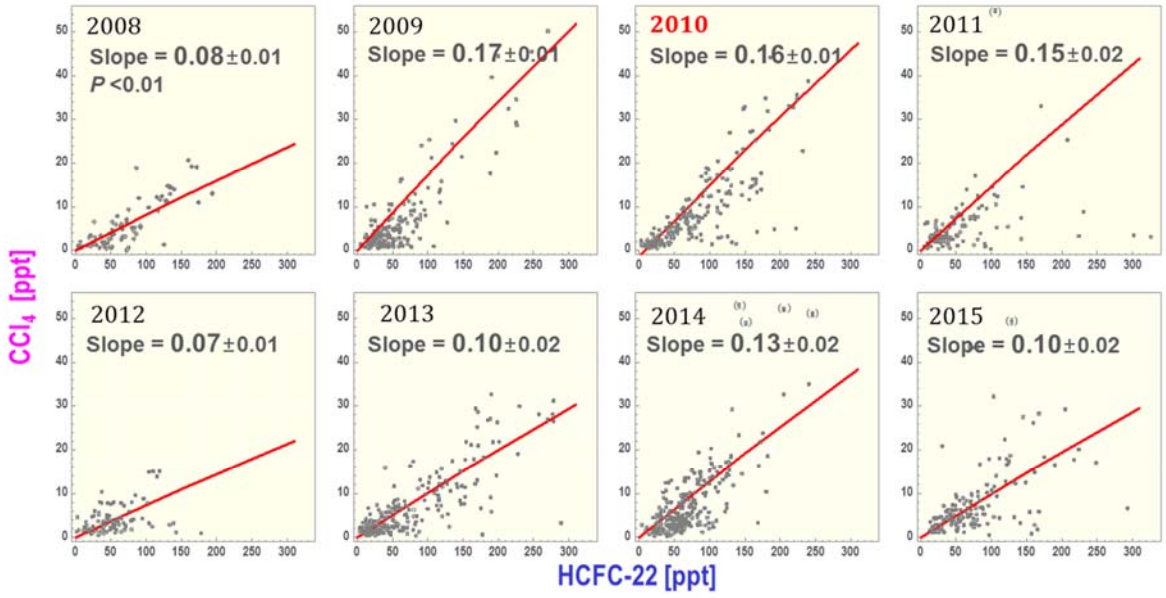

Fig. S7. The annual slopes of the empirical correlations between observed enhancements of CCl_4 vs. HCFC-22 (ΔCCl_4 vs. ΔHCFC-22). The slopes and uncertainties were calculated using a Williamson-York linear least-squares fitting method.

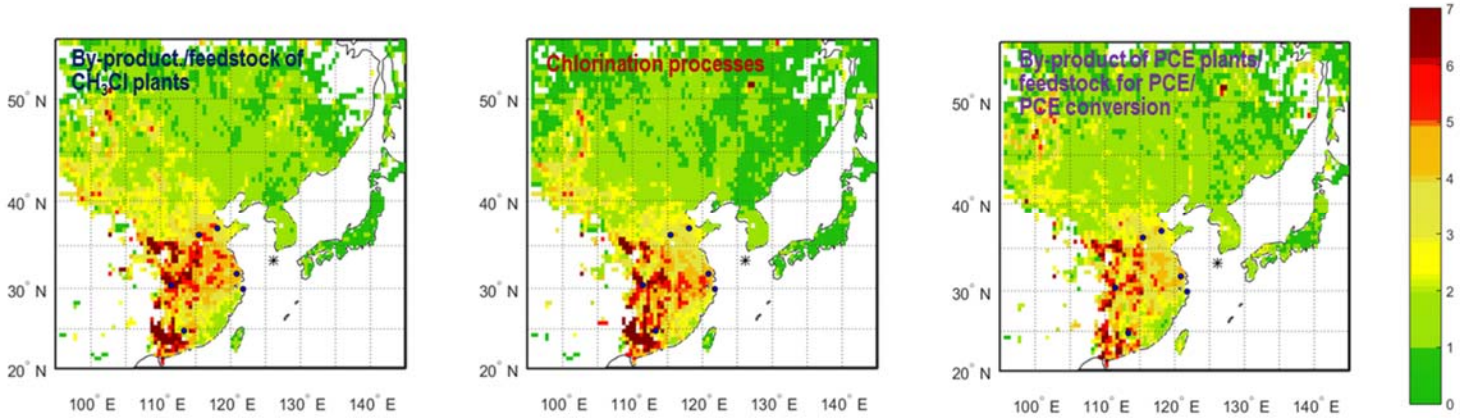

Fig. S8. Potential source region distributions of the three emission sources accounting for 89±5 % of CCl_4 enhancements observed at Gosan. The areas in and around Guangzhou of Guangdong, Wuhan of Hubei, Zhengzhou of Henan, and Xian of Shaanxi province were identified as the dominant contributors. The six blue dots indicate the locations of the main factories producing HFCs, HCFC-22 and fluorocarbons, which are given in http://eng.chinaiol.com/.

Positive Matrix Factorization Model Calculation

The PMF optimization uses a weighted least squares regression to obtain a best fit to the measured enhancements in the concentration data. Because the main constraints that need to be resolved during the analysis are "source factors", this is often called a factor analysis. The mathematical expression of the model is given by Eq. (1):

$$x_{ik} = \sum_{j=1}^{p} g_{ij} f_{jk} + e_{ik} (i = 1,2,..., m; j = 1,2,..., p; k = 1,2,..., n) \quad (1)$$

where x_{ik} represents enhanced concentrations in the time series of the "i" halogenated compound at the k^{th} sampling time; g_{ij} is the concentration fraction of the i^{th} compound from the j^{th} source; f_{jk} is the enhanced concentration from the j^{th} source contributing to the observation at the k^{th} time, which is given in ppt; e_{ik} is the model residual for i^{th} compound concentration measured in the k^{th} sampling time; and p is the total number of independent sources (i.e., the number of factors) (Paatero and Tapper, 1994). The optimal number of factors (p) should be determined by using a function Q, defined in Eq. S(5) below:

$$Q = \sum_{i=1}^{m} \sum_{k=1}^{n} \left(\frac{e_{ik}}{u_{ik} h_{ik}}\right)^2 \quad \text{S(5)}$$

where u_{ik} are the uncertainties corresponding to each measurement data point. Following the guideline provided by Polissar et al. (1998) for PMF model input uncertainties, we took into account the instrumental measurement uncertainty, monthly standard deviation (1σ) of background concentration, and 1/3 of the detection limit value as the overall uncertainty assigned to each data point. The PMF model input uncertainties (in ppt) were constructed as follows:

$$u_{ik} = \sqrt{\mu_{ik}^2 + \sigma_{ik}^2} + d_{ik}/3 \quad \text{S(6)}$$

where μ_{ik} is measurement uncertainty; σ_{ik} is the monthly standard deviation of the background; and d_{ik} is the analytical detection limit. The average values of these individual input error terms are listed for all species in Table S1. In Eq. S(5) h_{ik}=1 if $|e_{ik}/u_{ik}| < \alpha$, and otherwise h_{ik} is defined as $|e_{ik}/u_{ik}|/\alpha$. The α is the outlier threshold distance parameter. Appropriate down weighting of outliers in PMF datasets has been carried out in many studies (Polissar et al., 1998; Lee et al., 1999; Lee et al., 2002) using this parameter to reduce the influence of outliers and extreme values. We constrain the PMF analysis with α=4, which is most commonly used. In other words, when the scaled residual exceeded four times the standard deviation, the uncertainty, u_{ik}, was increased to down-weight that concentration.

The model runs with randomly selected initial values for f and g at a given number of factors (p) (varied from 5 to 10 factors) to obtain a minimum Q value in less than 20 iterations (Lee et al., 1999). As the number of factors increases, the corresponding minimum Q values decreases, with a level-off in this case near 7 factors. We carefully examined the solutions with 7, 8, and 9 factors and determined an optimal value based on both goodness of fit to the data and prior knowledge about halogenated compound emissions. The model's goodness-of-fit was estimated from a correlation plot between the measured and model-predicted concentrations. Most of the compounds (16 out of 18 species) showed good correlations ($R^2 > 0.6$, see Table S2) for the eight-factor solution. Another way to assess a PMF fit is to examine the distribution of scaled residuals (e_{ik}/u_{ik}). We found the most species except COS lie within ±4, which is considered a typical limit.

The seven-factor model cannot separate the foam-blowing-agent factor from the semiconductor/electronics sector factor, which are both well-known sources of halogenated compounds. For the nine-factor analysis, the sources for CH_2Cl_2 and $CHCl_3$ were split. Therefore, we concluded that an eight–source model provides the most relevant and meaningful interpretation for the enhanced CCl_4 concentrations observed at Gosan.

PMF input uncertainties

The uncertainties (in ppt) imposed on individual concentrations are typically determined typically in the PMF community as follows:

$$u_{ik} = \sqrt{\mu_{ik}^2 + \sigma_{ik}^2} + d_{ik}/3 \qquad \text{S(7)}$$

where μ_{ik} is the measurement uncertainty; σ_{ik} is the monthly standard deviation (1σ) of the background; and d_{ik} is analytical detection limit. The average values of these individual input error terms for all species are listed in Table S1.

Table S1. Three individual input error terms and their average values for all species.

Compounds	Analytical precision (ppt)	Background uncertainty (ppt)	Detection limit (ppt)
CFC-11	0.44	1.40	0.72
CFC-12	0.61	0.82	1.33
HCFC-22	0.54	1.70	1.47
HCFC-141b	0.10	0.70	0.22
HCFC-142b	0.07	0.55	0.12
HFC-23	0.13	0.26	0.26
HFC-134a	0.18	1.10	0.47
HFC-152a	0.08	0.53	0.15
HFC-32	0.13	0.29	0.28
HFC-125	0.05	0.20	0.11
HFC-143a	0.09	0.22	0.19
CF_4	0.09	0.19	0.20
C_2F_6	0.03	0.04	0.06
C_3F_8	0.01	0.02	0.03
SF_6	0.03	0.08	0.07
CH_3Cl	1.09	11.00	2.36
CH_2Cl_2	1.55	7.20	3.36
$CHCl_3$	0.17	1.40	0.74
CH_3Br	0.05	0.38	0.11
CCl_4	0.80	1.09	1.76
COS	2.92	14.00	6.36
PCE	0.02	0.42	0.04

Goodness of the PMF model fit

The goodness of fit of the PMF model can be assessed by comparing the predicted compound concentrations with the original measurements. We found R-squared of larger than 0.6 for most of the halogenated compounds, as shown in Table S2.

Table S2. Goodness-of-fit statistics for the plot of the observed concentrations versus the PMF model estimates at the number of factors (p) = 8.

CFCs and $HCFC_s$			HFC_s			PFC_s and SF_6			Others		
Compounds	R^2	p valve	Compounds	R^2	p valve	Compounds	R^2	p valve	Compounds	R^2	p valve
CFC-11	0.58	<0.01	HFC-23	0.70	<0.01	CF_4	0.67	<0.01	CCl_4	0.76	<0.01
HCFC-22	0.78	<0.01	HFC-134a	0.68	<0.01	C_2F_6	0.55	<0.01	$CHCl_3$	0.77	<0.01
HCFC-141b	0.74	<0.01	HFC-143a	0.27	<0.01	SF_6	0.81	<0.01	CH_2Cl_2	0.95	<0.01
HCFC-142b	0.75	<0.01	HFC-32	0.88	<0.01				CH_3Cl	0.99	<0.01
			HFC-125	0.86	<0.01				C_2Cl_4	0.99	<0.01
									COS	0.99	<0.01

Description of the PMF source factors

The sixth factor shown in Fig. 4 in the main text, interpreted as arising from refrigerant consumption, explains approximately 80±2 % of the HCFC-22 and 32±4 % of the HFC-134a enhancements observed. HCFC-22 and HFC-134a are the most abundant species in the HCFC and HFC families, respectively, showing increasing use in refrigeration units and air conditioning systems as CFCs replacements (Montzka et al., 2011). Many species contribute significantly to the seventh factor, in particular, 88±20 % for SF_6, 41±3 % for C_2F_6, and 40±13 % for CF_4. SF_6 is widely used in the high-voltage electrical equipment sector as a gaseous dielectric medium and is also used as an etching/cleaning agent in the semiconductor/electronics sector (Forster et al., 2007). Recently, use of PFCs (CF_4 and C_2F_6 foremost among them) for plasma etching and chamber cleaning in semiconductor/electronics manufacturing processes has been increasing (Mühle et al., 2010). Thus, the large contributions of SF6 and PFCs suggest that this source factor is related to processes in the semiconductor/electronics industry. The last factor shown in Fig. 4 is composed of 92±4 % HCFC-142b, the most widely used CFC replacement for foam blowing agents for extruded polystyrene boards (Derwent et al., 2007). The foam blowing factor also explains 23±2 % of CFC-11, indicating that this CFC is still emitted from remaining bank use or old building materials.

Chlorination reactions for CCl_4 production and use

CH_3Cl/ CCl_4 plants

$CH_3OH + HCl \rightarrow \mathbf{CH_3Cl} + H_2O$

$CH_3Cl + Cl_2 \rightarrow \mathbf{CH_2Cl_2} + HCl$

$CH_2Cl_2 + Cl_2 \rightarrow \mathbf{CHCl_3} + HCl$

$CHCl_3 + Cl_2 \rightarrow \mathbf{CCl_4} + HCl$

Feedstock for production of chloromethanes and PCE

$\mathbf{CCl_4} + H_2 \rightarrow CHCl_3 + HCl$

$3CCl_4 + CH_4 \rightarrow 4CHCl_3$

$CCl_4 + H_2 \rightarrow CHCl_3 + HCl$

$2CCl_4 + 2H_2 \rightarrow \mathbf{C_2Cl_4} + 4HCl$

$CCl_4 + 4Cl_2 + CH_4 \rightarrow \mathbf{C_2Cl_4} + 4HCl$

$2CCl_4 + H_2 \rightarrow C_2Cl_6 + 2HCl$

$CCl_4 + HF \rightarrow CCl_3F$[**CFC-11**] + HCl

References

Ashbaugh, L.L., Malm, W.C., Sadeh, W.Z.: A residence time probability analysis of sulfur concentrations at Grand Canyon National Park., Atmos. Environ., 19, 8, 1263–1270, 1985, https://doi.org/10.1016/0004-6981(85)90256-2, 1985.

Derwent, R.G., Simmonds, P.G., Greally, B.R., O'Doherty, S., McCulloch, A., Manning, A., Reimann, S., Folini, D., Vollmer, M.K.: The phase-in and phase-out of European emissions of HCFC-141b and HCFC-142b under the Montreal Protocol: Evidence from observations at Mace Head, Ireland and Jungfraujoch, Switzerland from 1994 to 2004, Atmos. Environ., 41, 4, 757–767, https://doi.org/10.1016/j.atmosenv.2006.09.009, 2007.

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D.W., Haywood, J., Lean, J., Lowe, D.C., Myhre, G., Nganga, J., Prinn, R., Raga, G., Schulz, M., Van Dorland, R.: Changes in Atmospheric Constituents and in Radiative Forcing, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L (Eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Lee, E., Chan, C.K., Paatero, P.: Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong, Atmos. Environ., 33, 19, 3201–3212, https://doi.org/10.1016/S1352-2310(99)00113-2, 1999.

Lee, J.H., Yoshida, Y., Turpin, B.J., Hopke, P.K., Poirot, R.L., Lioy, P.J., Oxley, J.C.: Identification of sources contributing to Mid-Atlantic regional aerosol, J. Air & Waste Manage. Assoc., 52, 10, 1186–1205, https://doi.org/10.1080/10473289.2002.10470850, 2002.

Montzka S.A., Reimann, S. (Eds): Ozone-depleting substances (ODSs) and related chemicals, in:

Scientific Assessment of Ozone Depletion 2010, World Meteorol Org, Geneva, 1−108., 2010.

Mühle, J., Ganesan, A. L., Miller, B. R., Salameh, P. K., Harth, C. M., Greally, B. R., Rigby, M., Porter, L. W., Steele, L. P., Trudinger, C. M., Krummel, P. B., O'Doherty, S., Fraser, P. J., Simmonds, P. G., Prinn, R. G., and Weiss, R. F.: Perfluorocarbons in the global atmosphere: tetrafluoromethane, hexafluoroethane, and octafluoropropane, Atmos. Chem. Phys., 10, 5145–5164, https://doi.org/10.5194/acp-10-5145-2010, 2010.

Paatero, P., Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, Environ., 5, 111-126, https://doi.org/10.1002/env.3170050203, 1994.

Poirot, R.L., Wishinski, P.R.: Visibility, sulfate and air-mass history associated with the summertime aerosol in northern Vermont, Atmos. Environ, 20, 7, 1457−1469, https://doi.org/10.1016/0004-6981(86)90018-1, 1986.

Polissar, A.V., Hopke, P.K., Paatero, P., Malm, W.C., Sisler, J.F.: Atmospheric aerosol over Alaska: 2. Elemental composition and sources, J. Geophys. Res., 103, D15, 19045–19057, https://doi.org/10.1029/98JD01212, 1998.

Reimann S., Schaub, D., Stemmler, K., Folini, D., Hill, M., Hofer, P., Buchmann, B., Simmonds, P.G., Greally, B.R., O'Doherty, S.: Halogenated greenhouse gases at the Swiss High Alpine Site of Jungfraujoch (3580 m asl): Continuous measurements and their use for regional European source allocation, J. Geophys. Res., 109, D05307J, https://doi.org/10.1080/10473289.2007.10465319109, 2004.

Wan, D., Xu, J., Zhang, J., Tong, X., Hu, J.: Historical and projected emissions of major halocarbons in China, Atmos. Environ., 43(36), 5822-5829, 2009.