Journal of Geophysical Research: Atmospheres

U.S. emissions of HFC-134a derived for 2008–2012 from an extensive flask-air sampling network

Lei Hu[1,2], Stephen A. Montzka[2], John B. Miller[1,2], Arlyn E. Andrews[2], Scott J. Lehman[3], Benjamin R. Miller[1,2], Kirk Thoning[2], Colm Sweeney[1,2], Huilin Chen[4], David S. Godwin[5], Kenneth Masarie[2], Lori Bruhwiler[2], Marc L. Fischer[6], Sebastien C. Biraud[7], Margaret S. Torn[7], Marikate Mountain[8], Thomas Nehrkorn[8], Janusz Eluszkiewicz[8], Scot Miller[9], Roland R. Draxler[10], Ariel F. Stein[10], Bradley D. Hall[2], James W. Elkins[2], and Pieter P. Tans[2]

[1]Cooperative Institute for Research in Environmental Science, University of Colorado Boulder, Boulder, Colorado, USA, [2]Global Monitoring Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, Colorado, USA, [3]Institute of Arctic and Alpine Research, University of Colorado Boulder, Boulder, Colorado, USA, [4]Centre for Isotope Research, University of Groningen, Groningen, Netherlands, [5]Alternatives and Emissions Reduction Branch, Stratospheric Protection Division, Office of Atmospheric Programs, Office of Air and Radiation, U.S. Environmental Protection Agency, Washington, District of Columbia, USA, [6]Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA, [7]Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA, [8]Atmospheric and Environmental Research, Lexington, Massachusetts, USA, [9]Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA, [10]Air Resources Laboratory, National Oceanic and Atmospheric Administration, College Park, Maryland, USA

Abstract U.S. national and regional emissions of HFC-134a are derived for 2008–2012 based on atmospheric observations from ground and aircraft sites across the U.S. and a newly developed regional inverse model. Synthetic data experiments were first conducted to optimize the model assimilation design and to assess model-data mismatch errors and prior flux error covariances computed using a maximum likelihood estimation technique. The synthetic data experiments also tested the sensitivity of derived national and regional emissions to a range of assumed prior emissions, with the goal of designing a system that was minimally reliant on the prior. We then explored the influence of additional sources of error in inversions with actual observations, such as those associated with background mole fractions and transport uncertainties. Estimated emissions of HFC-134a range from 52 to 61 Gg yr^{-1} for the contiguous U.S. during 2008–2012 for inversions using air transport from Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model driven by the 12 km resolution meteorological data from North American Mesoscale Forecast System (NAM12) and all tested combinations of prior emissions and background mole fractions. Estimated emissions for 2008–2010 were 20% lower when specifying alternative transport from Stochastic Time-Inverted Lagrangian Transport (STILT) model driven by the Weather Research and Forecasting (WRF) meteorology. Our estimates (for HYSPLIT-NAM12) are consistent with annual emissions reported by U.S. Environmental Protection Agency for the full study interval. The results suggest a 10–20% drop in U.S. national HFC-134a emission in 2009 coincident with a reduction in transportation-related fossil fuel CO_2 emissions, perhaps related to the economic recession. All inversions show seasonal variation in national HFC-134a emissions in all years, with summer emissions greater than winter emissions by 20–50%.

1. Introduction

Hydrofluorocarbons (HFCs) are used as replacements for chlorofluorocarbons (CFCs) and hydrochlorofluorocarbons in many societal applications because they do not deplete stratospheric ozone. However, long-lived HFCs are potent greenhouse gases with global warming potentials up to several thousand times larger than that of CO_2 over a 100 year time horizon [*Daniel et al.*, 2011; *Myhre et al.*, 2013]. Global atmospheric abundances of HFCs have been increasing rapidly since the middle 1990s [*Montzka et al.*, 1996, 2011; *O'Doherty et al.*, 2004, 2009, 2014; *Vollmer et al.*, 2011]. Between 2003 and 2008 the increase in direct radiative forcing from HFCs, along with SF_6 and perfluorocarbons (PFCs), was comparable to the increase in direct radiative forcing from CH_4 or N_2O, and it was equivalent to 6% of the increase in direct radiative forcing from CO_2 [*Montzka et al.*, 2011]. Without controls on production or emission, total radiative forcing from HFCs could be approximately 10–20% of the radiative forcing from CO_2 by 2050 in business-as-usual scenarios [*Gschrey et al.*, 2011; *Umweltbundesamt*,

2009; *Velders et al.*, 2009]. HFC-134a (CH_2FCF_3) is currently the most abundant HFC in the atmosphere and is used primarily in mobile air conditioners as a substitute for CFC-12 [*McCulloch et al.*, 2003]. The global warming potential of HFC-134a is about 1300 times that of CO_2 over a 100 year time horizon [*Myhre et al.*, 2013]. Radiative forcing contributed by HFC-134a accounted for about one half of total radiative forcing from all HFCs in 2008 [*Montzka et al.*, 2011].

Global emission estimates of long-lived gases such as HFC-134a can be relatively well constrained based on observed atmospheric growth rates and estimated global atmospheric lifetimes [e.g., *Daniel et al.*, 2007; *Montzka et al.*, 1999, 2014; *Saikawa et al.*, 2012, 2014]. Emissions on a regional scale are more difficult to quantify. Inventory-based approaches derive regional emissions based on production, sales and consumption data, and appropriate emissions factors [e.g., *Ashford et al.*, 2004; *McCulloch et al.*, 2003; *United States Environmental Protection Agency (US EPA)*, 2014]. Such "bottom-up" approaches require accurate information regarding production, sales, and consumption, and a comprehensive understanding of relevant emitting processes [*Winiwarter and Rypdal*, 2001]. Moreover, uncertainty in magnitudes of existing "banks" (materials in inventories, in in-use applications, and in disposed equipment) contributes substantially to the overall uncertainty of the bottom-up estimates [*Manning et al.*, 2003]. "Top-down" estimates based directly on atmospheric observations do not require prior knowledge of emission processes and intensities and can provide largely independent assessment of inventory-based bottom-up estimates. Different top-down approaches have been used to estimate HFC emissions on regional scales, such as various trace gas ratio flux scaling methods [e.g., *Barletta et al.*, 2011, 2013; *Miller et al.*, 2012; *Millet et al.*, 2009; *Yokouchi et al.*, 2006] and inverse modeling of observed atmospheric trace gas mole fractions [e.g., *Brunner et al.*, 2012; *Maione et al.*, 2014; *Manning et al.*, 2003; *Stohl et al.*, 2009].

Inverse modeling, which seeks to derive surface flux magnitude, and its spatiotemporal distribution, has become an increasingly common tool for inferring fluxes from atmospheric mole fraction observations [e.g., *Brunner et al.*, 2012; *Gourdji et al.*, 2012a, 2012b; *Jeong et al.*, 2012a, 2012b; *Lin et al.*, 2004; *Maione et al.*, 2014; *Manning et al.*, 2003; *O'Doherty et al.*, 2004; *Stohl et al.*, 2009]. When solving for fluxes at high temporal and spatial resolution, as is suggested to avoid so-called "aggregation error" [*Kaminski et al.*, 2001], the inverse approach requires additional information, which typically comes in the form of first guess (a priori) flux estimates and their errors. Additional information or effective reduction in the degree of freedom of unknowns comes via specifying spatial and/or temporal correlation among unknowns [e.g., *Michalak et al.*, 2004] or optimizing scaling factors of fluxes categorized and distributed by source [e.g., *Jeong et al.*, 2013] or (in the case of CO_2) ecosystem types [e.g., *Peters et al.*, 2007]. Furthermore, surface fluxes and their associated uncertainties derived with inverse methods are sensitive to choices of model setup parameters, such as values given to model-data mismatch errors, prior flux errors, and error correlations that are difficult to accurately estimate. Two common ways to determine model-data mismatch errors are (a) to estimate each error component individually from measurements and modeling [e.g., *Jeong et al.*, 2012a, 2012b; *Lin et al.*, 2004; *Zhao et al.*, 2009] and (b) to use the atmospheric observations to infer model-data mismatch errors based on a maximum likelihood estimation (MLE) technique [*Michalak et al.*, 2005]. The latter approach can also be used to determine prior flux errors and their correlation.

In this study, we developed an inverse modeling system to derive emissions of HFC-134a from the U.S. during 2008–2012 based on National Oceanic and Atmospheric Administration (NOAA) ongoing atmospheric measurements from 27 ground and aircraft sites across the U.S. that make up the U.S. portion of NOAA's Global Greenhouse Gas Reference network plus observations from a few remote sites over the Pacific basin. In this inversion system, we use the MLE technique to objectively estimate model-data mismatch errors and a prior flux error covariance matrix. A suite of synthetic data experiments was conducted in order to optimize the model assimilation time interval and temporal resolution and to assess the robustness of parameters derived from the MLE method. It was also used to evaluate the ability of our observational network and inversion system to derive credible flux estimates on regional and national scales in ways that were minimally reliant upon prior flux information. Minimal impact of the prior on the optimized fluxes is critical in our study so that inventories (which also serve as priors) can be independently compared to our "top-down" atmosphere-derived fluxes. With an optimized model design and increased confidence in the reliability of our derived fluxes from the synthetic data experiments, we estimated U.S. national and regional emissions of HFC-134a using authentic atmospheric mole fraction measurement data
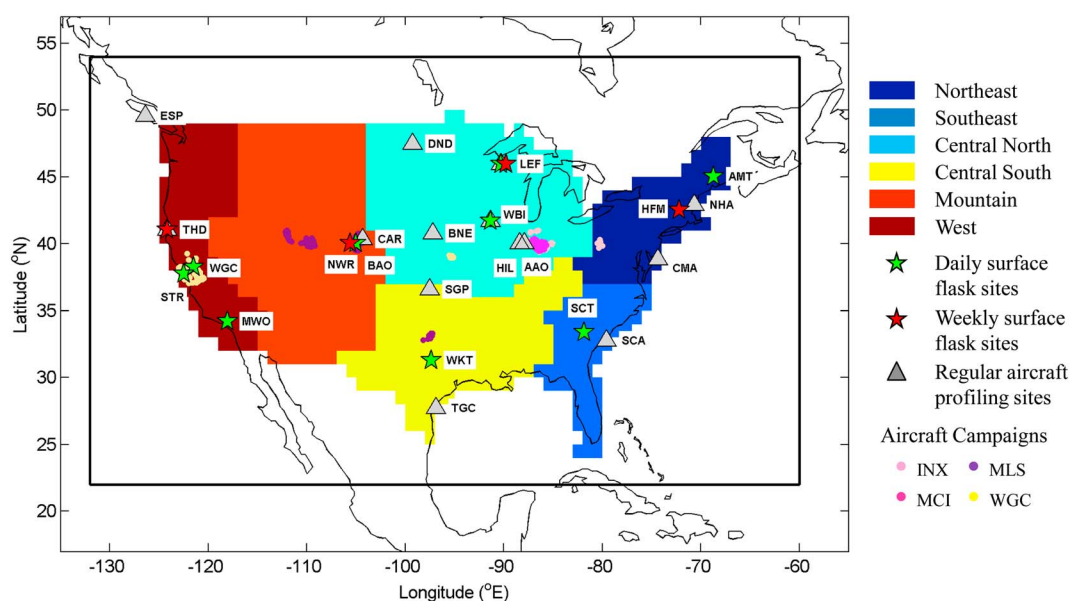
Figure 1. Site map. Green stars represent locations where ground-based, daily flask-air samples are collected; red stars represent ground-based, weekly stainless steel flask sampling sites; triangles are biweekly aircraft profiling sites; and circles indicate where flasks were collected during short-term aircraft campaigns during 2008–2012. The black box indicates the domain of our inversion system. Different colored areas denote subnational regions for which emissions are derived.

obtained from 2008 through 2012 with uncertainties determined from consideration of different priors, background determination methods, and multiple transport models.

2. Methodology

2.1. Observations, Measurements, and Data

Measurements of HFC-134a were made from air samples collected from ground-based inlets at 3–483 m above ground level (m agl) (including short and tall towers) and from aircraft sites across the U.S. during 2008–2012 within NOAA's Global Greenhouse Gas Reference Network (Figure 1 and Table 1). Approximately daily air samples were collected in glass flasks at ground sites including AMT, BAO, LEF, MWO, SCT, STR, WBI, WKT, and WGC [*Andrews et al.*, 2014]. Weekly air samples were collected in paired stainless steel flasks at ground sites including HFM, LEF, NWR, and THD (Table 1). Measurements at ground sites started before 2008 except at AMT, SCT, and MWO (Table 1). Air samples from aircraft profiles were collected approximately biweekly in glass flasks at altitudes ranging from 200 to 8000 m agl from AAO, BNE, CAR, CMA, DND, ESP, HIL, LEF, NHA, SCA, SGP [*Biraud et al.*, 2013], TGC, THD, and WBI (Figure 1 and Table 1). Measurements at AAO and BNE were terminated in 2009 and 2011, respectively (Table 1). Sampling frequency at tower sites like AMT, SCT, LEF, WBI, WKT, and WGC and a few aircraft sites such as CMA, NHA, TGC, and ESP was reduced by approximately 50% during mid-2011 to 2012 due to funding reductions.

Collected flask air samples were shipped back to the NOAA Earth System Research Laboratory's Global Monitoring Division in Boulder, Colorado, USA, and analyzed on one of two instruments using gas chromatography and mass spectrometry. Measured responses from samples are referenced to those determined from high-pressure (< 14 MPa) real air samples stored in silanized aluminum cylinders. Dry air mole fractions in glass flasks are typically determined from a single aliquot from a single flask; periodically in a subset of samplings, however, two glass flasks were filled in parallel to assess measurement reproducibility and precision. The median difference in HFC-134a mole fractions measured in these simultaneously filled flasks was 0.4% (90% of the time it was <1.5%). The NOAA HFC-134a standard scale is based on four standards prepared with gravimetric techniques at mole fractions between 5 and 225 parts per trillion (ppt, pmol mol^{-1}). The accuracy of this scale is estimated to be better than 2% based on consideration of impurities in the starting reagent material, and uncertainties in gravimetry and in the molecular mass of the diluent air. Intercomparison among laboratories that make the global measurements

Table 1. Information for Sites Used in the Inversion Analysis During 2008–2012[a]

| Sites	Latitude (°N)	Longitude (°E)	Altitude (m agl)	Number of Measurements[b]	Number of Data Used in Inversion[c]	Average Monthly Data Frequency[d]	Data Period[e]
			Ground-Based, Daily Glass Flask Sites				
AMT	45.03	−68.68	107	901	719	16	11/2008 to 12/2012
BAO	40.05	−105	300	1864	1334	23	1/2008 to 12/2012
LEF	45.95	−90.27	244	828	329	18	1/2008 to 5/2009
			396	1101	774		6/2009 to 12/2012
MWO	34.22	−118.06	43–46	948	344	10	4/2010 to 12/2012
SCT	33.41	−81.83	305	1259	902	18	8/2008 to 12/2012
STR	37.76	−122.45	232	2149	1182	21	1/2008 to 12/2012
WBI	41.72	−91.35	378–379	1675	994	17	1/2008 to 12/2012
WGC	38.27	−121.49	91	1220	907	16	1/2008 to 12/2012
			483	14	14		1/2008 to 3/2009
WKT	31.31	−97.33	457	1471	1001	18	1/2008 to 12/2012
			Ground-Based, Weekly Stainless Steel Flask Sites				
HFM	42.54	−72.17	29	212	208	3	1/2008 to 12/2012
LEF	45.95	−90.27	396	195	191	3	1/2008 to 12/2012
NWR	40.05	−105.59	3	200	182	3	1/2008 to 12/2012
THD	41.05	−124.15	13	231	214	4	1/2008 to 12/2012
			Biweekly Aircraft Profiling Sites				
AAO	40.05	−8.37	180 - 4628	1071	178	9	2/2008 to 9/2009
BNE	40.8	−97.18	135 - 7688	375	78	2	2/2008 to 4/2011
CAR	40.37	−104.3	267 - 7357	1241	107	2	1/2008 to 12/2012
CMA	38.83	−74.32	236 - 8050	1021	219	4	1/2008 to 12/2012
DND	47.5	−99.24	86–7654	566	90	2	3/2008 to 11/2009
							6/2010 to 1/2011
							9/2011 to 12/2012
ESP	49.58	−126.37	147–5787	1333	200	4	3/2008 to 12/2012
HIL	40.07	−87.91	349–7908	579	58	2	1/2008 to 4/2009 2/2011 to 12/2012
LEF	45.95	−90.27	92–3528	1039	213	4	1/2008 to 12/2012
NHA	42.95	−70.63	0–8069	1317	208	4	1/2008 to 12/2012
SCA	32.77	−79.55	218–8569	834	134	3	1/2008 to 12/2012
SGP	36.61	−97.49	0–12743	2002	342	6	2/2008 to 12/2012
TGC	27.73	−96.86	250–8107	811	133	3	1/2008 to 8/2011 3/2012 to 12/2012
THD	41.05	−124.15	205–8034	505	72	2	3/2008 to 7/2010 2/2011 to 11/2011 6/2012 to 8/2012
WBI	41.72	−91.35	338–7981	686	77	2	1/2008 to 2/2012
			Short-Term Aircraft Campaigns				
INX	n.a.	n.a.	99–3269	141	0	-	10/2010 to 12/2012
MLS	n.a.	n.a.	0–3530	207	0	-	2/2012 to 6/2012
MCI	n.a.	n.a.	104–2835	71	0	-	3/2008 to 5/2008
WGC	n.a.	n.a.	6–3167	155	0	-	5/2008 to 3/2009

[a]Notes: n.a. = not applicable.
[b]Total number of HFC-134a measurements made at individual sites.
[c]Number of observations used in the inversion that were considered independent. Here we only included day-time data (0700–1800 local standard time). We also averaged paired flask samples (collected within 30 min from tower sites or simultaneously at all sites) and airborne flask samples collected between 0–500 m agl and 500–1000 m agl.
[d]Average monthly data frequency was calculated based on the number of observations used in the inversion that were considered independent.
[e]Data gaps smaller than 6 months were not shown in the table.

of HFC-134a (e.g., Advanced Global Atmospheric Gases Experiment) suggests consistency within 2% with the NOAA results [*Hall et al.*, 2014; *Montzka et al.*, 2011]. Consistency in calibration is ensured over time by repeated analyses of a suite of gravimetric standards and a suite of archived real air samples stored at high pressure in silanized aluminum or humidified electropolished stainless steel cylinders (for more measurement details, see *Montzka et al.* [1996]). Results from the repeated analyses of archive tanks on each instrument
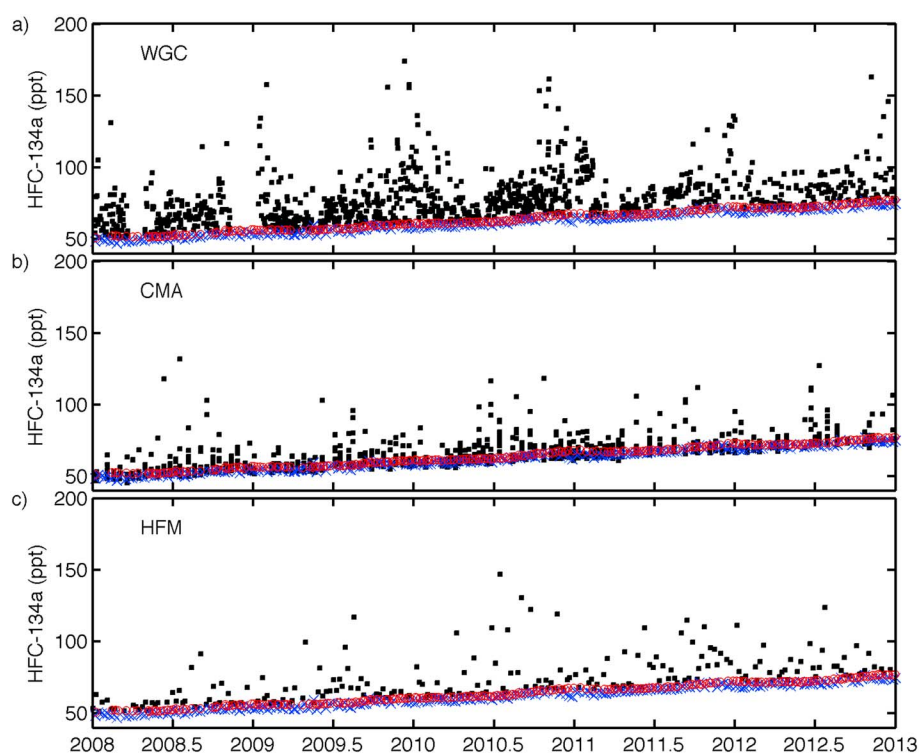
Figure 2. Time series of atmospheric observations for HFC-134a (black) at a selected daily flask site (a) (i.e., WGC), a selected biweekly aircraft profiling site (b) (i.e., CMA) below 1000 m agl, and a selected weekly stainless steel flask site (c) (i.e. HFM). Observed atmospheric mole fractions of HFC-134a at two remote sites, KUM (19.52°N, 154.82°W) (blue) and BRW (71.32°N, 156.61°W) (red) are shown to illustrate substantial enhancements at some of our continental sites due to recent emissions.

suggest that the calibration scale has been maintained within 0.25% during 2008–2012. Consideration of the analysis of these archive tanks on both instruments and of results from stainless steel flasks analyzed on both instruments ($n = \sim150$ pairs per year) suggests consistency in results from the different instruments within 0.5% on average.

Air samples collected in stainless steel flasks as pairs are part of a broader network at NOAA for characterizing long-term global changes in remote atmospheric mole fractions of trace gases including HFC-134 [*Montzka et al.*, 1996, 2014]. Results from remote Northern Hemisphere (NH) surface sites (i.e., ALT (82.45°N, 62.51°W), BRW (71.32°N, 156.61°W), THD (41.05°N, 124.15°W), KUM (19.52°N, 154.82°W)) and from low altitudes at remote NH aircraft sites (i.e. ESP (49.58°N, 126.37°W), ETL (54.35°N, 104.98°W), THD (41.05°N, 124.15°W); these results are from glass flasks), in combination with air back trajectories provide an estimate of background mole fractions in air reaching the continental U.S. against which regional and local enhancements occur. More details about estimation of background mole fractions are described in section 2.5. Examples of time series at selected continental ground and aircraft sites are shown along with data from two remote-atmosphere sites (Figure 2) to illustrate the substantial enhancements above background mole fractions that we observe for HFC-134a at some continental sites as a result of recent emissions.

Only the subset of available data thought to contain independent information about upstream fluxes was used in the inversion (i.e., samples with uncorrelated errors) (Table 1). Accordingly, we averaged data from flasks collected within 30 min of one another at each site. We also averaged results from flasks collected between 0 and 500 m agl and between 500 and 1000 m agl in the aircraft profiles. To minimize the influence of enhanced transport errors at night, we only included samples collected during 0700–1800 local standard time [*Geels et al.*, 2007; *Gerbig et al.*, 2008; *Gourdji et al.*, 2010]. Samples collected above 1000 m agl from aircraft sites, along with those from short-term aircraft campaigns, were excluded from the inverse analysis and used later to provide an independent evaluation of derived fluxes.

2.2. Inversion Framework

We employed a Bayesian approach to estimate monthly fluxes of HFC-134a at $1° \times 1°$ resolution using observations obtained from 2008 to 2012. The linkage between atmospheric observations and fluxes is through the linear equation

$$z = fF\lambda + \varepsilon \tag{1}$$

where z is an observed enhancement in atmospheric mole fractions above background values owing to surface emissions (ppt); the dimension of z is $n \times 1$, where n represents the number of independent observations; f represents the spatially resolved ($1° \times 1°$) sensitivity of observed enhancements to upstream fluxes (ppt (pmol $m^{-2} s^{-1})^{-1}$), or sample "footprints" [*Lin et al.*, 2003]; the dimension of f is $n \times m$, where m stands for the number of gridded fluxes in space and time; F stands for the $1° \times 1°$ prior emissions (pmol $m^{-2} s^{-1}$, dimension: $m \times 1$); λ is a vector that includes scaling factors applied to the prior emission field (F) (dimension: $m \times 1$) that will be optimized in the inversion, and ε represents model-data mismatch errors (in ppt) that include both measurement and model errors. In our application, measurements of HFC-134a are very precise (the agreement among paired flasks and between two instruments is typically better than 1.5%; see section 2.1); thus, model errors, such as inaccurate simulation of atmospheric transport, dominate the model-data mismatch errors.

Substituting fF with K, equation (1) can be written as

$$z = K\lambda + \varepsilon \tag{2}$$

Based on Bayes' theorem, along with assumptions of Gaussian errors [*Rodgers*, 2000], the scaling factors λ can be solved with equation (3):

$$\lambda = \lambda_p + QK^T \left(KQK^T + R\right)^{-1} \left(z - K\lambda_p\right) \tag{3}$$

The final solution of λ is a balance between a prior guess (λ_p, a vector with values of 1) and the emission scaling informed by atmospheric data. This balance is determined by the model-data mismatch (R; dimension: $n \times n$) and prior flux error covariance (Q; dimension: $m \times m$) matrices. We assume that transport errors are uncorrelated and all observations used in the analysis provide independent information about fluxes and so that R is a diagonal matrix with site-dependent, seasonally and interannually varying values (σ_r^2). Because it is difficult to test the influence of these assumptions, we consider multiple transport models in this study to explore the influence of systematic errors of transport on the derived emission (see section 2.3 below). We also construct the prior flux error covariance matrix (Q) to obtain an exponentially decaying covariance function (equation (4)) with a prior flux error of σ_q and spatial and temporal correlation scales of τ_l and τ_t (the 95% correlation scales are approximately 3 τ_l and 3 τ_t) [*Gourdji et al.*, 2010; *Michalak et al.*, 2004]:

$$Q = \sigma_q^2 \begin{vmatrix} 1 & \exp\left(-\frac{h_{s,1,2}}{\tau_l}\right)\exp\left(-\frac{h_{t,1,2}}{\tau_t}\right) & \cdots & \exp\left(-\frac{h_{s,1,m}}{\tau_l}\right)\exp\left(-\frac{h_{t,1,m}}{\tau_t}\right) \\ \exp\left(-\frac{h_{s,2,1}}{\tau_l}\right)\exp\left(-\frac{h_{t,2,1}}{\tau_t}\right) & 1 & \cdots & \exp\left(-\frac{h_{s,2,m}}{\tau_l}\right)\exp\left(-\frac{h_{t,2,m}}{\tau_t}\right) \\ \vdots & \vdots & \vdots & \vdots \\ \exp\left(-\frac{h_{s,m,1}}{\tau_l}\right)\exp\left(-\frac{h_{t,m,1}}{\tau_t}\right) & \exp\left(-\frac{h_{s,m,2}}{\tau_l}\right)\exp\left(-\frac{h_{t,m,2}}{\tau_t}\right) & \cdots & 1 \end{vmatrix} \tag{4}$$

where, h_s and h_t are separation distances between grid cells in space and time.

The posterior error covariance matrix (V) of the $1° \times 1°$ scale factors (λ) was solved with equation (5) [*Rodgers*, 2000; *Yadav and Michalak*, 2013]:

$$V = Q - QK^T \left(KQK^T + R\right)^{-1} KQ \tag{5}$$

With optimized scaling factors (λ) and their covariance matrix (V), a posterior emission (E) and its covariance matrix (C) can be calculated with $E = F\lambda$ and $C = FVF^T$. To improve the computational efficiency, we use algorithms given by *Yadav and Michalak* [2013] for inverting large matrices.

2.3. Calculating Footprints

Footprints (f) were calculated for each sampling event at all sites during 2008–2012 using the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model (http://ready.arl.noaa.gov/HYSPLIT_disp.php) [*Draxler and Hess*, 1997, 1998; *Draxler*, 1999] driven with archived meteorological data from the 12 km resolution North American Mesoscale Forecast System model (NAM12, domain: ~ 60–140°W, 20–60°N). Footprint-generation capability is a recent addition to HYSPLIT and implementation was based on the Stochastic Time-Inverted Lagrangian Transport (STILT) model [*Lin et al.*, 2003]. STILT was originally derived from HYSPLIT and there are many similarities, but some of the underlying parameterizations are different. For comparison, a second set of footprints was also calculated for samples collected during 2008–2010 using STILT driven by a Weather Research and Forecasting model (WRF) simulation customized for Lagrangian modeling [*Nehrkorn et al.*, 2010]. The WRF meteorology has nested domains with resolution of 10 km over the continental U.S. and 40 km over the rest of the North America. Both HYSPLIT-NAM12 and STILT-WRF were run for each sample with an ensemble of 500 particles that were transported backward in time for 10 days with hourly output.

Hegarty et al. [2013] evaluated forward-in-time simulations of tracer release experiments using HYSPLIT, STILT, and FLEXible PARTicle dispersion (FLEXPART) model [*Stohl et al.*, 2009] and found that the particle dispersion models behaved similarly, with differences primarily attributable to meteorological driver data. As part of the current study, we assessed the influence of different transport models and meteorological data on derived fluxes (section 3.2).

2.4. Estimating Model-Data Mismatch Errors and Prior Flux Error Covariances

In any inversion calculation, it is critical to appropriately estimate model-data mismatch errors and prior flux error covariances because they determine the relative weight given to the observations versus the prior flux field in the determination of retrieved fluxes and their uncertainties [*Gourdji et al.*, 2010; *Peters et al.*, 2005]. For example, underestimated model-data mismatch errors may yield fluxes that are anticorrelated in space and time and exhibit a "dipole" phenomenon [*Powell*, 2013; *Rodgers*, 2000], whereas overestimating model-data mismatch errors or underestimating prior flux errors would take insufficient advantage of the information contained within the observations and result in an unnecessary bias toward the prior in posterior fluxes. Furthermore, underestimating correlation scales (temporally and spatially) in the prior flux error covariance matrix would likely reduce the ability of an inversion to correct large-scale biases in a prior flux field, whereas overestimating them might reduce the temporal and spatial heterogeneity in derived fluxes. In this study, we use the maximum likelihood estimation (MLE) method [*Michalak et al.*, 2005] to determine the prior flux error (σ_q), spatial and temporal correlation scales (τ_l and τ_t) of prior flux errors, and site-dependent model-data mismatch errors (σ_r) that vary seasonally, interannually, and by site. The MLE method is designed to objectively derive these parameters given a suite of atmospheric observations. This method first uses the a priori emissions to predict atmospheric mole fractions. It then separates model-data mismatch error and prior flux error by assuming that (1) the difference between predicted and actual observations results from model-data mismatch and prior flux errors and (2) the effect of model-data mismatch error on predicted observation errors is expected to be random, whereas the main influence of prior flux errors is systematic [*Michalak et al.*, 2005]. The MLE implementation ensures that residuals between observed and simulated atmospheric mole fractions and residuals between posterior and prior fluxes follow a χ^2 distribution [*Michalak et al.*, 2005]. Use of this method can be computationally intensive; thus, only a simplistic model of errors can be considered. Limitations inherent to this approach include (1) the assumption that model-data mismatch errors are random, despite the possible presence of systematic errors in the simulated transport (e.g., boundary layer height, and wind speed); (2) possible correlation of estimated model-data mismatch error and prior flux error covariance matrices [*Ganesan et al.*, 2014]; and (3) inability to propagate uncertainties of these estimated covariance parameters into the inversion calculations [*Ganesan et al.*, 2014].

2.5. Estimating Background Mole Fractions

Regional emission magnitudes were derived from measured enhancements in atmospheric mole fractions at sites near sources relative to mole fractions measured in the background atmosphere upwind of these sources. Hence, errors in estimated background or up-wind atmospheric mole fractions of HFC-134a may lead to unresolved biases in the retrieved fluxes. In order to evaluate the sensitivity of calculated fluxes to this uncertainty, we employed three different approaches to the calculation of background mole fractions. In the first, we derived background mole fractions at individual surface sites based on the lowest 10th percentile

of observed mole fractions within a 3 month moving window. At aircraft sites, background mole fractions were estimated on a flight-by-flight basis as the average mole fraction in the free troposphere (3000–7000 m agl). This approach (hereinafter, referred to as "the 10th percentile + average free troposphere" or background 1) has the advantage of relying only on the data from each site. However, cases exist where this approach may lead to a bias. For example, for species with a strong latitudinal gradient in the remote atmosphere, the background values sampled entirely or dominantly from a particular wind sector may not be representative of background mole fractions for air transported from other directions. Similarly for the aircraft profiles, significant wind shear may exist above the planetary boundary layer and the transport history of free tropospheric air may differ from that of air near the surface.

In the second approach, we estimated background mole fractions using the marine boundary layer reference (http://www.esrl.noaa.gov/gmd/ccgg/mbl/) [*Masarie and Tans*, 1995] sampled at the latitudes of individual sites. The marine boundary layer reference for HFC-134a was created as a time- and latitude- varying surface of mole fractions that was based on smoothed interpolations of observations made primarily at remote sites in the Pacific basin (i.e., ALT, BRW, ESP, ETL, THD, and KUM; section 2.1) as described in *Masarie and Tans* [1995]. This is because air flow into the U.S. is predominantly from the west (due to westerly geostrophic flow in midlatitudes). This method neglects vertical gradients in mole fractions in the remote atmosphere that may influence near-surface observations. This approach also does not consider the actual transport history of the air (air may originate from a different latitude) that was sampled at individual sites. Hereinafter, this approach is referred to as "the marine boundary layer reference" or background 2.

In the third approach, we extended the Pacific marine boundary layer reference vertically to produce a background mole fraction field varying across latitudes, altitudes, and time. This three-dimensional background "curtain" represents mole fractions of HFC-134a in the remote atmosphere between 10° and 70°N and from 0 to 7000 m above sea level. It was derived from atmospheric observations of HFC-134a made at remote ground-based sites and from aircraft samples collected in the free troposphere, using the same curve-fitting algorithms described in *Masarie and Tans* [1995]. Similar background fields have been used in regional inverse-modeling studies of CH_4, CO_2, and other gases [e.g., *Gourdji et al.*, 2012a, 2012b; *Jeong et al.*, 2012a, 2012b; *Miller et al.*, 2013; *Zhao et al.*, 2009]. For each observation, 500 air particles were transported backward in time for 10 days to estimate when and where air masses intersected the boundary of the background curtain (which is assumed at the boundary of the NAM12 domain) during HYSPLIT-NAM12 footprint simulations. On average, 78–98% (a range of averages for all sites) of particles exited the boundary of the NAM12 domain within 10 days. For these particles, background mole fractions were obtained from the curtain at the intersection positions. For particles that remained inside of the domain after 10 days, we used the background mole fractions from the curtain at the latitudes and altitudes of the particles at their positions 10 days back, thereby neglecting any possible variation in background mole fractions with longitude. The average background mole fraction over the 500 particles was used as the background for each observation. Hereinafter, the third background approach is referred as "curtain + air back trajectories" or background 3.

2.6. Design of Synthetic Data Experiments

A suite of synthetic data experiments was conducted in order to optimize the model assimilation time interval and temporal resolution and to assess model-data mismatch errors and prior flux error covariances derived with MLE. More importantly, it was used to investigate the ability of the available observations within our sampling network, along with our inversion system, to accurately quantify flux magnitudes and their seasonality and interannual variability on multispatial and multitemporal scales (i.e., national, regional, and a 1° × 1° grid scales; monthly, yearly, and five-yearly scales). This is to make sure that derived fluxes using this framework are minimally reliant upon the prior that is generally from a bottom-up inventory.

In these experiments, we used a specified emission field as the "synthetic-true" emission (Figure 3 and Table 2). We created "synthetic observations" by multiplying the synthetic-true emission and surface-sensitivity footprints (calculated by HYSPLIT-NAM12) and then perturbing the synthetic observations with normally distributed noise. The frequency and location of the synthetic observations were identical to those of the actual observations. We evaluated the sensitivity of derived fluxes to a range of prior flux fields having flux magnitudes (e.g., Prior 1) and/or distributions (e.g., Priors 2–3) very different than the synthetic-true emission. We also specified a fourth prior (Prior 4) with prior flux error magnitude and prior flux error correlation scales that are more comparable to those of the real data inversions relative to other priors (Priors 1–3) (Table 3). In all cases,
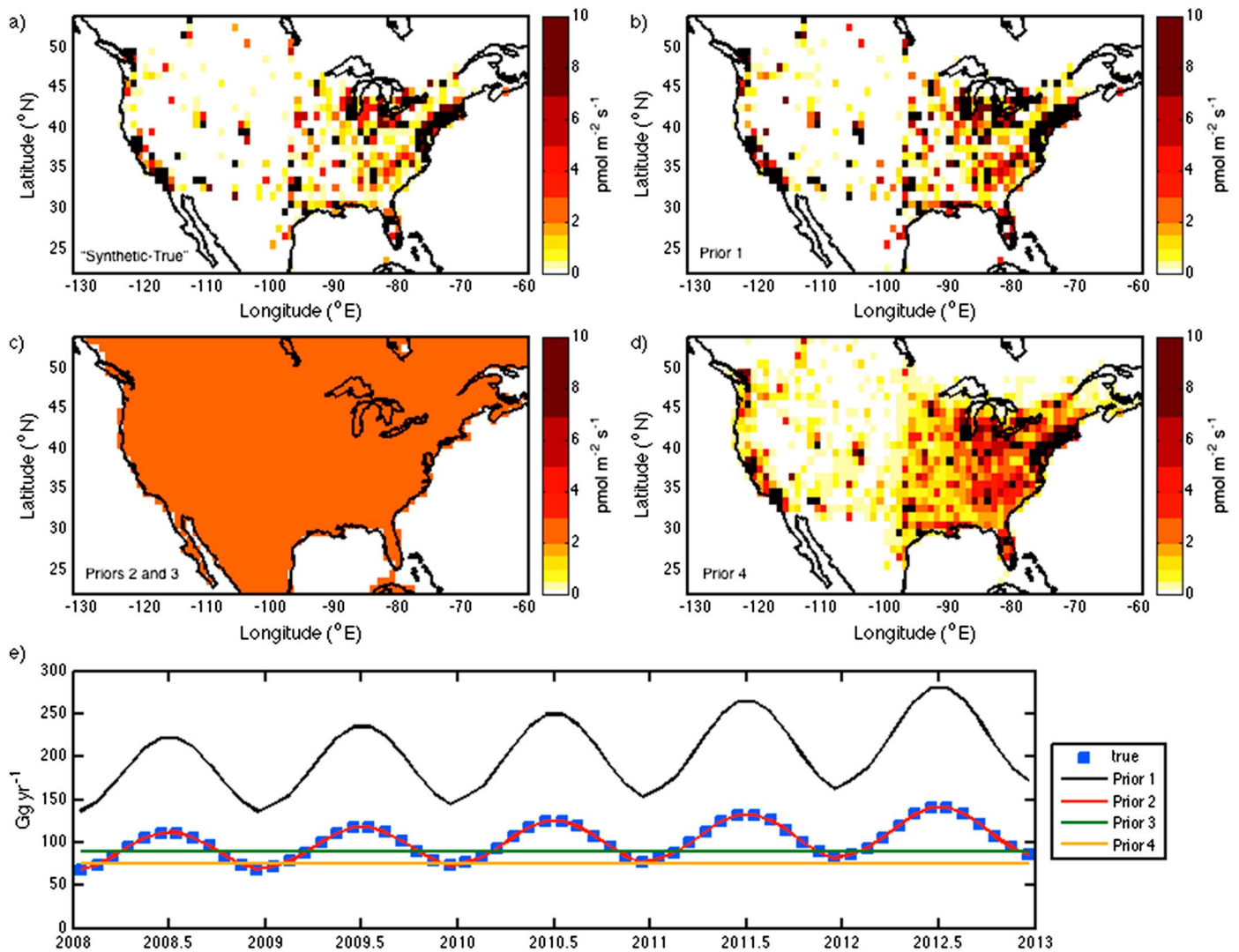
Figure 3. (a–d) Spatial distributions and (e) monthly total emissions of synthetic-true and prior emissions that were used in the synthetic data experiments (see also Table 2). Monthly total emissions are aggregated emissions throughout the entire model domain.

Table 2. Description of Various Emission Fields Used in Synthetic Data Experiments

Emission Fields	Description
Synthetic-true emission	Derived with 2008 HFC-134a gridded emissions from the EDGARv4.2 inventory, with an imposed 40% seasonality (peak-to-peak relative to the mean flux), and a 5.7 Gg yr^{-1} increasing emission trend during 2008–2012; it was then perturbed with a random multiplicative scaling factor between 0.1 and 2 over 2° × 2° regions each month.
Prior 1	Derived with the same emission distribution as in the synthetic-true emission but includes a scaling factor of 2 relative to the synthetic-true emission in all 1° × 1° grid cells. This prior is briefly referred to "doubled synthetic-true."
Prior 2	A seasonally varying, spatially constant emission prior (constant emission rate per surface area over all land grid cells). Monthly total fluxes aggregated over the entire domain and their seasonality are equal to those from the synthetic-true emission. This prior is referred to as the "seasonal, flat-distribution" prior.
Prior 3	An aseasonal, spatially constant emission prior (constant emission rate per surface area over all land grid cells). Annual total fluxes aggregated over the entire domain are set to the 2008 annual mean monthly flux from synthetic-true emission for all years. This prior is referred to as the "aseasonal, flat-distribution" prior.
Prior 4	An aseasonal prior, which was derived with the 2005 HFC-134a emission from the EDGARv4.1 inventory. This prior is referred to as the "EDGARv4.1" prior.

Table 3. Average Prior Flux Errors (in Fraction Relative to the "Synthetic-True" Emission) (σ_q) and Their Spatial and Temporal Correlation Scales (τ_s in km and τ_t in Days) obtained From a Variogram Technique and Maximum Likelihood Estimation (MLE) Over Five Assimilation Time Intervals in Synthetic Data and Real Data Inversions

Parameters	Synthetic Data (Prior 1, "Doubled Synthetic-True")		Synthetic Data (Prior 2, "Seasonal, Flat-Distribution")		Synthetic Data (Prior 3, "Aseaonal, Flat-Distribution")		Synthetic Data (Prior 4, "EDGARv4.1")		Real Data (Prior = "EDGARv4.2")
	Variogram Derived	MLE Derived	Variogram Derived	MLE Derived	Variogram Derived	MLE Derived	Variogram Derived	MLE Derived	MLE Derived
σ_q	0	0.4	4.2	4.4	4.6	4.4	0.7	1.3	1.4
τ_s	∞	5.4×10^5	112	99	120	124	118	97	138
τ_t	∞	1.1×10^7	545	2186	544	1196	777	67	328

the goal was to determine our ability to recover the synthetic-true emission using the synthetic observations, footprints, and multiple prior emission fields. We expected that posterior flux errors for Prior 4, which most closely resembled the error structure in the prior that was used in the real data inversion, would be more representative of those in a real data inversion than those derived from the other priors.

The magnitude (1σ) of the applied normally distributed noise in each season ($\sigma_{r,s}$) relative to the seasonally averaged synthetic observations (enhancements) at each site (z_s) was set by the ratio of the MLE-derived seasonal model-data mismatch errors to the seasonally averaged observed enhancements in a real data inversion using the 2008 HFC-134a EDGARv4.2 emission as a prior and HYSPLIT-NAM12 footprints (equation (6) and Table 4).

$$\left(\frac{\sigma_{r,s}}{z_s}\right)_{synthetic\ data} = \left(\frac{\sigma_{r,s}}{z_s}\right)_{real\ data} \tag{6}$$

The purpose of adding this additional noise was to (1) simulate model-data mismatch errors that arise from measurement and model errors in the real data inversions, (2) test whether the MLE methodology can return the assigned model-data mismatch errors (the applied random noise), and (3) determine the accuracy of derived fluxes obtainable from our actual air sampling network when the size of enhancements relative to model-data mismatch errors are similar to those in real data inversions.

3. Results and Discussion

3.1. Synthetic Data Inversion
3.1.1. Optimizing the Assimilation Time Interval and Temporal Resolution
In inverse modeling, the size of the matrix to be inverted increases geometrically as the number of observations and the size of a discretized flux vector increase, which may lead to practical computational constraints [*Bruhwiler et al.*, 2005]. To achieve a balance between computational efficiency and estimation accuracy, we tested multiple assimilation time intervals (i.e., 3 months, 1 year, and 18 months) with various temporal resolutions for derived fluxes (i.e., every three days vs. monthly) at a fixed spatial resolution of 1° × 1°. Our results suggest that using a shorter assimilation time interval (i.e., three months) may cause discontinuity in derived fluxes between sequential "batch" inversions and systematic biases in derived fluxes (relative to the synthetic-true flux), whereas using longer assimilation time intervals (e.g., 1 year or 18 months) can mitigate these problems. Moreover, a larger assimilation time interval provides flux estimates constrained by more observations [*Peters et al.*, 2005].

With respect to the temporal resolution of derived fluxes, we did not observe a significant difference in derived fluxes averaged over a month for inversions performed at monthly resolution or at substantially greater (every 3 days) resolution. In an attempt to balance computational efficiency and estimation accuracy, we divided our study period (2008–2012) into five batch inversion intervals. Each batch had a time interval of 15–18 months and a temporal resolution of 1 month in the state vector. Between two sequential batches, there was a 6 month overlap period that minimizes the problem of discontinuity in derived fluxes.
3.1.2. Assessing Model-Data Mismatch Errors and Prior Flux Error Covariances Given by MLE
To determine the accuracy of model-data mismatch errors derived from MLE and their sensitivity to prior fluxes, we ran the MLE calculation with the four different priors in synthetic data inversions (Table 2). The results suggest that, with a prior emission field that has a similar emission distribution to the synthetic-true emission (Priors 1 and 4) (Table 2 and Figure 3), the difference between the magnitude of noise added to

Table 4. Average Enhancements, MLE-Derived Model-Data Mismatch Errors Using HYSPLIT-NAM12 and Their Ratios Over 2008–2012 and 2008–2010 at Individual Sites in Real Data Inversions, Along With the Ratios of MLE-Derived Model-Data Mismatch Errors Using STILT-WRF Relative to Average Enhancements Over 2008–2010[a]

| Regions	Sites	2008–2012			2008–2010	
		Enhancements (ppt)	Model-Data Mismatch (HYSPLIT-NAM12) (ppt)	Ratio (HYSPLIT-NAM12)	Ratio (HYSPLIT-NAM12)	Ratio (STILT-WRF)
Northeast	AMT	5.09	3.38	0.66	0.60	0.63
	CMAa	12.70	5.88	0.46	0.58	0.43
	NHAa	8.78	5.16	0.59	0.52	0.57
	HFMs	8.98	5.17	0.58	0.49	0.54
Southeast	SCT	9.75	8.30	0.85	0.80	0.78
	SCAa	10.89	6.92	0.64	0.66	0.68
Central North	LEF	3.75	2.14	0.57	0.48	0.55
	LEFa	2.64	1.38	0.52	0.36	0.57
	LEFs	3.90	2.49	0.64	0.58	0.51
	WBI	6.45	5.02	0.78	0.78	0.77
	WBIa	4.93	3.25	0.66	0.58	0.59
	AAOa	8.72	6.60	0.76	0.76	0.75
	BNEa	5.21	3.17	0.61	0.60	0.58
	DNDa	2.49	1.79	0.72	0.67	0.53
	HILa	9.96	5.43	0.55	0.66	0.82
Central South	WKT	8.11	6.22	0.77	0.80	0.90
	SGPa	6.83	3.59	0.53	0.52	0.62
	TGCa	5.86	4.25	0.72	0.78	0.74
Mountain	BAO	8.73	11.59	1.33	1.25	1.35
	CARa	2.95	2.47	0.84	0.92	0.63
	NWRs	1.73	1.52	0.88	0.70	0.74
West	STR	10.99	17.35	1.58	1.67	1.65
	WGC	11.95	9.80	0.82	0.86	0.90
	MWO	54.55	59.44	1.09	1.05	0.90
	THDa	2.32	2.06	0.89	1.05	0.95
	THDs	2.15	2.53	1.18	1.35	0.86
	ESPa	1.00	1.13	1.13	0.99	1.27

[a]Notes: The magnitudes of model-data mismatch errors control the root-mean-square errors of simulated mole fraction enhancements using derived posterior fluxes (relative to observed mole fraction enhancements) for data that are included in the inversion analysis. Model-data mismatch errors derived with MLE were using the 2008 HFC-134a emission from EDGARv4.2 as a prior. When using the 2005 HFC-134a emission from EDGARv4.1 as a prior, the average agreement in MLE-derived model-data mismatch errors with those using the EDGARv4.2 as a prior is 8%. Tower sites, weekly stainless steel flask sites and aircraft sites are denoted with site code, site code + "s" and site code + "a."

synthetic observations and MLE-derived model-data mismatch errors is small (12 ± 12%). When using a prior that has a very different distribution than the synthetic-true emission distribution, a larger discrepancy between the MLE-derived model-data mismatch errors and the noise added to the synthetic observations was observed at most sites (19 ± 21%); at some sites (e.g., AMT, CMA, and MWO), the difference was as large as to 100%. This indicates that using a prior that has a spatial distribution very different from the synthetic-true emission may result in larger random errors in predicted mole fractions at certain locations (relative to synthetic observations). These random errors augment the model-data mismatch errors derived by MLE.

To evaluate the prior flux error covariance parameters derived by MLE, we first calculated the true prior flux errors (σ_q) as the ratio of the standard deviation of prior flux errors relative to the synthetic-true emissions. We derived the spatial and temporal correlation scales (τ_l and τ_t) of prior flux errors using a variogram technique and considered them as true correlation scales [*Cressie and Hawkins*, 1980]. The variogram (γ) here refers to correlation in prior flux errors between two grid cells separated in space or all grid cells in the entire space separated by a certain time. By fitting the variogram with the following function:

$$\gamma = \exp\left(-\frac{h}{\tau}\right) \tag{7}$$

where h (h_s or h_t) and τ (τ_s or τ_t) represent separation distance and correlation scales in space or time, we obtain estimates of the spatial and temporal correlation scales of prior flux errors for the four prior emission fields (Table 3). The fundamental difference between the variogram technique and the MLE is that the
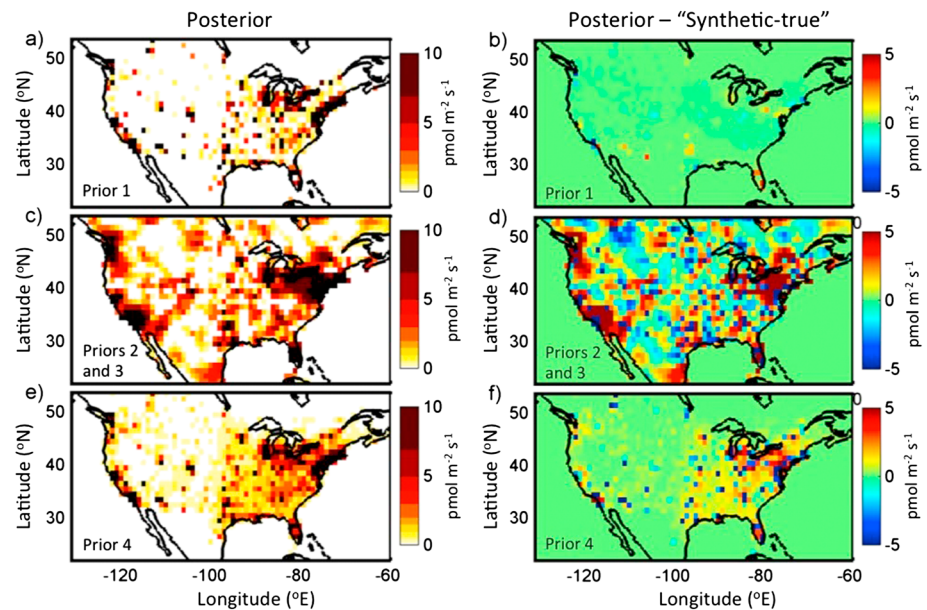
Figure 4. Left-hand side panels: 5 year averaged posterior fluxes that were calculated with (a) Prior 1, (c) Priors 2 and 3, and (e) Prior 4 in synthetic data experiments. Right-hand side panels: Grid-scale differences between posterior and synthetic-true emissions averaged over 5 years for results derived from (b) Prior 1, (d) Priors 2 and 3, and (f) Prior 4. Results from Priors 2 and 3 are shown in one panel because their derived posterior fluxes or the differences between their posterior fluxes and synthetic-true fluxes are not distinguishable on the plotted color scale.

variogram technique infers the correlation scales from the prior fluxes themselves, whereas the MLE derives those parameters using atmospheric observations. For Prior 1 in which the synthetic-true flux magnitude is simply scaled by a factor of 2 (Table 2), the variogram technique is not necessary because the prior flux errors relative to the synthetic-true fluxes are perfectly correlated and the correlation scales are infinite in time and space. Comparing to the true prior flux errors and their spatial correlation scales derived with the variogram technique, MLE-derived prior flux errors agree well within 20% in most cases (Table 3). For the MLE-derived temporal correlation scales of prior errors, their disagreements with the variogram-derived parameters are up to a factor of 3 (Table 3).

### 3.1.3. Assessing Derived Fluxes in Synthetic Data Experiments

Given the available observations from our air sampling network, potential inaccuracies in MLE-derived model-data mismatch errors and prior flux error covariances (section 3.1.2), and known shortcomings of this inversion framework (sections 2.2 and 2.4), it is important to investigate how accurately we can quantify fluxes on various spatial and temporal scales using our inversion system. For Prior 1 (double the magnitude and identical distribution of the synthetic-true emission), 95% of the derived 1° × 1° posterior fluxes agree with synthetic-true fluxes within 10% (Figure 4), suggesting that when the prior has a correct distribution but a consistently biased magnitude, we can reliably derive the synthetic-true fluxes on a 1° × 1° grid scale provided an appropriate prior error correlation scale is used (Table 3). Note that without any correlation in the prior flux error covariance matrix, derived fluxes had substantial errors compared to synthetic-true fluxes even on a national scale (not shown). This is because without correlation in prior flux errors, the fluxes to be derived would have a much higher degree of freedom that requires more observations to constrain the flux estimates. Compared to Prior 1, other priors (Priors 2–4) have different spatial distributions than the synthetic-true emission (Table 2). Although the spatial distributions of the associated posterior flux estimates from Priors 2 to 4 are more similar to the synthetic-true emission (Figure 4) than their priors (Figure 3), the differences between posterior and synthetic-true fluxes for these priors are large on a 1° × 1° grid scale (Figure 4) but become smaller when aggregated to larger scales (Figure 5).

In the real data inversion, we expect to have an imperfect prior, so defining the spatial scales over which derived fluxes are reliably derived is important. On a national scale, we find that derived posterior fluxes are fairly insensitive to prior flux distribution and prior flux magnitude with the priors we tested. Posterior flux
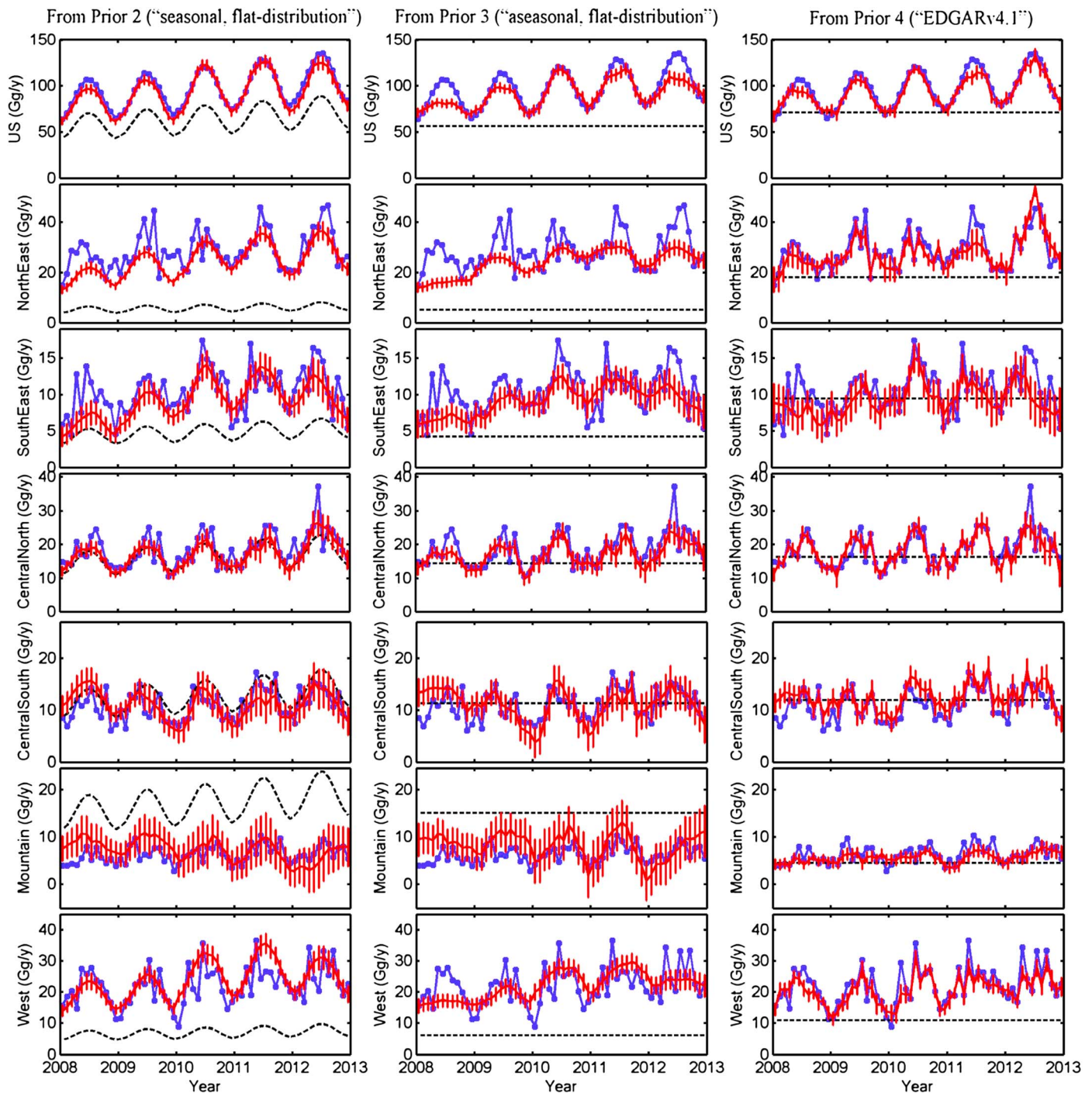
Figure 5. Monthly national and regional total fluxes from synthetic-true (blue), prior (black), and posterior (red) emissions for different prior emission fields used in the synthetic data inversions. Panels from left to right show results from Prior 2, Prior 3, and Prior 4. Panels from the top to the bottom show total emissions from the U.S. and for the regions specified in Figure 1.

errors on national annual and five-yearly scales are less than 11% and 6% given any prior we tested. On a monthly scale, the sensitivity of derived national fluxes to different priors is associated with the extent and intensity of the footprint of our sampling network. For example, in 2008, the footprint intensity of our sampling network was relatively low, especially in the east and southwest due to fewer observations made in these regions (Figure 6). During this period, the discrepancy between derived monthly posterior fluxes from the least informative prior (Prior 3, Table 2) and monthly synthetic-true emissions was as large as to 24%
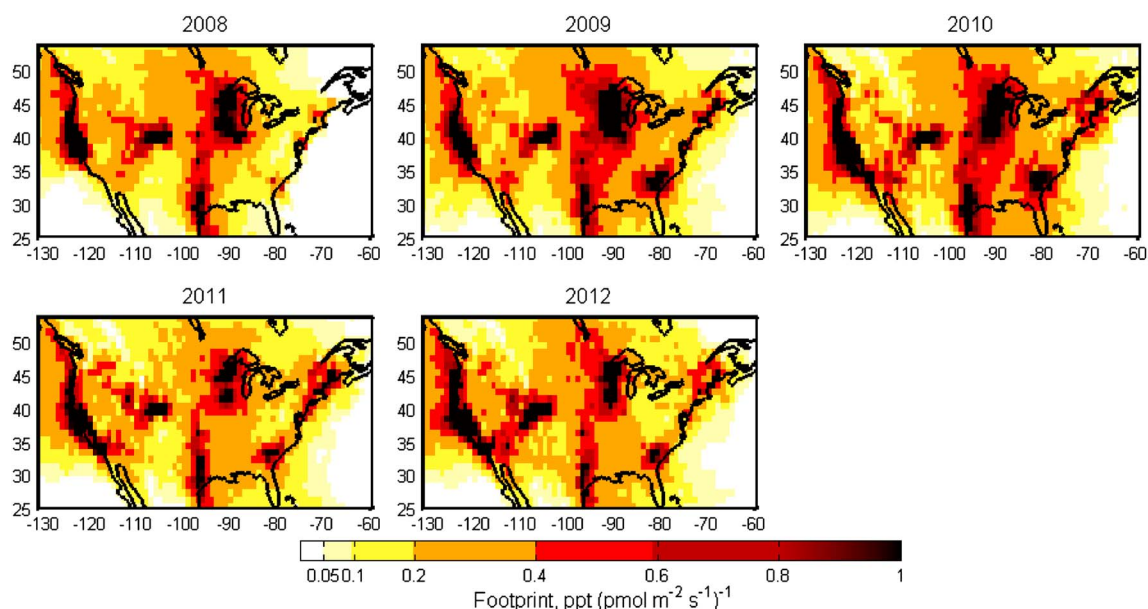
Figure 6. Annual average sensitivity of observations to upstream fluxes (footprint) derived with HYSPLIT-NAM12 based on available observations of HFC-134a each year within our air sampling network between 2008 and 2012.

(Figure 5); for monthly posterior fluxes derived from a more realistic prior with temporal variation or spatial distribution more similar to synthetic-true emission (Prior 2 and Prior 4), relative errors (compared to monthly synthetic-true emissions) are < 15% (Figure 5). This suggests that, when the footprint intensity of our sampling network throughout the study region is low, derived emission estimates are more dependent upon the spatial distribution and temporal variability imposed by the prior. In 2010, when the largest number of observations were available relative to other years over our study period, the footprint intensity of our sampling network was relatively high throughout our study region (Figure 6) and derived monthly fluxes from even the least informative prior (Prior 3) are consistent with monthly synthetic-true emissions within 10% (Figure 5). Furthermore, in 2010, errors of derived posterior fluxes were at the lowest compared to those in other years for any prior that has an imperfect spatial distribution (i.e., Priors 2–4). Interestingly, we found a strong correlation between errors of derived monthly posterior fluxes (expressed as root-mean-square errors of derived monthly posterior fluxes relative to monthly synthetic-true fluxes) and the number of available observations when we used a prior with constant emissions over space (i.e., Priors 2 and 3) ($R^2 = 0.86$–0.96; $n = 5$). As the accuracy of the prior increases, this dependence becomes less tightly correlated. Funding reductions in 2011 and 2012 necessitated a reduction in sampling frequency that appears as a substantial reduction in extent and intensity of footprints (Figure 6), which explains the increased dependence to the prior and elevated errors of derived posterior fluxes with a spatially constant prior (i.e., Priors 2 and 3) during that period (Figure 5).

On a regional scale (indicated in Figure 1) (on the order of 10^6 km^2), we were able to derive 5 year averaged fluxes within 32% of the synthetic-true flux (or 5 Gg yr^{-1} on an absolute scale) (Figures 5 and 7) using any prior described above (Table 2). At monthly or annual scales, derived regional fluxes are more sensitive to prior flux distribution than those derived on a national scale (Figures 5 and 7), especially in regions where the footprint intensity is relatively low, such as the northeast and mountain regions (Figures 6), or where the model-data mismatch errors are high relative to the specified or observed enhancements, such as the west and mountain regions (Table 4). In such regions, the relative errors of derived posterior fluxes for relatively unrealistic priors that have a constant emission rate over space (Priors 2 and 3) may be up to 3 times those derived from a prior whose distribution is more similar to the synthetic-true emission (Prior 4) (Figure 7).

One common feature we found in derived fluxes on both national and regional scales is that when using a prior that has a constant emission rate over space (Priors 2 and 3), the 1σ or 3σ errors of posterior fluxes calculated with the Bayesian method as implemented here does not always encompass the synthetic-true
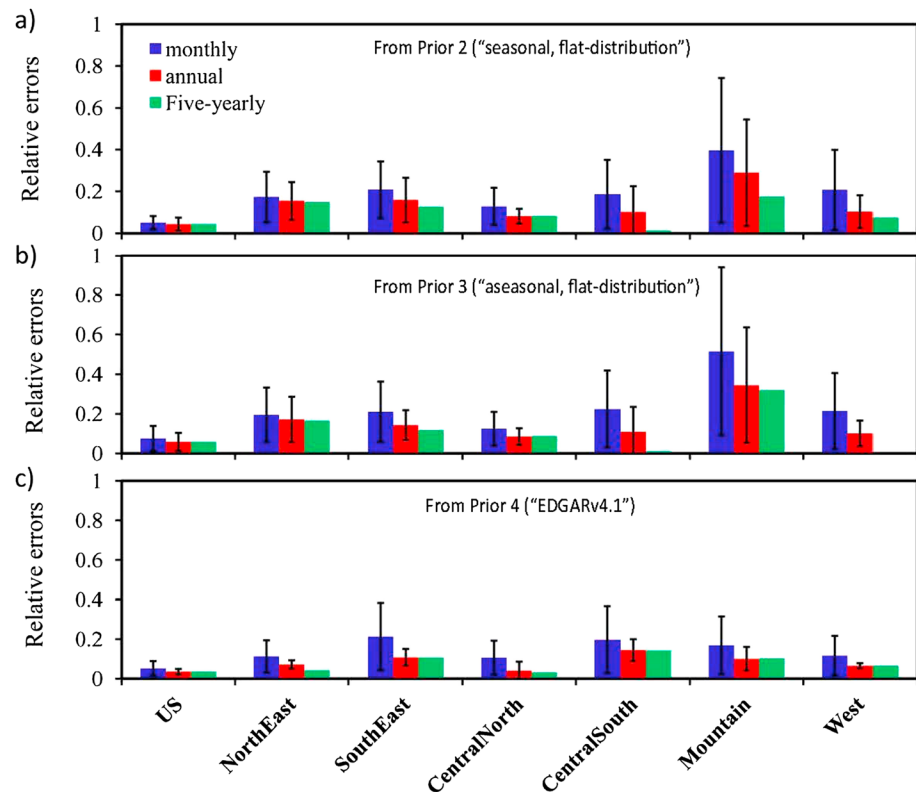
Figure 7. Relative errors between derived posterior and synthetic-true fluxes from various regions on monthly (blue), annual (red), and five-yearly (green) time scales using Priors (2–4). Error bars stand for one standard deviation of relative errors, determined from their monthly (blue) and interannual (red) variations.

emission (Figure 5). Since errors of gridded posterior fluxes (relative to gridded synthetic-true fluxes) are not normally distributed, we cannot reliably assess the probability that derived posterior fluxes encompass the true fluxes. However, as shown in Table 3, MLE-derived temporal correlation scales for prior flux errors in Priors (2 and 3) are a few factors larger than those calculated with the variogram technique, which would cause an underestimate in derived posterior flux errors and, potentially, the lack of overlap between true and derived emissions for these priors during some months (Figure 5).

3.1.4. Summary and Limitations of Synthetic Data Inversion

Results of the synthetic data experiments have helped us to optimize our inverse model design given our observational data, to assess the parameters derived from MLE, and most importantly, to test the sensitivity of derived fluxes at various spatial and temporal scales to a range of different priors. These experiments imply that, on a national scale, fluxes derived from this observational network and inversion system are largely insensitive to the prior. On a regional scale, derived 5 year average fluxes from four widely different priors are consistent with the synthetic-true fluxes within 32%. For national monthly or regional monthly and annual scales, the sensitivity of derived fluxes to the priors tested here appear to be dependent upon the footprint of the sampling network over the studied region and the ratio of model-data mismatch errors to enhancements.

In reality, the extent to which we can retrieve the true fluxes on various scales is not only dependent on the factors mentioned above (i.e., prior emissions, footprints, model-data mismatch errors, and prior flux error covariances); it is also associated with uncertainty in background and transport. In real data inversions, we have presubtracted background mole fractions from observations to obtain mole fraction enhancements, so the assumed background modifies the calculated enhancements directly. In synthetic data inversions, we created enhancements (synthetic observations) by convolving footprints with a given emission field (a synthetic-true emission), so the background term is zero and therefore perfectly known. In addition, the transport used to create synthetic observations is the same transport that was employed in the inversion
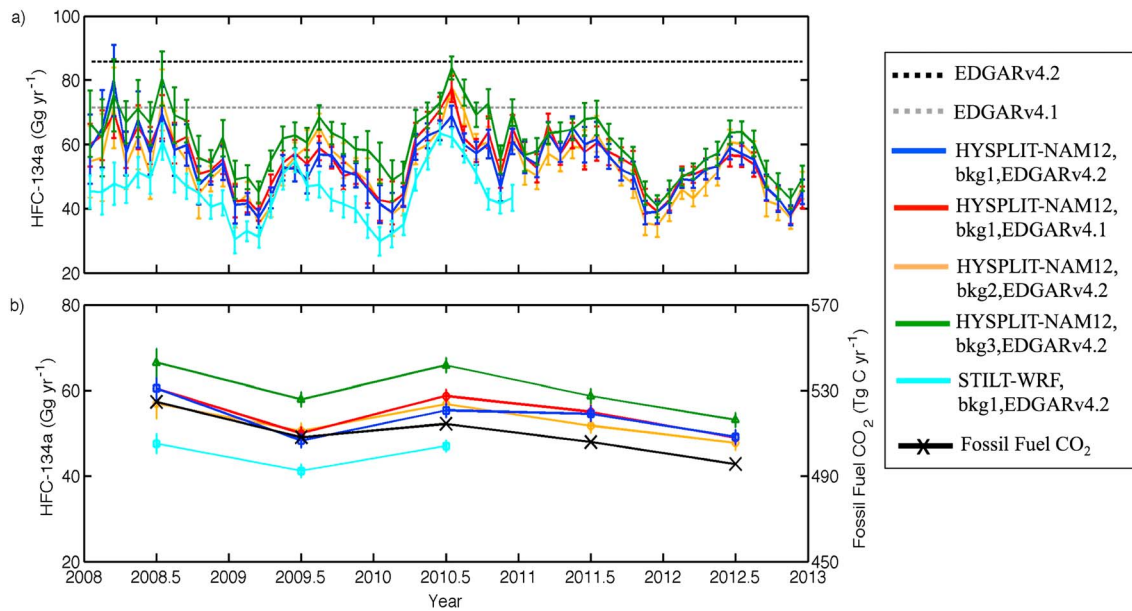
Figure 8. Posterior (a) monthly and (b) annual U.S. national emissions of HFC-134a derived from atmospheric observations and different priors (blue and red), different background (blue, orange, and green) and different transport models (blue and cyan). Dashed lines represent two priors from the EDGARv4.2 (black) and EDGARv4.1 (grey). Blue and red represent posterior fluxes derived from the two priors given background 1 (the 10th percentile + average free troposphere) and transport from HYPSLIT-NAM12. Orange and green stand for posterior fluxes that were derived based on background 2 (marine boundary reference) and background 3 (curtain + air back trajectories), given the EDGARv4.2 as a prior and transport calculated from HYSPLIT-NAM12. Cyan represents derived posterior fluxes with transport from STILT-WRF, background 1 and the EDGARv4.2 as a prior. Black crosses in panel Figure 8b represent national fossil fuel CO_2 emissions from transportation during 2008–2012 (in Tg C yr^{-1}) [US EIA, 2014] and are plotted relative to the right-hand y-axis.

calculation (i.e., the transport is perfectly known), whereas in reality the transport used in the inversion analysis will contain both random and systematic errors. These issues have led us to determine enhancements with different methods for estimating background mole fractions and to consider multiple models for simulating air transport in the inverse analysis of real atmospheric observations.

3.2. Real Data Inversion Using Actual Observations

3.2.1. National and Regional Emissions

Real data inversions were performed on the actual atmospheric measurements of HFC-134a using two different aseasonal prior emissions fields with no interannual variability through the entire study period. The two priors were based on 2008 annual emissions of HFC-134a from EDGARv4.2 and 2005 annual emissions from EDGARv4.1 (spatial distributions of these two priors are similar to the synthetic-true emission and Prior 4 in synthetic data experiments, Figure 3). The difference between derived annual national fluxes for the two priors is less than 6% despite the fact that these priors contain different spatial distributions (Figure 3) and annual national totals for the US that differ by approximately 20% (Figure 8). This result is consistent with our synthetic data inversion results that suggest national-scale posterior fluxes are relatively insensitive to the prior, given our observational network and analytical framework.

Compared to the small difference in posterior fluxes derived with two different priors, a larger difference was observed in derived national fluxes of HFC-134a with different background mole fractions or with footprints derived from different atmospheric transport models (Figure 8). When using the same prior (EDGARv4.2) and transport (HYSPLIT-NAM12), monthly posterior fluxes calculated from background 3 that considers the vertical gradient in background mole fractions and air transport history ("curtain + air back trajectories") are 10–20% higher than those derived with other methods (Figure 8a). Note that, using the 5th or 15th percentile of the surface data to derive background mole fractions result in emissions that are within 1% of emissions derived from the background determined by the 10th percentile (background 1). When using the same prior (EDGARv4.2) and the same background (in this case background 1), derived monthly national fluxes of HFC-134a from STILT-WRF are on average 20% lower than those derived using

Table 5. Estimated National and Regional Emissions for HFC-134a Averaged Over 2008–2012 and 2008–2010 (Gg yr^{-1})[a]

Regions	HYSPLIT-NAM12 (2008–2012)[b]	HYSPLIT-NAM12 (2008–2010)[c]	STILT-WRF (2008–2010)[c]
US total	52–61	54.7 ± 1.5	45.2 ± 1.2
US Northeast	9–12	11.5 ± 1.2	8.6 ± 0.7
US Southeast	7–11	8.4 ± 0.9	6.7 ± 0.8
US Central North	13–15	12.9 ± 0.5	9.8 ± 0.4
US Central South	12–14	12.8 ± 0.6	10.8 ± 0.5
US Mountain	3–5	3.3 ± 0.5	2.9 ± 0.5
US West	5–7	5.7 ± 0.7	6.4 ± 0.7

[a]Notes: Boundaries of various regions are shown in Figure 1.
[b]The ranges of estimated emissions represent uncertainties associated with prior emissions and background mole fractions that were derived from multiple inversion scenarios.
[c]Derived emissions are based on the 2008 HFC-134a emission field from EDGARv4.2 as a prior and background 1. Indicated uncertainties are 1σ errors of derived posterior emissions.

HYSPLIT-NAM12 (Figure 8 and Table 5). Despite these differences, posterior emissions derived using multiple priors, background mole fractions, and transport models consistently show seasonally varying emissions, which are 20–50% (the range represents the variation in different years) higher during summer than those in winter (Figure 8a).

One goal of NOAA's ongoing measurement program is to detect interannual changes or trends in trace gas emissions from the U.S. Inversion results derived from different background mole fractions, priors, and transport models all suggest that U.S. national emissions of HFC-134a were 10–20% lower in 2009 relative to the annual means for 2008 and 2010 (Figure 8b). We note that 2009 was the year of a U.S.-wide economic recession [*Gore*, 2010] and was apparently associated with a 2–3% decline in national fossil fuel CO_2 emissions from transportation [*US Energy Information Administration (US EIA)*, 2014] (Figure 8b). The correlation (R^2) between national annual HFC-134a emissions and transportation-related national annual fossil fuel CO_2 emissions during 2008–2012 is 0.7–0.9 based on the multiple inverse estimates. High correlation between these two is consistent with the fact that HFC-134a is mainly used in mobile air conditioning. As two thirds of HFC-134a emissions are estimated to come from the transportation sector [*US EPA*, 2014], the calculated absolute emissions of HFC-134a from transportation suggests that the global warming potential (GWP) associated with HFC-134a emissions from this sector is equivalent to about 3% of that for transportation-related CO_2 emissions on a 100 year time horizon [*Myhre et al.*, 2013]. Emissions associated with automobile air conditioning have been considered in the regulation of climate-related automotive emissions in some developed countries (e.g., the U.S. [*US EPA*, 2012] and the EU [*European Union*, 2006]). The climate-related influence of these emissions depends in large part on the GWP of the refrigeration fluid.

As with national emissions, derived posterior fluxes at regional scales are more sensitive to background mole fractions and transport uncertainties than to the different priors considered here (EDGARv4.1 and EDGARv4.2) (Figure 9). Derived regional emissions with different priors, background mole fractions, and multiple transport simulations are consistent within 50% (2σ relative to the mean) on an annual basis according to results from an ensemble of inversion scenarios (Figure 9). When compared to regional emissions of HFC-134a specified in the priors, our results indicated that EDGARv4.2 emissions are overestimated by up to a factor of 1–2 in the northeast and west regions for the year of 2008 (Figure 9). They also suggest a consistent seasonality in emissions from the northeast, southeast, central north, and west regions (Figure 9). In the central south and mountain regions, seasonality of HFC-134a emissions is small (Figure 9).

Considering multiple prior emissions and background values, we estimate 52–61 Gg yr^{-1} of HFC-134a emissions averaged for 2008–2012 from the contiguous U.S. with HYSPLIT-NAM12. On regional scales, we estimate 9–12 Gg yr^{-1}, 7–11 Gg yr^{-1}, 13–15 Gg yr^{-1}, 12–14 Gg yr^{-1}, 3–5 Gg yr^{-1}, and 5–7 Gg yr^{-1} from the northeast, southeast, central north, central south, mountain, and west regions, respectively (Table 5). When we use transport simulated from STILT-WRF for the years from 2008 to 2010, estimated emissions are 20% lower on a national scale, 10–30% lower in the northeast, southeast, central north, central south, and mountain regions and 12% higher in the west, compared to results derived from HYSPLIT-NAM12 for the same time period (Table 5).
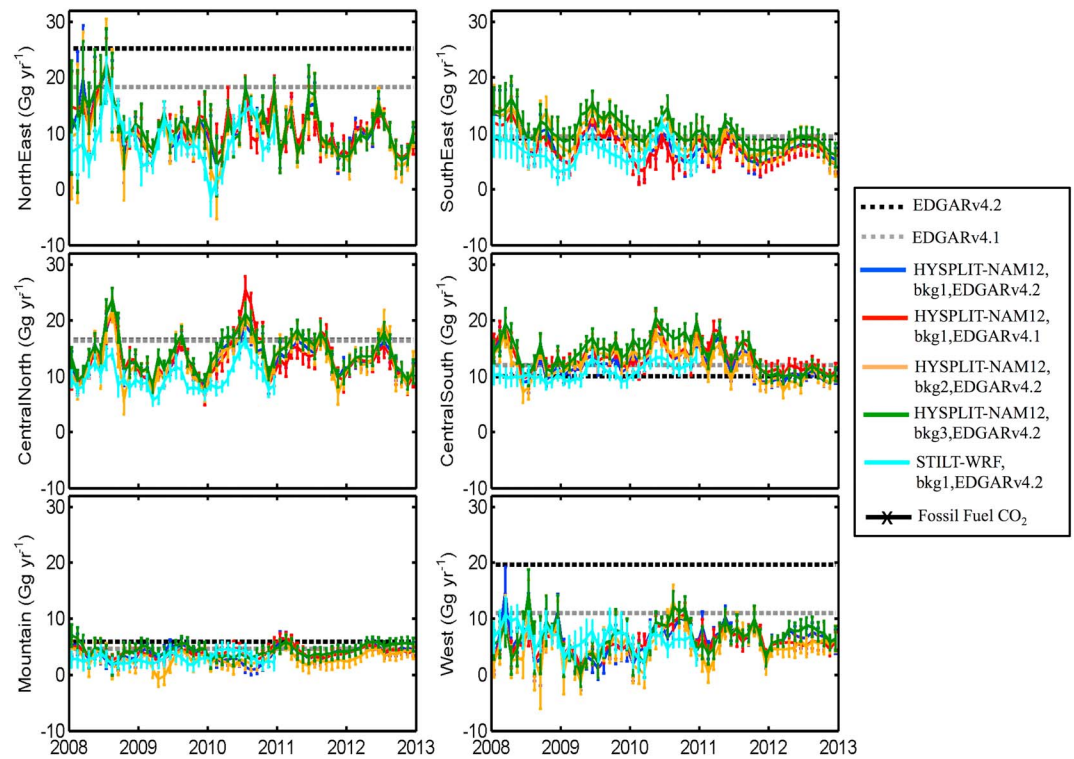
Figure 9. Monthly regional emissions of HFC-134a from posterior fluxes derived with different priors (blue and red), different background (blue, orange, and green) and different transport models (blue and cyan). Dashed lines represent two priors from the EDGARv4.2 (black) and EDGARv4.1 (grey). Blue and red represent posterior fluxes derived from the two priors, given background 1 (the 10th percentile + average free troposphere) and transport from HYPSLIT-NAM12. Orange and green stand for posterior fluxes that were derived based on background 2 (marine boundary reference) and background 3 (curtain + air back trajectories), given the EDGARv4.2 as a prior and transport from HYSPLIT-NAM12. Cyan represents derived posterior fluxes using transport from STILT-WRF, background 1 and the EDGARv4.2 as a prior.

3.2.2. Evaluation of Inversion Results With Actual Observations

Results from synthetic data experiments show that our inversion system is able to bring derived fluxes closer to the synthetic-true emission for various priors that have flux magnitudes and distributions very different from synthetic-true emission, given perfectly determined background mole fractions and transport. In inversions using actual observations, we apply other metrics to evaluate derived posterior fluxes (since true fluxes are not known), such as the difference between observed and simulated enhancements (expressed as root-mean-square errors or RMSE), the correlation between observed and simulated enhancements (R^2), and the slope of a best fit between observed and simulated enhancements.

When using HYSPLIT-NAM12, for data that were included in the inversion analysis, the site-specific RMSE of simulated enhancements using derived posterior fluxes is about 40 ($\pm$20, 1σ)% lower than the enhancements simulated using the prior EDGAR inventories (Figure 11). The correlation (R^2) between simulated and observed enhancements increases from 0.2 ($\pm$0.2, 1σ) for results derived from EDGAR inventories to 0.5 ($\pm$0.2, 1σ) for those derived from our posterior fluxes (Figures 10 and 11). Calculated site-specific slopes from a best fit between simulated and observed enhancements using an orthogonal distance regression [*Boggs and Rogers*, 1990] average 2.8 $\pm$ 2.8 (1σ) at individual sites when simulated enhancements derived from unoptimized EDGAR inventories, whereas they are 0.6 $\pm$ 0.3 (1σ) when using optimized posterior fluxes (Figures 10 and 11).

If we consider observations that were excluded from the inversion analysis (i.e., those from short-term aircraft campaigns and from 1000 to 3000 m agl at regular aircraft sites) and which therefore provide a quasi-independent assessment of the flux retrieval, the RMSE of simulated enhancements from derived posterior fluxes is about 30 ($\pm$20, 1σ)% smaller than that for the EDGAR inventories when using
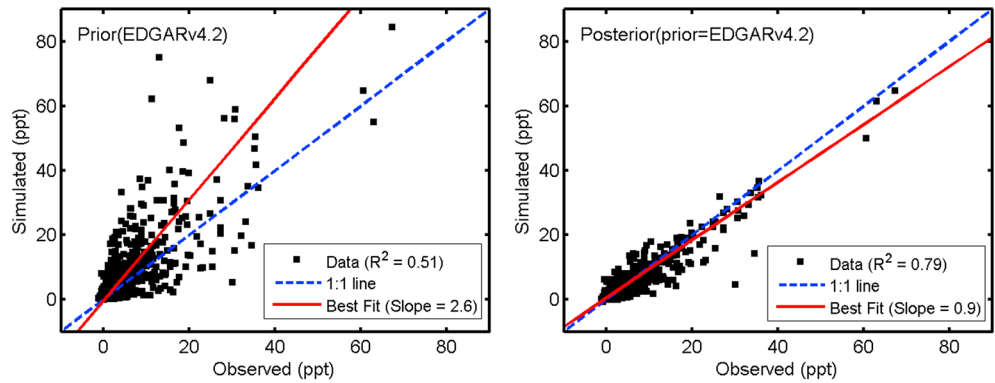
Figure 10. Simulated versus observed enhancements at AMT (black dots). Simulated enhancements were computed using the 2008 HFC-134a emission from (left) EDGARv4.2 and corresponding posterior fluxes derived from (right) this study. Red lines represent a best fit between simulated and observed enhancements using an orthogonal distance regression. Blue lines stand for the 1:1 line. Data plotted here were used in the inversion analysis.
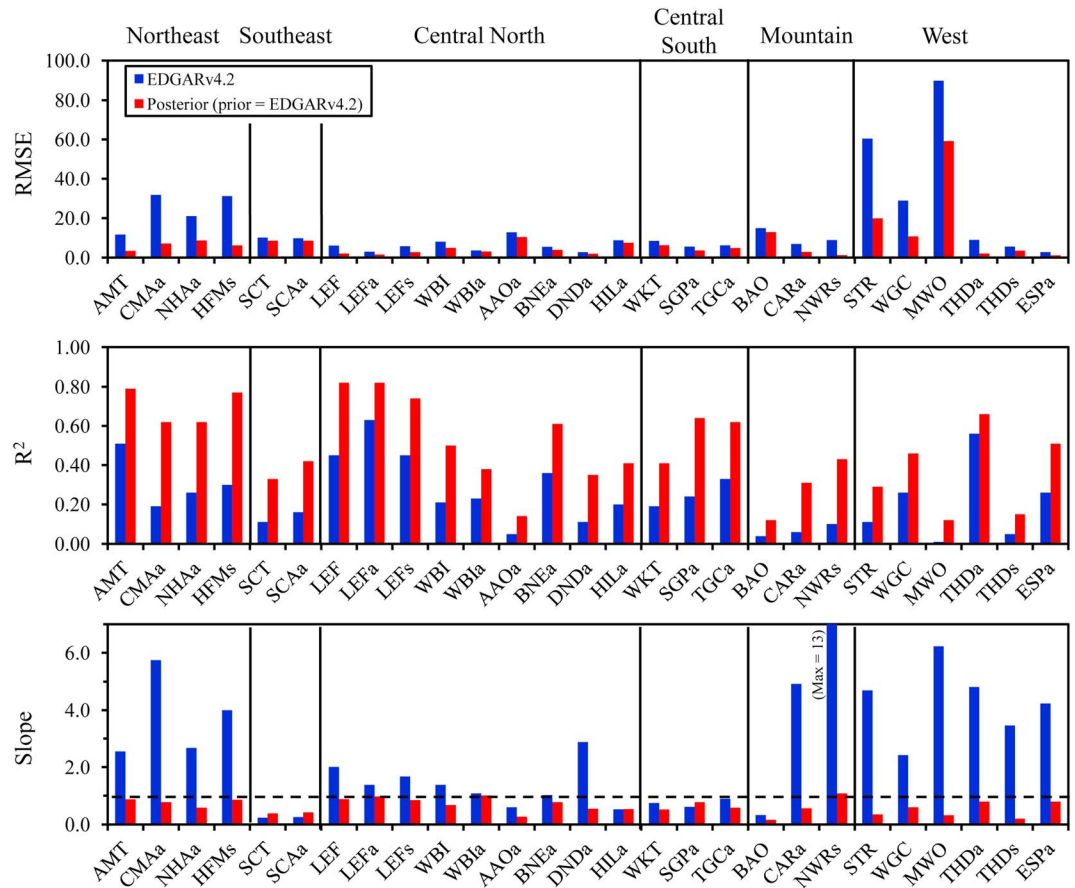
Figure 11. Root-mean-square errors (RMSE) of simulated enhancements, correlation between simulated and observed enhancements (R^2), and slopes from a best fit between simulated and observed enhancements for data used in inversion at daily glass flask sites (site code), weekly stainless steel flask sites (site code + "s") and aircraft profiling sites (site code + "a"). Blue bars represent results using simulated enhancements with the EDGARv4.2 inventory and red bars represent results using simulated enhancements with derived posterior fluxes. Sites are sorted by regions defined in Figure 1.
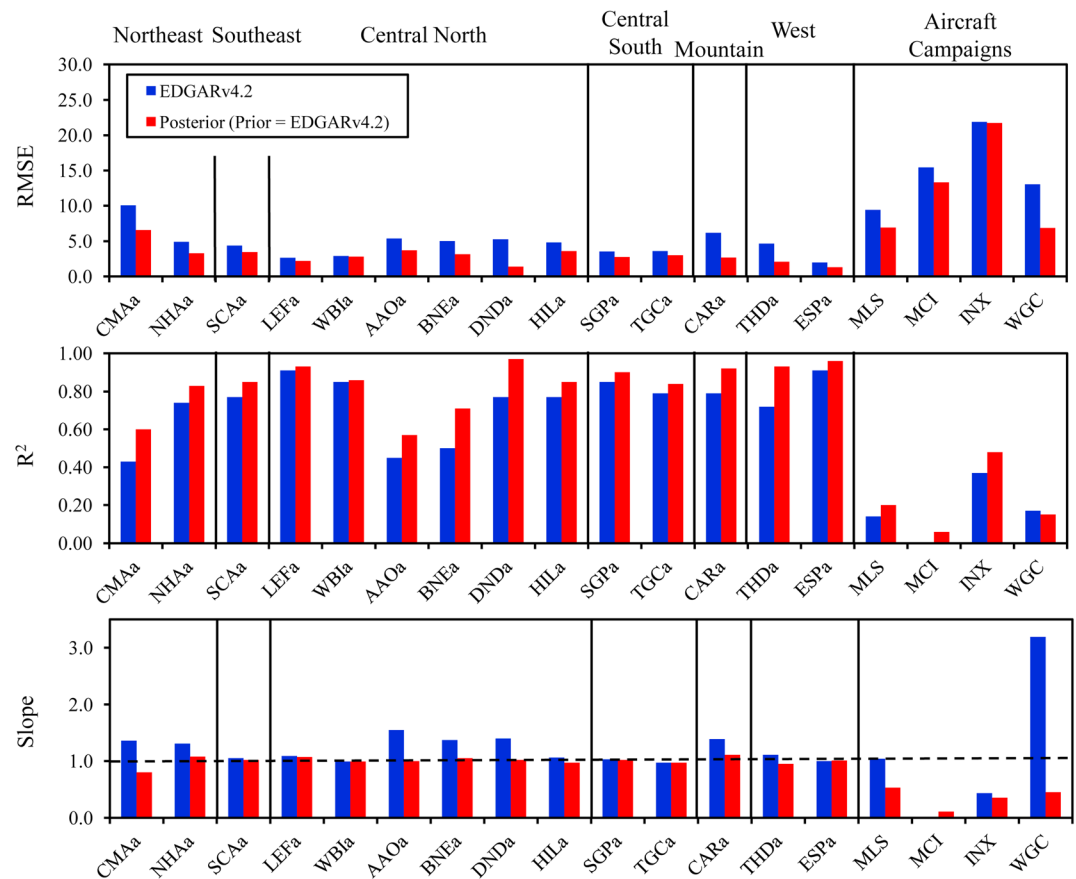
Figure 12. Root-mean-square errors (RMSE) of simulated enhancements, correlation between simulated and observed enhancements (R^2), and slopes from a best fit between simulated and observed enhancements for data excluded from inversion at aircraft profiling sites (site code + "a") and aircraft campaigns. Blue bars represent results using simulated enhancements with the EDGARv4.2 inventory, and red bars represent results using simulated enhancements with derived posterior fluxes. Sites are sorted by regions defined in Figure 1.

HYSPLIT-NAM12 (Figure 12). The correlation (R^2) between simulated and observed enhancements from derived posterior fluxes is improved by 20 ($\pm$20, 1σ)% relative to those for the EDGAR inventories (Figure 12). Slopes of a best fit between simulated and observed enhancements at all sites changed from 1.2 ± 0.6 (1σ) when simulated enhancements were calculated using the EDGAR inventories to 0.9 ± 0.3 (1σ) when using the optimized fluxes (Figure 12).

Results computed for both included and excluded observations and using STILT-WRF transport also consistently show reduced RMSE of simulated enhancements, higher correlation between simulated and observed enhancements, and a slope closer to 1 from the best fit between simulated and observed enhancements with derived posterior fluxes, compared to those derived with unoptimized inventories (EDGARv4.2 and EDGARv4.1). These metrics (RMSE, R^2 and slope) suggest that fluxes derived from our inversion system were significantly improved relative to those reported by EDGAR inventories. Since our system optimizes enhancements that were obtained by subtracting estimated background mole fractions from observations, any systematic errors in background mole fractions may present themselves as systematic offsets between posterior fluxes derived using different background values (Figures 8 and 9). However, the metrics considered here do not vary appreciably in results derived with the different background mole fractions. Thus, they do not allow any assessment on the systematic errors associated with these different background approaches.

To assess transport errors, we compared simulated and observed mole fractions from aircraft profiles using the two different transport models. Simulated mole fractions averaged over all profiles or profiles at each site for optimized fluxes derived from both transport models display vertical variation that agrees with the
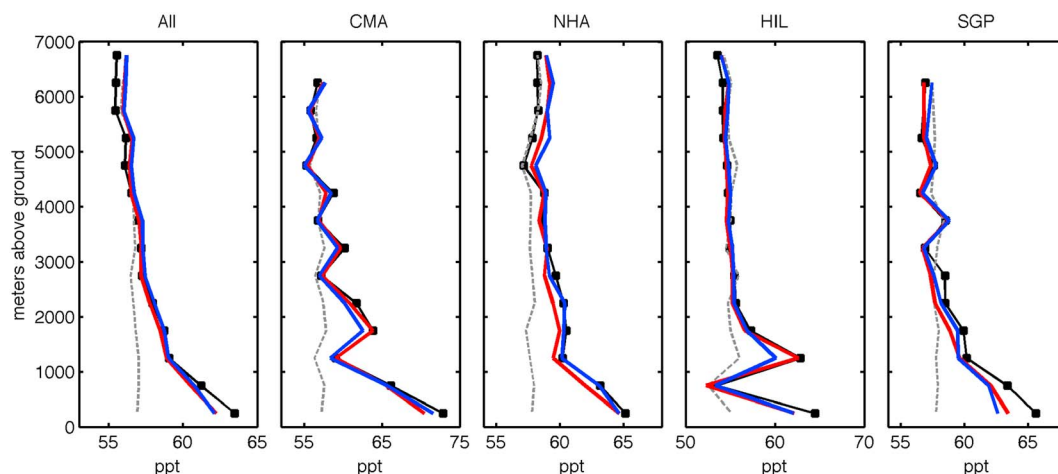
Figure 13. Averaged aircraft profiles from (first panel) all sites and (second to fifth panels) selected sites during 2008–2010 using observed mole fractions (black squares connected with a solid line) and simulated mole fractions that were calculated with posterior fluxes, footprints from HYSPLIT-NAM12 (red solid line) and STILT-WRF (blue solid line) and background 1 (gray dashed line). Only data between 1000 and 3000 m agl are independent data that were excluded from inversion. Data below 1000 m agl were used in the inversion calculation and data above 3000 m agl were used to derive background mole fractions at these sites.

observations (Figure 13), implying that vertical mixing is represented reasonably well in both transport models. In addition, the averaged RMSE, R^2 and slope for simulated and observed mole fractions from all sites are comparable for results calculated from both transport models using data included and excluded in the inversion (e.g., RMSE averaged over all sites for data excluded from the inversion is 3.3 ($\pm$ 2.2, 1σ) ppt with HYSPLIT-NAM12, whereas it is 3.5 ($\pm$ 2.2, 1σ) ppt with STILT-WRF). However, there is a noticeable difference in the performance between the two transport models at different locations. For example, for sites in the northeast and west coast (i.e., CMA, MWO, and THD), the MLE-derived model-data mismatch errors using STILT-WRF are 10–35% lower than those with HYSPLIT-NAM12 (Table 4), suggesting, at these sites, the transport calculated from STILT-WRF is more realistic, whereas at sites in the central area (i.e., LEF, WKT, SGP, and HIL), the transport calculated from HYSPLIT-NAM12 yields 10–40% lower model-data mismatch errors (Table 4). Furthermore, MLE-derived model-data mismatch errors from both transport models relative to observed enhancements are significantly higher at BAO, MWO, and STR (Table 4) where transport is difficult to model due to complex terrain and/or the presence of sea breeze. At THD and ESP, the relative model-data mismatch errors (to enhancements) are also large (Table 4). This is likely due to the relatively small enhancements measured at these sites.

3.2.3. Comparison With Other National Emission Estimates

Figure 14 summarizes national emission estimates of HFC-134a from inventory-based bottom-up and atmosphere-based top-down results. The two inventory-based bottom-up estimates (from EDGARv4.2 and U.S. EPA) are based on national activity data and associated emission factors. Due to an incomplete understanding of emission processes and the size of existing reservoirs [*US EPA*, 2014], uncertainties in national emissions from bottom-up estimates are not well determined and therefore not provided by either inventory. The difference in annual national total emissions between the two inventories in 2008 is as large as 60% relative to U.S. EPA estimate (Figure 14).

Multiple approaches for deriving HFC-134a emissions from atmospheric data have been used previously, including a "CO"-based tracer ratio method [*Barletta et al.*, 2011; *Millet et al.*, 2009], a "$^{14}CO_2$"-based tracer ratio method [*Miller et al.*, 2012], and a Bayesian inversion [*Stohl et al.*, 2009]. The difference among different top-down estimates in the U.S. national emission rate they derive is about 100% (relative to the lowest estimate [*Millet et al.*, 2009]) and the discrepancy between bottom-up and top-down estimates is as large as a factor of 2 (relative to the lowest estimate [*Millet et al.*, 2009]). Besides the fundamental differences in methodology for deriving fluxes from measured atmospheric mole fractions, some of the differences in the derived national estimates are likely attributable to the spatial and temporal extrapolations required in these short-duration or localized studies for estimating national-scale fluxes. For example, data used in the study of
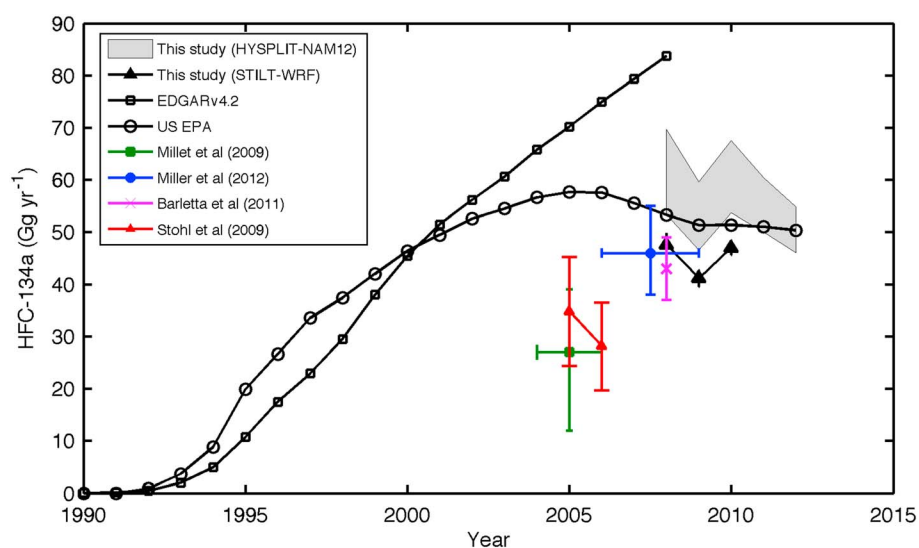
Figure 14. U.S. HFC-134a emission estimates derived in this work compared to previous top-down studies and existing emission inventories. The grey area shows the estimated possible range of emissions derived with transport from HYSPLIT-NAM12 using different prior emission fields and background mole fractions (shown in Figure 8). Black-filled triangles show our derived emissions with transport from STILT-WRF, given background determined from background approach (1) and the EDGARv4.2 as a prior. Black circles and black squares are HFC-134a emissions reported by U.S. EPA [*US EPA*, 2014], and EDGARv4.2 (http://edgar.jrc.ec.europa.eu/overview.php?v=42). Green, red, magenta, and blue symbols represent emissions estimated by *Millet et al.* [2009], *Stohl et al.* [2009], *Barletta et al.* [2011], and *Miller et al.* [2012], respectively.

Millet et al. [2009] were mostly collected in July–August of 2004 and September–October of 2006 over the northeast and southeast U.S. during short-term aircraft campaigns; U.S. emission of HFC-134a estimated by *Stohl et al.* [2009] was constrained only by atmospheric data at a remote site in the northern coast of California from January 2005 to March 2007; *Barletta et al.* [2011] extrapolated their data from aircraft campaigns during June 2008 over the South Coast Air Basin of California to the entire U.S. to provide a U.S. national emission of HFC-134a; and the emission estimate given by *Miller et al.* [2012] was based on year-round data collected at two sites in the northeast of the U.S. from 2006 to 2009.

In this study, we estimate U.S. emission of HFC-134a averaged over 2008–2012 at 52–61 Gg yr^{-1} (Figure 14) using an ensemble of inversion scenarios that consider uncertainties from prior emission fields and background, with transport calculated from HYSPLIT-NAM12 (Table 5). This is highly consistent with the U.S. EPA estimate (50–53 Gg yr^{-1}) [*US EPA*, 2014] for these same years but 30–40% lower than the estimate from the EDGARv4.2 inventory for 2008 (86 Gg yr^{-1}) (Figure 14). In addition, estimates from this study and U.S. EPA consistently suggest no significant emission trend of HFC-134a during 2008–2012 (Figure 14). The national total HFC-134a emission reported by U.S. EPA is based on a "Vintaging Model" [*Godwin*, 2012] that considers motor vehicle sales and registration data, an estimate of a refrigerant charge size by vehicle make, model and year, an estimated emission factor derived from limited data [*Atkinson*, 2008; *Clodic et al.*, 2010; *Vincent et al.*, 2004], and emissions for other lesser (~30%) applications of HFC-134a (i.e., foam blowing and aerosol propellant). The consistency of our atmosphere-based top-down estimate and U.S. EPA inventory-based bottom-up estimate implies that estimated HFC-134a emission factors and activities from the bottom-up U.S. EPA inventory are broadly correct or contain compensating errors.

In comparison to other atmosphere-based top-down studies for deriving national-scale emissions, we have incorporated data that have the most extensive coverage in both time and space across the contiguous U.S. during a 5 year period. Results from this air network should contain emission signals from most all regions in all four seasons. Therefore, annual national emissions determined in the present study are likely to be more representative of total U.S. emissions than previous top-down studies mentioned above. Because our analysis provides spatially and temporally resolved fluxes, our results can be used to quantify the potential biases that might exist in an annual national emission estimate based on data collected from

localized regions or short-duration aircraft campaigns. For example, our results suggest that when using data collected in California and extrapolating them to the entire U.S. based on population [e.g., *Barletta et al.*, 2011], the calculated U.S. national emissions of HFC-134a could be underestimated by 30–40% as the calculated flux per capita we derive for the west region is 30–40% lower than the national average. When using data collected in summer and extrapolating to the whole year, the derived annual national HFC-134a emission could be overestimated by 20–50%.

4. Conclusions

A regional inverse model was developed to quantify regional fluxes of long-lived trace gases. This inversion system uses a maximum likelihood estimation (MLE) to objectively determine model-data mismatch errors and a prior flux error covariance matrix thus avoiding potential biases with subjective assumptions about these uncertainties. Results from synthetic data experiments suggest that the model-data mismatch errors given by MLE are dependent on the prior emission distribution. When the prior has a distribution that is similar to that in the synthetic-true emission, calculated model-data mismatch errors from MLE were found to be very consistent with the magnitude of random noise added to the synthetic observations. Results from synthetic data experiments also suggest that there are certain differences between the MLE-derived (atmosphere-derived) and variogram-derived (flux-derived) parameters in prior flux error covariance. Inaccuracies in MLE-derived parameters may result in certain biases and unrepresented errors in derived posterior fluxes, especially when the prior has a distribution that is very different from the synthetic-true emission.

Given our air sampling network and the available measurements made during 2008–2012, annual national fluxes for HFC-134a derived with an optimized inversion system are fairly insensitive to the assumed prior emission magnitude and distribution. This was evident from both the synthetic data and real data inversions. The magnitudes of errors of derived posterior fluxes strongly correlate with the number of available independent observations with a poor prior that has a constant emission (i.e., Priors 2 and 3). Furthermore, derived fluxes for regions with fewer observations (therefore lower footprint intensity) or where transport is difficult to model (thus higher model-data mismatch errors relative to enhancements) show higher dependence on prior fluxes. These results highlight the importance of well-distributed and frequent atmospheric observations for deriving accurate fluxes on national and regional scales.

Derived fluxes with multiple combinations of prior emission fields, background mole fractions, and transport simulations are consistent within 25% on an annual national scale and within 50% on an annual regional scale (at a 95% confidence interval). All inversion scenarios suggest emissions of HFC-134a during summer that are 20–50% greater than emissions during winter. Derived national emissions of HFC-134a averaged over 2008–2012 were in a range of 52–61 Gg yr^{-1} with air transport calculated from HYSPLIT-NAM12. Estimated national emissions during 2008–2010 were 20% lower when specifying alternative transport from STILT-WRF. Using data that were included or excluded in inversions suggests, overall, simulated mole fractions with transport from HYSPLIT-NAM12 and STILT-WRF agree comparably well with observed mole fractions. However, the performance of one transport model could be noticeably better than the other at different locations.

Estimated U.S. national emissions during 2008–2012 from this study, 52–61 Gg yr^{-1} (with transport from HYSPLIT-NAM12), agree with the reported emissions from U.S. EPA (50–53 Gg yr^{-1}) for the same time period. They are about 30–40% lower than the emission from the EDGARv4.2 inventory for 2008 (the most recent year for which EDGAR emissions are available). This is mainly due to an overestimated emission by a factor of 1 – 2 by EDGARv4.2 in the northeast and west regions. Large discrepancies were observed between inventories (i.e., U.S. EPA and EDGARv4.2) and some previous atmosphere-based top-down estimates (Figure 14). This is likely due to errors arising from spatial and temporal extrapolations that were required to scale results from short-duration aircraft campaigns or localized regional studies to national totals. When we use data from multiple sites across the U.S. for multiple years that requires minimal extrapolations, national emissions we derive are consistent with the U.S. EPA inventory. Furthermore, our estimated absolute annual HFC-134a emissions, along with independent estimates of the fraction of HFC-134a emissions from the transportation sector, suggest that the global warming potential (GWP) associated with recent HFC-134a emissions from transportation is equivalent to about 3% of that for transportation-related CO_2 emissions (for an assumed 100 year time horizon).

References

Andrews, A. E., et al. (2014), CO_2, CO, and CH_4 measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: Instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts, *Atmos. Meas. Tech.*, 7, 647–687, doi:10.5194/amt-7-647-2014.

Ashford, P., D. Clodic, A. McCulloch, and L. Kuijpers (2004), Emission profiles from the foam and refrigeration sectors comparison with atmospheric concentrations. Part 1: Methodology and data, *Int. J. Refrig.*, 27(7), 687–700, doi:10.1016/j.ijrefrig.2004.07.025.

Atkinson, W. (2008), Presentation from the 2008 Mobile Air Conditioning Leadership Summit, Scottsdale, Ariz. [Available at http://www.epa.gov/cpd/mac/Atkinson%202008%20MAC%20summit.pdf.]

Barletta, B., P. Nissenson, S. Meinardi, D. Dabdub, F. Sherwood Rowland, R. A. VanCuren, J. Pederson, G. S. Diskin, and D. R. Blake (2011), HFC-152a and HFC-134a emission estimates and characterization of CFCs, CFC replacements, and other halogenated solvents measured during the 2008 ARCTAS campaign (CARB phase) over the South Coast Air Basin of California, *Atmos. Chem. Phys.*, 11(6), 2655–2669, doi:10.5194/acp-11-2655-2011.

Barletta, B., et al. (2013), Emission estimates of HCFCs and HFCs in California from the 2010 CalNex study, *J. Geophys. Res. Atmos.*, 118, 2019–2030, doi:10.1002/jgrd.50209.

Biraud, S. C., M. S. Torn, J. R. Smith, C. Sweeney, W. J. Riley, and P. P. Tans (2013), A multi-year record of airborne CO_2 observations in the U.S. Southern Great Plains, *Atmos. Meas. Tech.*, 6, 751–763, doi:10.5194/amt-6-751-2013.

Boggs, P. T., and J. E. Rogers (1990), Orthogonal distance regression, *Contemp. Math.*, 112, 183–194.

Bruhwiler, L. M. P., A. M. Michalak, W. Peters, D. F. Baker, and P. Tans (2005), An improved Kalman Smoother for atmospheric inversions, *Atmos. Chem. Phys.*, 5(10), 2691–2702, doi:10.5194/acp-5-2691-2005.

Brunner, D., S. Henne, C. A. Keller, S. Reimann, M. K. Vollmer, S. O'Doherty, and M. Maione (2012), An extended Kalman-filter for regional scale inverse emission estimation, *Atmos. Chem. Phys.*, 12(7), 3455–3478, doi:10.5194/acp-12-3455-2012.

Clodic, D., S. Barrault, and S. Saba (2010), Global Inventories of the worldwide fleets of refrigerating and air conditioning equipment in order to determine refrigerant emissions. The 1990 to 2006 updating. Extracts from the Final Report. Centre Energetique et Procedes, Mines ParisTech, Armines, Paris, France. [Available at http://www.ereie-sas.fr/docs/Short%20report%20-%20complete.pdf.]

Cressie, N., and D. Hawkins (1980), Robust estimation of the variogram: I, *Math. Geol.*, 12(2), 115–125, doi:10.1007/bf01035243.

Daniel, J. S., G. J. M. Velders, S. Solomon, M. McFarland, and S. A. Montzka (2007), Present and future sources and emissions of halocarbons: Toward new constraints, *J. Geophys. Res.*, 112, D02301, doi:10.1029/2006JD007275.

Daniel, J. S., et al. (2011), A focus on information options for policymakers, Chapter 5, in *Scientific Assessment of Ozone Depletion 2010*, Global Ozone Research and Monitoring Project – Report No. 52, 516 pp, World Meteorological Organization, Geneva, Switzerland.

Draxler, R. R. (1999), HYSPLIT4 user's guide, *NOAA Tech. Memo. ERL ARL-230,* NOAA Air Resources Laboratory, Silver Spring, Md.

Draxler, R. R., and G. D. Hess (1997), Description of the HYSPLIT4 modeling system, *NOAA Tech. Memo. ERL ARL-224,* 24 pp., NOAA Air Resources Laboratory, Silver Spring, Md.

Draxler, R. R., and G. D. Hess (1998), An overview of the HYSPLIT4 modeling system of trajectories, dispersion, and deposition, *Aust. Meteorol. Mag.*, 47, 295–308.

European Union (2006), DIRECTIVE 2006/40/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 17 May 2006 relating to emissions from air-conditioning systems in motor vehicles and amending Council Directive 70/156/EEC. [Available at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:161:0012:0018:en:PDF.]

Ganesan, A. L., et al. (2014), Characterization of uncertainties in atmospheric trace gas inversions using hierarchical Bayesian methods, *Atmos. Chem. Phys.*, 14(8), 3855–3864, doi:10.5194/acp-14-3855-2014.

Geels, C., et al. (2007), Comparing atmospheric transport models for future regional inversions over Europe—Part 1: mapping the atmospheric CO_2 signals, *Atmos. Chem. Phys.*, 7(13), 3461–3479, doi:10.5194/acp-7-3461-2007.

Gerbig, C., S. Körner, and J. C. Lin (2008), Vertical mixing in atmospheric tracer transport models: Error characterization and propagation, *Atmos. Chem. Phys.*, 8(3), 591–602, doi:10.5194/acp-8-591-2008.

Godwin, D. S. (2012), Demand for ozone-depleting substances and hydrofluorocarbons estimated by a Tier 2 emission inventory model compared to top-down chemical consumption data for the US, *J. Integr. Environ. Sci.*, 9(sup1), 81–95, doi:10.1080/1943815X.2012.693090.

Gore, C. (2010), The global recession of 2009 in a long-term development perspective, *J. Int. Dev.*, 22(6), 714–738, doi:10.1002/jid.1725.

Gourdji, S. M., A. I. Hirsch, K. L. Mueller, V. Yadav, A. E. Andrews, and A. M. Michalak (2010), Regional-scale geostatistical inverse modeling of North American CO2 fluxes: A synthetic data study, *Atmos. Chem. Phys.*, 10(13), 6151–6167, doi:10.5194/acp-10-6151-2010.

Gourdji, S. M., et al. (2012a), North American CO2 exchange: Inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, *Biogeosciences*, 9(1), 457–475, doi:10.5194/bg-9-457-2012.

Gourdji, S. M., et al. (2012b), North American CO2 exchange: Inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, *Biogeosciences*, 9(1), 457–475, doi:10.5194/bg-9-457-2012.

Gschrey, B., W. Schwarz, C. Elsner, and R. Engelhardt (2011), High increase of global F-gas emissions until 2050, *Greenhouse Gas Meas. Manage.*, 1(2), 85–92, doi:10.1080/20430779.2011.579352.

Hall, B. D., et al. (2014), Results from the International Halocarbons in Air Comparison Experiment (IHALACE), *Atmos. Meas. Tech.*, 7(2), 469–490, doi:10.5194/amt-7-469-2014.

Hegarty, J., R. R. Draxler, A. F. Stein, J. Brioude, M. Mountain, J. Eluszkiewicz, T. Nehrkorn, F. Ngan, and A. Andrews (2013), Evaluation of Lagrangian particle dispersion models with measurements from controlled tracer releases, *J. Appl. Meteorol. Climatol.*, 52(12), 2623–2637, doi:10.1175/jamc-d-13-0125.1.

Jeong, S., C. Zhao, A. E. Andrews, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012a), Seasonal variation of CH4 emissions from central California, *J. Geophys. Res.*, 117, D11306, doi:10.1029/2011JD016896.

Jeong, S., C. Zhao, A. E. Andrews, E. J. Dlugokencky, C. Sweeney, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012b), Seasonal variations in N2O emissions from central California, *Geophys. Res. Lett.*, 39, L16805, doi:10.1029/2012GL052307.

Jeong, S., Y.-K. Hsu, A. E. Andrews, L. Bianco, P. Vaca, J. M. Wilczak, and M. L. Fischer (2013), A multitower measurement network estimate of California's methane emissions, *J. Geophys. Res. Atmos.*, 118, 11,339–11,351, doi:10.1002/jgrd.50854.

Kaminski, T., P. J. Rayner, M. Heimann, and I. G. Enting (2001), On aggregation errors in atmospheric transport inversions, *J. Geophys. Res.*, 106(D5), 4703–4715, doi:10.1029/2000JD900581.

Lin, J. C., C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, K. J. Davis, and C. A. Grainger (2003), A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, *J. Geophys. Res.*, 108(D16), 4493, doi:10.1029/2002JD003161.

Lin, J. C., C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, C. A. Grainger, B. B. Stephens, P. S. Bakwin, and D. Y. Hollinger (2004), Measuring fluxes of trace gases at regional scales by Lagrangian observations: Application to the CO_2 Budget and Rectification Airborne (COBRA) study, *J. Geophys. Res.*, 109, D15304, doi:10.1029/2004JD004754.

Maione, M., et al. (2014), Estimates of European emissions of methyl chloroform using a Bayesian inversion method, *Atmos. Chem. Phys., 14*, 9755–9770, doi:10.5194/acp-14-9755-2014.

Manning, A. J., D. B. Ryall, R. G. Derwent, P. G. Simmonds, and S. O'Doherty (2003), Estimating European emissions of ozone-depleting and greenhouse gases using observations and a modeling back-attribution technique, *J. Geophys. Res., 108*(D14), 4405, doi:10.1029/2002JD002312.

Masarie, K. A., and P. P. Tans (1995), Extension and integration of atmospheric carbon dioxide data into a globally consistent measurement record, *J. Geophys. Res., 100*(D6), 11,593–11,610, doi:10.1029/95JD00859.

McCulloch, A., P. M. Midgley, and P. Ashford (2003), Releases of refrigerant gases (CFC-12, HCFC-22 and HFC-134a) to the atmosphere, *Atmos. Environ., 37*(7), 889–902, doi:10.1016/S1352-2310(02)00975-5.

Michalak, A. M., L. Bruhwiler, and P. P. Tans (2004), A geostatistical approach to surface flux estimation of atmospheric trace gases, *J. Geophys. Res., 109*, D14109, doi:10.1029/2003JD004422.

Michalak, A. M., A. Hirsch, L. Bruhwiler, K. R. Gurney, W. Peters, and P. P. Tans (2005), Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, *J. Geophys. Res., 110*, D24107, doi:10.1029/2005JD005970.

Miller, J. B., et al. (2012), Linking emissions of fossil fuel CO_2 and other anthropogenic trace gases using atmospheric $^{14}CO_2$, *J. Geophys. Res., 117*, D08302, doi:10.1029/2011JD017048.

Miller, S. M., et al. (2013), Anthropogenic emissions of methane in the United States, *Proc. Natl. Acad. Sci. U.S.A., 110*, 20,018–20,022, doi:10.1073/pnas.1314392110.

Millet, D. B., E. L. Atlas, D. R. Blake, N. J. Blake, G. S. Diskin, J. S. Holloway, R. C. Hudman, S. Meinardi, T. B. Ryerson, and G. W. Sachse (2009), Halocarbon emissions from the United States and Mexico and their global warming potential, *Environ. Sci. Technol., 43*(4), 1055–1060, doi:10.1021/es802146j.

Montzka, S. A., R. C. Myers, J. H. Butler, J. W. Elkins, L. T. Lock, A. D. Clarke, and A. H. Goldstein (1996), Observations of HFC-134a in the remote troposphere, *Geophys. Res. Lett., 23*(2), 169–172, doi:10.1029/95GL03590.

Montzka, S. A., J. H. Butler, J. W. Elkins, T. M. Thompson, A. D. Clarke, and L. T. Lock (1999), Present and future trends in the atmospheric burden of ozone-depleting halogens, *Nature, 398*(6729), 690–694.

Montzka, S. A., et al. (2011), Ozone-Depleting Substances (ODSs) and Related Chemiscals, Chapter 1, in *Scientific Assessment of Ozone Depletion 2010*, Global Ozone Research and Monitoring Project – Report No. 52, 516 pp., World Meteorological Organization, Geneva, Switzerland.

Montzka, S. A., M. McFarland, S. O. Andersen, B. R. Miller, D. W. Fahey, B. D. Hall, L. Hu, C. Siso, and J. W. Elkins (2014), Recent trends in global emissions of hydrochlorofluorocarbons and hydrofluorocarbons: Reflecting on the 2007 adjustments to the Montreal Protocol, *J. Phys. Chem. A*, doi:10.1021/jp5097376.

Myhre, G., et al. (2013), Anthropogenic and natural radiative forcing, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker et al., Cambridge Univ. Press, Cambridge, U. K., and New York.

Nehrkorn, T., J. Eluszkiewicz, S. Wofsy, J. Lin, C. Gerbig, M. Longo, and S. Freitas (2010), Coupled weather research and forecasting–stochastic time-inverted Lagrangian transport (WRF–STILT) model, *Meteorol. Atmos. Phys., 107*(1–2), 51–64, doi:10.1007/s00703-010-0068-x.

O'Doherty, S., et al. (2004), Rapid growth of hydrofluorocarbon 134a and hydrochlorofluorocarbons 141b, 142b, and 22 from Advanced Global Atmospheric Gases Experiment (AGAGE) observations at Cape Grim, Tasmania, and Mace Head, Ireland, *J. Geophys. Res., 109*, D06310, doi:10.1029/2003JD004277.

O'Doherty, S., et al. (2009), Global and regional emissions of HFC-125 (CHF2CF3) from in situ and air archive atmospheric observations at AGAGE and SOGE observatories, *J. Geophys. Res., 114*, D23304, doi:10.1029/2009JD012184.

O'Doherty, S., et al. (2014), Global emissions of HFC-143a (CH3CF3) and HFC-32 (CH2F2) from in situ and air archive atmospheric observations, *Atmos. Chem. Phys., 14*(17), 9249–9258, doi:10.5194/acp-14-9249-2014.

Peters, W., J. B. Miller, J. Whitaker, A. S. Denning, A. Hirsch, M. C. Krol, D. Zupanski, L. Bruhwiler, and P. P. Tans (2005), An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations, *J. Geophys. Res., 110*, D24304, doi:10.1029/2005JD006157.

Peters, W., et al. (2007), An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, *Proc. Natl. Acad. Sci. U.S.A., 104*(48), 18,925–18,930, doi:10.1073/pnas.0708986104.

Powell, B. (2013), Treating nonlinearities in data-space variational assimilation, in *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, vol. II, edited by S. K. Park and L. Xu, pp. 233–250, Springer, Berlin.

Rodgers, C. D. (2000), *Inverse Methods for Atmospheric Sounding*, World Science, Tokyo.

Saikawa, E., et al. (2012), Global and regional emission estimates for HCFC-22, *Atmos. Chem. Phys., 12*(21), 10,033–10,050, doi:10.5194/acp-12-10033-2012.

Saikawa, E., et al. (2014), Corrigendum to "Global and regional emission estimates for HCFC-22," Atmos. Chem. Phys., 12, 10033–10050, 2012, *Atmos. Chem. Phys., 14*(10), 4857–4858, doi:10.5194/acp-14-4857-2014.

Stohl, A., et al. (2009), An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons, *Atmos. Chem. Phys., 9*(5), 1597–1620, doi:10.5194/acp-9-1597-2009.

Umweltbundesamt (2009), Projections of global emissions of fluorinated greenhouse gases in 2050, by Gschrey B., W. Schwarz, Öko-Recherche, on behalf of the German Federal Environmental Agency (Umweltbundesamt), Report-no. UBA-FB001318.

United States Environmental Protection Agency (US EPA) (2012), EPA and NHTSA set standards to reduce greenhouse gases and improve fuel economy for model years 2017–2025 cars and light trucks, EPA-420-F-12-051. [Available at http://www.epa.gov/otaq/climate/documents/420f12051.pdf.]

United States Environmental Protection Agency (US EPA) (2014), Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2012, EPA 430-R-14-003.

US Energy Information Administration (US EIA) (2014), Monthly Energy Review *Rep. DOE/EIA-0035*(*2014/01*).

Velders, G. J. M., D. W. Fahey, J. S. Daniel, M. McFarland, and S. O. Andersen (2009), The large contribution of projected HFC emissions to future climate forcing, *Proc. Natl. Acad. Sci. U.S.A., 106*(27), 10,949–10,954, doi:10.1073/pnas.0902817106.

Vincent, R., K. Cleary, A. Ayala, and R. Corey (2004), Emissions of HFC-134a from light-duty vehicles in California, SAE International Technical Paper 2004-01-2256, doi:10.4271/2004-01-2256.

Vollmer, M. K., et al. (2011), Atmospheric histories and global emissions of the anthropogenic hydrofluorocarbons HFC-365mfc, HFC-245fa, HFC-227ea, and HFC-236fa, *J. Geophys. Res., 116*, D08304, doi:10.1029/2010JD015309.

Winiwarter, W., and K. Rypdal (2001), Assessing the uncertainty associated with national greenhouse gas emission inventories: A case study for Austria, *Atmos. Environ., 35*(32), 5425–5440, doi:10.1016/S1352-2310(01)00171-6.

Yadav, V., and A. M. Michalak (2013), Improving computational efficiency in large linear inverse problems: An example from carbon dioxide flux estimation, *Geosci. Model Dev., 6*(3), 583–590, doi:10.5194/gmd-6-583-2013.

Yokouchi, Y., S. Taguchi, T. Saito, Y. Tohjima, H. Tanimoto, and H. Mukai (2006), High frequency measurements of HFCs at a remote site in east Asia and their implications for Chinese emissions, *Geophys. Res. Lett., 33*, L21814, doi:10.1029/2006GL026403.

Zhao, C., A. E. Andrews, L. Bianco, J. Eluszkiewicz, A. Hirsch, C. MacDonald, T. Nehrkorn, and M. L. Fischer (2009), Atmospheric inverse estimates of methane emissions from Central California, *J. Geophys. Res., 114*, D16302, doi:10.1029/2008JD011671.