Designing optimal greenhouse gas observing networks that consider performance and cost

D. D. Lucas[1], C. Yver Kwok[2], P. Cameron-Smith[1], H. Graven[3,4], D. Bergmann[1], T. P. Guilderson[1], R. Weiss[4], and R. Keeling[4]

[1]Lawrence Livermore National Laboratory, Livermore, CA 94550, USA
[2]Laboratoire des Sciences du Climat et de l'Environnement, Gif-sur-Yvette, France
[3]Department of Physics and Grantham Institute, Imperial College London, London, UK
[4]Scripps Institution of Oceanography, UC San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0244, USA

Correspondence to: D. D. Lucas (ddlucas@alum.mit.edu)

Abstract. Emission rates of greenhouse gases (GHGs) entering into the atmosphere can be inferred using mathematical inverse approaches that combine observations from a network of stations with forward atmospheric transport models. Some locations for collecting observations are better than others for constraining GHG emissions through the inversion, but the best locations for the inversion may be inaccessible or limited by economic and other non-scientific factors. We present a method to design an optimal GHG observing network in the presence of multiple objectives that may be in conflict with each other. As a demonstration, we use our method to design a prototype network of six stations to monitor summertime emissions in California of the potent GHG 1,1,1,2-tetrafluoroethane (CH_2FCF_3, HFC-134a). We use a multiobjective genetic algorithm to evolve network configurations that seek to jointly maximize the scientific accuracy of the inferred HFC-134a emissions and minimize the associated costs of making the measurements. The genetic algorithm effectively determines a set of "optimal" observing networks for HFC-134a that satisfy both objectives (i.e., the Pareto frontier). The Pareto frontier is convex, and clearly shows the tradeoffs between performance and cost, and the diminishing returns in trading one for the other. Without difficulty, our method can be extended to design optimal networks to monitor two or more GHGs with different emissions patterns, or to incorporate other objectives and constraints that are important in the practical design of atmospheric monitoring networks.

1 Introduction

Greenhouse gas (GHG) emissions are difficult to measure directly, which has led to the development of two indirect methods to estimate their emission rates. "Bottom-up" methods stitch together data on economic activity, fuel consumption, emission factors, and other disparate sources to form GHG emissions inventories (e.g., EDGAR, 2009). Alternatively, "top-down" methods estimate the emissions by combining measurements of GHG concentrations in the atmosphere from a network of stations with information about the atmospheric transport of the gases from their source region to the measurement location (e.g., Weiss and Prinn, 2011; Nisbet and Weiss, 2010). Bottom-up and top-down methods are both expected to play important roles in verifying GHG emissions policies at the state, national, and international levels (e.g., Ciais et al., 2010, 2014; National Research Council, 2010; Jonietz et al., 2011; Prinn et al., 2011; Fischer and Jeong, 2013).

The viability of using a top-down approach to constrain GHG emissions hinges on the network of observing stations. Measurements from the network are compared objectively to simulations from an atmospheric transport model using inverse methods (e.g., Prinn, 2000; Enting, 2000). If the observations contain useful information, and the atmospheric model provides an accurate representation of transport, inverse methods yield estimates of emissions that give the best agreement between the simulations and observations. Inverse methods often also provide an estimate of the uncertainties in

the inferred emissions. Various details of the observing network (e.g., station location, instrument precision, and measurement frequency) can profoundly impact the quality of the inversion. For example, if the stations in a GHG observing network are all located upwind of an emitting region of interest, the inversion algorithm will not provide any information on the emissions for that region.

Optimization techniques can be used to strategically place stations and select sampling strategies in a network, in order to maximize the information obtained for top-down inversion systems. Quantitative methods for designing "optimal" observing networks have been described for inferring carbon dioxide (CO_2) emissions, improving weather forecasts, collecting oceanographic data, and monitoring air quality and climate change (e.g., Barth and Wunsch, 1990; Morss et al., 2001; Patra and Maksyutov, 2002; Gloor et al., 2000; Carmichael et al., 2008; Stuart et al., 2007; Mauger et al., 2013). Ziehn et al. (2014) and Nickless et al. (2015) illustrate recent applications of using optimization methods to design GHG observing networks. Without requiring actual observations, so-called observing system simulation experiments (e.g., Masutani et al., 2010) can be used to create synthetic observations and assess the scientific value of adding new observations at various locations and times.

Network optimization studies typically construct and optimize a single objective function, which is usually related to the performance of the observing network (e.g., Mauger et al., 2013; Ziehn et al., 2014; Nickless et al., 2015). Although single objective optimization problems can consider several aggregated quantities, they still reduce the problem down to a single objective. Real-world observing networks, however, are generally faced with multiple, potentially conflicting objectives. The networks may measure more than one quantity, and there can be different strategies to optimize the separate quantities. For example, the problem of adding a new observing station to the Advanced Global Atmospheric Gases Experiment network (AGAGE, Prinn et al., 2000) inherently has multiple objectives. AGAGE measures a large suite of GHGs, including 1,1,1,2-tetrafluoroethane (CH_2FCF_3, HFC-134a), and uses these observations to estimate GHG emissions on both global and regional scales. Because many trace gases measured by AGAGE have distinct emissions patterns, it is not possible, in general and particularly at the regional level, to find a single location for a new station that will be best for monitoring all of the gases.

Economic and operational factors also heavily influence the design of observing networks (Morss et al., 2005). Without cost or instrumentation constraints, the overall goal of network design is to optimize the performance of the top-down network. However, some locations that optimize performance may be remote and may require new construction and infrastructure, which rapidly drive up costs. Alternatively, existing observing locations could be leveraged to make new measurements and keep costs low (e.g., using the AGAGE network), but these locations could be sub-

optimal for performance. Thus, there exists a natural tradeoff between performance and cost in optimal network design. Quantitative analyses of this tradeoff are needed to design practical GHG observing networks.

We apply a multiobjective genetic algorithm to quantify and optimize the performance-cost tradeoff curve for a prototypical top-down GHG observing network. Multiobjective optimization is a powerful generalization of standard, single objective optimization methods (Schaffer, 1985; Kursawe, 1991; Fonseca and Fleming, 1993; Zitzler and Thiele, 1999). Solutions to multiobjective problems are represented by a set of optimal points known as a Pareto frontier (Pareto, 1896), rather than a single point best case. Multiobjective methods have been used to solve many complex design and optimization problems (e.g., Jia et al., 2009; Jourdan et al., 2009; Judy et al., 2009; Kasprzyk et al., 2009), but they have been applied only sparingly to environmental monitoring (e.g., Trujillo-Ventura and Hugh Ellis, 1991; Sarigiannis and Saisana, 2008; Carnevale et al., 2012).

Extending the work of Yver et al. (2013), we design optimal networks to monitor HFC-134a emissions in California by combining the following three elements: forward atmospheric transport simulations of HFC-134a (see Sect. 2), top-down estimates of HFC-134a emissions using a Bayesian inversion scheme (see Sect. 3), and multiobjective optimization of network performance and cost using genetic algorithms (see Sect. 4). We use this framework to quantify the performance and cost tradeoffs between adding measurement stations at new locations in California vs. using existing stations designed for other purposes.

2   Forward atmospheric simulations

2.1   Model configuration

The forward atmospheric simulations used in this study were conducted as part of an effort to constrain HFC-134a emissions in California using atmospheric measurements and an inverse method (Yver et al., 2011, 2013). The network design methods presented and demonstrated here require only an archive of simulation output, not additional simulations. This archive contains time series of HFC-134a simulated throughout California over a 90-day period with an output frequency of 2 h. The HFC-134a was emitted using an emissions inventory and tagged by the region it originated from.

The archive was constructed using version 3.4 of the Weather Research and Forecasting model with coupled chemistry (WRF-Chem) (Grell et al., 2005), which uses the Advanced Research WRF dynamical core (Skamarock et al., 2005; Skamarock and Klemp, 2007). The model configuration for our specific archive had $129 \times 159 \times 27$ grid boxes in the longitudinal, latitudinal, and vertical directions, respectively, and did not use grid nesting. The domain was centered over the western United States using a horizontal resolution

of 12 km (see Fig. 1). Lateral boundary conditions were specified using ERA-Interim reanalysis data available from the European Centre for Medium Range Weather Forecasting, and the simulations were run over the period from May 2010 to the end of July 2010. All of the subsequent analysis utilizes the simulations from the second vertical level in WRF-Chem, which lies about 50 m above the surface. Air masses from this level contain useful information about emissions from distant regions and so are well suited for designing a network to constrain emissions through a top-down approach.

The HFC-134a time series in the archive were generated using version 4.1 of the gridded emissions inventory from EDGAR (EDGAR, 2009). The emissions, which are shown on the right hand side in Fig. 1, emit $8.6 \, \text{Gg yr}^{-1}$ of HFC-134a into California. To constrain the emissions through inverse modeling, 15 spatially separated HFC-134a tracers were emitted, tagged, and transported in WRF-Chem. These tracers were emitted from different regions based on the air basins defined by the California Air Resources Board (CARB). These basins, which are shown and numbered in the left hand side in Fig. 1, divide California into individual regions based on meteorological and air quality characteristics. A separate tracer for HFC-134a emitted from outside of California is also included.

2.2 Synthetic HFC-134a observations

Candidate sites for new observing stations are assessed by creating synthetic observations of HFC-134a from the forward model simulations and then using the observations in the Bayesian inversion scheme described in Sect. 3. The symbols ξ_m and ξ_o are used to represent time series of HFC-134a mole fractions at a candidate site from the forward model and synthetic observations, respectively. Hereafter, the terms "observations" and "synthetic observations" are used interchangeably, even though all of the observations are based on WRF-Chem simulations and not on actual measurements.

The background air advected into our simulation domain from the west (see Fig. 1) is typically well mixed in terms of HFC-134a, so we ignore the background levels and create synthetic observations of the enhancements of HFC-134a above the background that result from the prescribed emissions. For reference, background mole fractions of HFC-134a measured at the remote coastal site of Trinidad Head, California, ranged between 50 and 65 parts per trillion (ppt) over the period 2008–2010 (Prinn et al., 2000). The removal of the background level from our analysis has no effect on our results.

In order to produce synthetic observations that are reasonably consistent with actual observations, we first uniformly scale all of the simulated HFC-134a mole fractions using

$$\xi_m^* = 0.7 \xi_m, \tag{1}$$

where ξ_m^* represents the time series of the scaled modeled mole fraction enhancements at a candidate site. The mole fractions are reduced because version 4.1 of the EDGAR inventory overestimates HFC-134a emissions in California by about a factor of 1.4 (Yver et al., 2011). Noisy synthetic observations are then generated using the expression

$$\xi_o = \begin{cases} \xi_m^* + \rho\epsilon & \text{if } \epsilon \geq -\xi_m^*/\rho \\ \xi_m^* & \text{if } \epsilon < -\xi_m^*/\rho \end{cases}, \tag{2}$$

where ξ_o is the time series of "observed" mole fraction enhancements at a candidate site, ϵ is a set of random numbers drawn from a standard Gaussian distribution (one random number per datum in the time series), and ρ is the amplitude of the noise. The conditions in the expression apply to individual points in the time series, but these are not explicitly indexed to keep the notation compact. A noise amplitude of $\rho = 20 \, \text{ppt}$ is prescribed that is constant in space and time, and the noise is spatially uncorrelated by drawing different random numbers for ϵ at different locations using site-specific, unique grid cell integers (see Eq. 16) as input seeds to a random number generator.

The purpose of the noise is to inject uncertainty into the problem that can arise from a variety of factors, including imprecise measurements, scale representation errors, model imperfections, and other sources. Depending upon the relative magnitude of the noise amplitude (ρ) to the scaled mole fractions (ξ_m^*), the piecewise nature of Eq. (2) creates synthetic observations with different characteristics. For cases with relatively small noise levels (i.e., $\rho \ll \xi_m^*$), most of the random numbers drawn from ϵ satisfy the upper condition. This results in observation–model residuals that are approximately normally distributed, and values for ξ_o that are generally less than ξ_m because of Eq. (1). For cases with relatively large noise levels (i.e., $\rho \gg \xi_m^*$), many of the random numbers in ϵ satisfy the lower condition, which filters out the negative values (but keeps the positive values). The lower condition therefore causes the distribution of observation–model residuals to be highly non-Gaussian and skewed toward positive values, and leads to situations in which ξ_o can be greater than ξ_m.

Figure 2 compares examples of simulated and "observed" time series of HFC-134a at Walnut Grove and Trinidad Head using Eqs. (1) and (2). These sites are located at the southern edge of basin 3 and along the coast in basin 2, as shown in Fig. 1. The time series are displayed using a sampling frequency of $1 \, \text{sample day}^{-1}$ to clearly show the model and observation differences, though higher frequency sampling strategies are also tested and evaluated (see Eq. 17). For the reasons noted in the previous paragraph, the model tends to overestimate the synthetic observations at Walnut Grove because the noise amplitude is relatively small at that location (i.e., $\rho = 20 \, \text{ppt}$ is less than $\xi_m^* \approx 100 \, \text{ppt}$). Furthermore, the model underestimates the synthetic observations at Trinidad Head because the noise amplitude is relatively large at that coastal location (i.e., $\rho = 20 \, \text{ppt}$ is greater than $\xi_m^* \approx 10 \, \text{ppt}$). The synthetic observations in the figure also appear qualitatively similar to actual observational time se-

www.geosci-instrum-method-data-syst.net/4/121/2015/

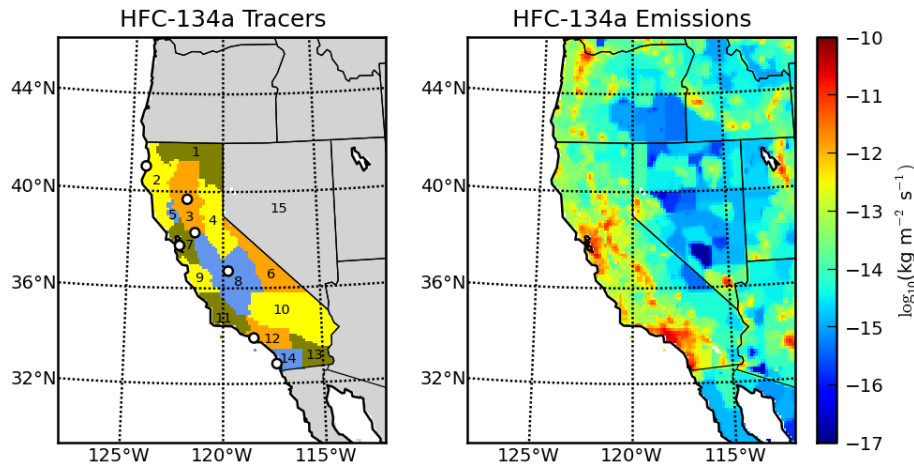Geosci. Instrum. Method. Data Syst., 4, 121–137, 2015

Figure 1. Both figures show the spatial domain and model grid used for the simulations of HFC-134a using WRF-Chem. The figure on the left shows the 15 regions used for tagging the HFC-134a tracers (regions 1–14 in California, 15 outside of California), and the locations of seven existing measurement sites (white dots). The figure on the right shows the spatial distribution of HFC-134a emissions from the version 4.1 EDGAR inventory on the WRF-Chem model grid.
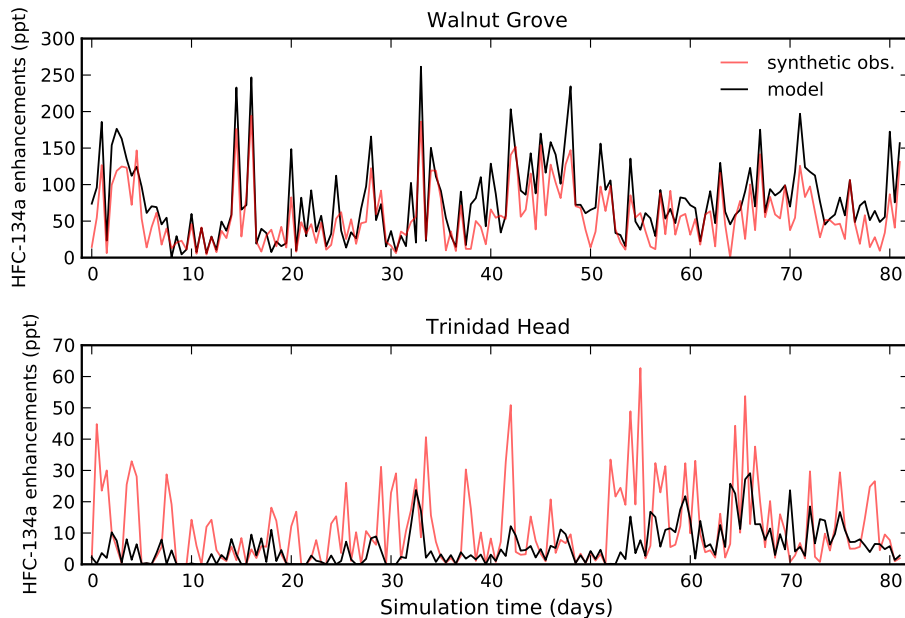

Figure 2. The figure shows the HFC-134a time series from the forward model simulations (black lines, ξ_m) and synthetic observations (red lines, ξ_o) at Walnut Grove (upper) and Trinidad Head (lower). Synthetic observations are generated using Eqs. (1) and (2). The time series are displayed using 1 sample day^{-1} to clearly show the observation–model differences, though higher frequency sampling strategies are available and tested.

ries (Yver et al., 2011) because they exhibit localized high-concentration events that are not captured by the model and EDGAR emissions.

These synthetic observations thus provide a realistic challenge for inversion algorithms. The skewed, non-Gaussian component reflects model structural errors or systematic biases that can affect the source inversion (e.g., Baker et al., 2006) and, consequently, the design of the observing network. The presence of these errors implies that there is no

emission configuration in our setup that can simulate HFC-134a to perfectly match the synthetic observations. Additional terms could be included in the inversion method in Sect. 3 to estimate and account for biases, but these considerations are outside of the scope of this work and do not impact the network optimization methodology that is the main focus of this report. Moreover, because the "true" emissions values corresponding to the observations are known, the impact of these errors on the performance of our inversion algorithm

can be verified. As shown later (Sect. 5.2), these errors have a small effect because our inversion algorithm successfully retrieves both the true emissions and prescribed noise level.

3   Bayesian inversion

The surface emissions of HFC-134a (model inputs) are inferred by solving an inverse problem that minimizes the differences between observed and simulated mole fractions in the atmosphere (model outputs). The target "observations" are taken as the values from Eq. (2) (e.g., the red lines in Fig. 2), while the simulations are the values produced by the model and unscaled EDGAR emissions (e.g., the black lines in Fig. 2). Differences between these values are used to estimate individual scaling factors (or weights) that apply to the original emissions from the 15 spatial regions shown in Fig. 1.

Due to the linear relationship between emission levels and atmospheric mole fractions, the net time series of HFC-134a simulated by the model is a weighted sum of the time series of the individual HFC-134a tracers emitted and tagged from the separate regions. This relationship is expressed as

$$\boldsymbol{\xi}_{\mathrm{m}} = \mathbf{X}_{\mathrm{m}} \boldsymbol{w}, \tag{3}$$

where the boldface notation is used to denote vectors and matrices. The symbol $\boldsymbol{\xi}_{\mathrm{m}}$ is a column vector containing the data points in the net time series, $\mathbf{X}_{\mathrm{m}}$ is a matrix with individual columns containing the time series of the corresponding 15 tagged tracers, and $\boldsymbol{w}$ is a column vector of scaling factors (weights) for emissions from the 15 regions. The symbol $\boldsymbol{\xi}_{\mathrm{o}}$, which is used in expressions below, is similarly defined as the vector of the time series of synthetic observations. Equation (3) is applicable to the time series at a single site or, by concatenating vectors together, at many locations in an observing network.

The goal of the inversion is to determine the values of the emissions weights, $\boldsymbol{w}$, that minimize differences between the model $\boldsymbol{\xi}_{\mathrm{m}}$ and observations $\boldsymbol{\xi}_{\mathrm{o}}$. Because there is uncertainty in these quantities, a probabilistic Bayesian approach is adopted that estimates the probability distribution of weights by incorporating uncertainty (i.e., see the terms α and β introduced below). Bayesian inversion schemes normally supply "prior" emissions up front (e.g., Patra and Maksyutov, 2002; Gurney et al., 2003; Thompson et al., 2011), so that the algorithm is constrained when the observations are non-informative. The resulting "posterior" emissions can be highly sensitive to the prescribed prior emissions. To circumvent this issue, we employ an iterative Bayesian technique known as *evidence approximation* (MacKay, 1992; Bishop, 2007). Evidence approximation, which is described in more detail at the end of this section, uses the data to estimate the values of distribution parameters that are usually prescribed in the inversion, resulting in posterior emissions that are insensitive to the priors.

Given observations of HFC-134a, the probability distribution of weights for the emissions is obtained from Bayes' rule,

$$p\left(\boldsymbol{w}|\boldsymbol{\xi}_{\mathrm{o}}\right) \propto p\left(\boldsymbol{\xi}_{\mathrm{o}}|\boldsymbol{w}\right) p(\boldsymbol{w}), \tag{4}$$

where $p(a|b)$ denotes the conditional probability of a given b, $p(\boldsymbol{w})$ is the prior distribution for the emission weights, $p(\boldsymbol{\xi}_{\mathrm{o}}|\boldsymbol{w})$ is the likelihood that the simulation matches the observations for a given set of emission weights, and $p(\boldsymbol{w}|\boldsymbol{\xi}_{\mathrm{o}})$ is the posterior distribution of the weights.

The prior distribution of weights for the emissions is modeled as

$$p(\boldsymbol{w}) = \mathcal{N}\left(\boldsymbol{w}|\boldsymbol{m}_0, \mathbf{S}_0\right), \tag{5}$$

where $\mathcal{N}(\boldsymbol{w}|\boldsymbol{m}_0, \mathbf{S}_0)$ denotes a normal distribution over variable $\boldsymbol{w}$ with a mean of $\boldsymbol{m}_0$ and covariance of $\mathbf{S}_0$. The prior distribution is further modified by setting $\boldsymbol{m}_0 = \mathbf{1}$ and $\mathbf{S}_0 = \alpha^{-1} \mathbf{I}$, though these settings do not lead to any loss in generality. The latter setting yields prior emissions uncertainties that are independent between the regions and have variances of α^{-1}. The range for α allows for prior emissions distributions that are infinitely wide or narrow, or anything in between ($0 < \alpha < \infty$). The value of α is not prescribed, it is determined from the data using evidence approximation as described below.

For differences between simulated and observed mole fractions that are normally distributed, the likelihood function is given by the product of probabilities,

$$p\left(\boldsymbol{\xi}_{\mathrm{o}}|\boldsymbol{w}\right) = \prod_{i=1}^{N_{\mathrm{d}}} \mathcal{N}\left(\xi_{\mathrm{o},i}|\xi_{\mathrm{m},i}, \beta^{-1}\right), \tag{6}$$

where the product is over N_{d} data points in the time series at all of the stations, and β represents observation and model uncertainty (i.e., β^{-1} is the variance). Because the noise comes from Eq. (2), there is a relationship between β and $\epsilon \rho$. Instead of prescribing β from this relationship, we also use evidence approximation to estimate the value of β directly from the data.

Using these forms for the prior distribution and likelihood function, the posterior distribution of weights for the emissions is also Gaussian,

$$p\left(\boldsymbol{w}|\boldsymbol{\xi}_{\mathrm{o}}\right) = \mathcal{N}\left(\boldsymbol{w}|\boldsymbol{m}_N, \mathbf{S}_N\right), \tag{7}$$

with a mean value ($\boldsymbol{m}_N$) and covariance ($\mathbf{S}_N$) given by

$$\boldsymbol{m}_N = \mathbf{S}_N \left(\alpha \mathbf{1} + \beta \mathbf{X}_{\mathrm{m}}^T \boldsymbol{\xi}_{\mathrm{o}}\right), \tag{8}$$

$$\mathbf{S}_N^{-1} = \alpha \mathbf{I} + \beta \mathbf{X}_{\mathrm{m}}^T \mathbf{X}_{\mathrm{m}}. \tag{9}$$

Equations (8) and (9) constitute the solution to the Bayesian inversion problem for the emissions of HFC-134a. The posterior emissions, however, still require uncertainty information about the prior fluxes (α) and observation–model noise (β).

Evidence approximation is used to estimate α and β from the simulations and synthetic observations. A detailed derivation of the method is given in MacKay (1992) and Bishop (2007). The method is iterative, starting with initial guesses for α and β. These are used to calculate $\boldsymbol{m}_N$ and $\mathbf{S}_N$ from Eqs. (8) and (9), and to calculate the quantity

$$\gamma = \sum_{j=1}^{N_\mathrm{r}} \frac{\beta \lambda_j}{\alpha + \beta \lambda_j}, \tag{10}$$

where N_r is the number of regions (15 for our problem) and the λ_j are the eigenvalues of $\mathbf{X}_\mathrm{m}^T \mathbf{X}_\mathrm{m}$. Equation (10) provides a measure of the number of regions with constrained emissions. In the limit of infinitely wide prior emissions and narrow observation–model uncertainties ($\alpha \to 0$ and $\beta \to \infty$), all of the regions can be constrained and $\gamma = N_\mathrm{r}$. For infinitely narrow prior emissions and wide observation–model uncertainties ($\alpha \to \infty$ and $\beta \to 0$), none of the regions can be constrained and $\gamma = 0$. After computing γ, the values for α and β are updated using

$$\alpha = \frac{\gamma}{\boldsymbol{m}_N^T \boldsymbol{m}_N} \tag{11}$$

and

$$\beta^{-1} = \frac{\mathbf{1}^T \left(\boldsymbol{\xi}_\mathrm{o} - \mathbf{X}_\mathrm{m} \boldsymbol{m}_N \right)}{N_\mathrm{d} - \gamma}, \tag{12}$$

where the differences between the observed and simulated mole fractions use the posterior emission weights from the current iteration. The updated values for α and β are then used to re-calculate $\boldsymbol{m}_N$, $\mathbf{S}_N$, and γ. The process is repeated until convergence is achieved. Note that the eigenvalues λ_j only have to be computed once at the beginning of the scheme. For convergence, we iterate through the procedure until neither α nor β change by more than 5 %, which usually requires only a few iterations. Using this procedure, the uncertainty in the prior emissions (α) and observation–model noise (β) are discovered from the data, and converge on values that are close to their true values, given the skewed, non-Gaussian nature of Eq. (2).

4   Multiobjective optimization for network design

A primary goal of network design problems is to determine the best locations and sampling strategies for a collection of instruments or sensors that optimize a given set of objectives. In designing a wireless communications network, for example, the objectives may be to achieve complete coverage over a given area using a limited number of transmitters (e.g., Jia et al., 2009). Network design problems often involve two or more conflicting objectives that need to be optimized simultaneously (e.g., cost and performance), which falls into a class of problems known as multiobjective optimization.

Multiobjective optimization problems are formulated mathematically as

$$\underset{z}{\text{minimize}} \left[f_1(z), f_2(z), \ldots, f_n(z) \right] (n \geq 2)$$

$$\text{subject to } \{ g(z) = 0, h(z) \leq 0, z_\ell \leq z \leq z_u \}, \tag{13}$$

where z represents design parameters that need to be optimized (e.g., measurement locations) and the $f_i(z)$ are the multiple objectives of interest (e.g., inversion errors and measurement costs). A set of constraints can be applied to the design parameters, including lower and upper bounds that are placed on the parameter values (z_ℓ and z_u), and linear or non-linear equality or inequality constraints that must be satisfied ($g(z) = 0$ and $h(z) \leq 0$). There is usually not a single set of z that minimizes all of the objectives in a multiobjective problem. Rather, there are multiple sets of optimal points known as a *Pareto frontier*. The points along a Pareto frontier are optimal in the sense that moving to other locations in the design parameter space may improve one or more objectives, but will worsen at least one of the other objectives and lead to an overall less desirable solution (Pareto, 1896). Conversely, for all design parameters that satisfy the constraints but that are not on the Pareto frontier, there exist other points in the design space that improve one or more of the objectives.

4.1   Simple multiobjective example

To better illustrate the concept of multiobjective optimization and the Pareto frontier, consider the simple example given below and shown in Fig. 3:

$$\underset{z_1, z_2}{\text{minimize}} \left[\begin{array}{l} f_1 = (z_1 - 0.35)^2 + (z_2 - 0.35)^2, \\ f_2 = (z_1 - 0.65)^2 + (z_2 - 0.65)^2 \end{array} \right] \tag{14}$$

subject to $0 \leq z_1 \leq 1, 0 \leq z_2 \leq 1$.

The goal of this problem is to determine the design points (z_1, z_2) that minimize the two quadratic functions f_1 and f_2, subject to the constraint that z_1 and z_2 are bounded by 0 and 1. As shown in the left and center panels of Fig. 3, there is not a single design point (z_1, z_2) that minimizes both functions simultaneously. By inspecting Eq. (14), it is easy to see that f_1 is minimized at point (0.35, 0.35), which yields function values of $f_1 = 0.0$ and $f_2 = 0.18$. In contrast, f_2 is minimized at point (0.65, 0.65), which yields function values of $f_1 = 0.18$ and $f_2 = 0.0$. Between these cases exist other design points that are also considered to be "optimal" because they minimize preferred combinations of f_1 and f_2. The Pareto frontier is the set of function values plotted in the objective space for the optimal design points.

The right panel in Fig. 3 displays the solution to Eq. (14). The light blue shaded region shows non-optimal values of the objective functions for feasible combinations of z_1 and z_2, while the red line along the lower left edge shows the Pareto frontier. For this simple example, an analytical expression for the Pareto frontier can be derived by setting $z_1 = z_2$ based on
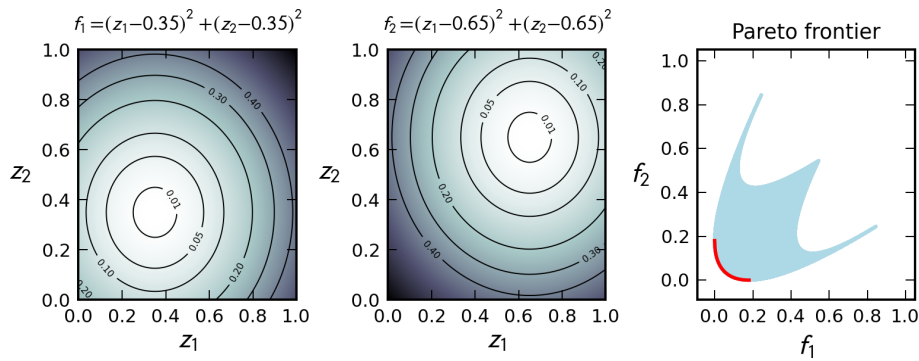
Figure 3. The figures illustrate the simple multiobjective problem described in Sect. 4.1. The figures on the left and in the middle show contours and shadings of the two quadratic objective functions f_1 and f_2 as a function of design variables z_1 and z_2 (low and high values are indicated by light and dark shades, respectively). The figure on the right shows non-optimal solutions (light blue shaded region) and optimal points along the Pareto frontier (red line) for the problem given in Eq. (14).

the symmetry of the problem, and then eliminating z_1 between f_1 and f_2. This leads to the following expression for the Pareto frontier:

$$f_2 = 2\left(\sqrt{\frac{f_1}{2}} - 0.3\right)^2 \text{ for } 0 \leq f_1 \leq 0.18. \tag{15}$$

The points along the Pareto frontier are clearly optimal because there is no way to improve the combination of f_1 and f_2 by moving to other points in the design space.

4.2 Designing a multiobjective HFC-134a observing network

The goal for the multiobjective HFC-134a network design demonstration problem is to select "optimal" locations for placing six observing stations to monitor summertime emissions of HFC-134a from California. Optimal locations are determined by jointly maximizing the scientific performance and minimizing the measurement costs of the observing network. Seven "existing sites" are available that have related measurement capabilities. Including any of these existing sites in the network will reduce the costs, but may decrease the performance. This section provides further mathematical details of the optimization problem (design variables, search space, and objectives) and describes the numerical algorithm used to solve the problem. Given the size and complexity of the problem, and the nature of the numerical optimization algorithm, it is important to keep in mind that the resulting observing networks are not global optimal solutions. Instead, they represent plausible local optimal designs that are significantly better than a random selection of sites. Moreover, we also caution against using these designs as a basis for a real-world HFC-134a observing network, as many factors were not included in this demonstration (e.g., biases in WRF-Chem transport, inter-seasonal variations of HFC-134a, year-to-year changes in meteorology and emissions, and terms not represented in the idealized cost model).

4.2.1 Design variables and search space

Two types of design variables are considered for our HFC-134a observing network test problem. We consider different locations for placing the six observing stations ($z_1, \ldots, z_6$) and alternate frequencies for making measurements at the sites (z_7). The eligible locations for the observing network (also referred to as candidate sites) are taken as the discrete grid boxes in the WRF-Chem domain that fall within California, excluding offshore sites (e.g., Catalina Island). At the spatial resolution used for the WRF-Chem model runs, there are 2921 eligible sites (see Fig. 1). The locations of candidate sites are inherently two dimensional (latitude and longitude), but we encode them as one-dimensional integer-valued design variables:

$$z_{1-6} \in \{1, 2, \ldots, 2921\}. \tag{16}$$

Candidate site 1 is set as the grid box at the southernmost and westernmost part of the domain. The remaining candidate sites are incremented moving from west to east, followed by south to north. Candidate site 2921 therefore corresponds to the northernmost and easternmost grid box. This encoding scheme is straightforward but it loses information about latitudinal gradients. For example, candidate sites 556 and 604 are adjacent in physical space, but not in design space. Better methods can be used to encode multiple spatial dimensions into one-dimensional design variables for optimization (e.g., using Hilbert space-filling curves; Sergeyev et al., 2013), but these fall outside the scope of this paper. In future work we plan to investigate the effects of using different spatial encodings on geophysical optimization problems.

As with the station locations, the daily measurement frequency is also represented as an integer-valued design variable, though we use the same frequency for all six of the locations. Measurement frequency is included as a design variable because changing the number of measurements leads to an interesting tradeoff between network performance and

cost. This variable takes integer values 1–6 and maps them to the six different ways of dividing 24 h into regular sampling intervals using the 2-hourly WRF-Chem output:

$$z_7 \in \{1, 2, 3, 4, 5, 6\} \longmapsto \left\{ 1, 2, 3, 4, 6, 12 \, \text{samples day}^{-1} \right\}. \quad (17)$$

Values of $z_7 = 4$ and 5, for example, correspond to 4 and 6 samples day^{-1}. Note that a case involving 8 samples day^{-1} is not included because it involves 3-hourly measurements that would require interpolation of the WRF-Chem output archive.

These design variables are independent directions in a seven-dimensional integer-valued search space. Brute force search methods are impractical for searching through a space this large. To illustrate, first consider the simple case of choosing a location for just a single monitoring station with a fixed measurement frequency. For this case, 2921 candidate sites need to be assessed to optimize the objectives. Choosing the locations for a pair of fixed-frequency stations, however, yields a search space containing roughly 4.2 million design points. The number of ways of selecting r stations out of s candidate sites is a combinatorial counting problem, which is calculated from the binomial coefficient $\binom{s}{r}$. The full search space for our problem therefore contains more than 5×10^{18} design points. Directly evaluating all of these points is not feasible with current computers, so we apply a global stochastic numerical optimization algorithm that is effective for solving multiobjective problems in large search spaces (Sect. 4.2.3).

4.2.2 Performance and cost objectives

Two objectives are jointly optimized in the network design. These are to find design points that maximize performance, f_1, and minimize measurement costs, f_2. These objectives may be in conflict with each other in our demonstration and in other network design problems. For example, a candidate site may be well positioned to sample air masses from an important emissions basin (e.g., downwind of Los Angeles), but the site could be expensive to set up and maintain if it requires new construction and is located in an isolated and rugged area. The compromise of placing a station at an existing site with existing infrastructure would reduce the costs, but may provide less information for the inversion, which would decrease the performance.

Performance is optimized by minimizing

$$f_1 = \text{Tr} \left(\mathbf{S}_N \right), \quad (18)$$

which is the trace of the covariance matrix of the posterior distribution of HFC-134a emissions. This objective is equivalent to minimizing the mean squared estimation error (Huber, 2009). Alternate performance objectives based on the covariance matrix could be formulated and applied, common choices being the determinant of $\mathbf{S}_N$ or weighted versions of the trace or determinant (e.g., Huber, 2009). To keep the discussion brief, we use only Eq. (18) in our demonstration.

For the cost objective, we assume that it is less expensive to set up HFC-134a monitoring capabilities near sites where infrastructure already exists and atmospheric measurements or soundings are routinely taken (e.g., sites in the National Oceanic and Atmospheric Administration's Cooperative Air Sampling Network). The following seven locations in California are considered as "existing sites" where costs can be minimized: Trinidad Head, Chico, Walnut Grove, Sutro Tower, Fresno, Los Angeles, and Scripps. The locations of these sites are shown by the white circles in Fig. 1.

The total cost for the six-station observing network is calculated using

$$f_2 = \sum_{i=1}^{6} c_s (z_i) + c_o (z_7), \quad (19)$$

where $c_s(z_i)$ denotes a one-time "setup" cost for adding a station at location z_i, and $c_o(z_7)$ is a continuing "operational" cost for making measurements at all of the stations with a frequency of z_7. The total cost is normalized to the range [0, 1]. A minimum cost of 0 coincides with a network that uses only existing sites and that make one measurement per day. A maximum total cost of 1 occurs with stations that are far from existing sites and that make 12 measurements per day, with half of the total coming from the setup cost and half from the operational cost.

For a candidate site z_i, the setup cost is assumed to vary with the distance, $d(z_i)$, to the nearest existing site, up to a maximum distance, d_{max}. This is expressed as

$$c_s (z_i) = \frac{1}{12} \min \left[1, \left(\frac{d(z_i)}{d_{max}} \right)^q \right] \text{ for } q > 0, \quad (20)$$

where q is the distance dependency (e.g., $q = 1$ for linear, and $q = 2$ for quadratic). If an existing station is used for monitoring, $d(z_i) = 0$ and there is no setup cost. If a candidate site is more than d_{max} away from the nearest existing station, $d(z_i) > d_{max}$ and the setup cost is $\frac{1}{12}$. If all of the candidate sites are more than d_{max} away from existing sites, the setup cost is 0.5. For the HFC-134a demonstration problem, we set $d_{max} = 150 \, \text{km}$ and $q = 2$. The former setting provides a radial window of about 12 grid boxes around existing sites over which candidate sites can "sense" the effects of existing sites, while the quadratic dependency provides a spatial gradient strong enough to drive candidate sites toward existing sites to minimize costs.

The operational cost is assumed to depend linearly on z_7 through the expression

$$c_o (z_7) = 0.5 (z_7 - 1) / 5, \quad (21)$$

where z_7 is the design variable, as opposed to the measurement frequency, given in Eq. (17). There is no operational cost for making the first measurement per day in this formulation, and c_o is half of the maximum total cost at the maximum measurement frequency. Embedded within

this expression is the assumption that it is more cost effective to operate at higher measurement frequencies, because the marginal cost of increasing from 1 to 3 samples day^{-1} (i.e., from $z_7 = 1$ to 3) is the same as increasing from 4 to 12 samples day^{-1} (i.e., from $z_7 = 4$ to 6).

Although the cost model described by Eqs. (19)–(21) is idealistic and does not include specific prices for GHG measurement instruments, site permits, personnel, and so forth, it is still useful for demonstrating our multiobjective network design methodology. If available, a realistic economic cost model could be substituted and used to compute f_2. Without a loss in generality, the same techniques would be used to optimize cost and performance and quantify the tradeoffs between the two.

4.2.3 Multiobjective genetic algorithms

Unlike the Pareto frontier that was derived analytically for the simple example in Sect. 4.1, numerical algorithms must be used for moderately complex multiobjective optimization problems. We use a genetic algorithm to design the HFC-134a observing network. Genetic algorithms (also known as evolutionary algorithms) have only recently been adapted to multiobjective optimization problems (e.g., Deb et al., 2002; Zitzler et al., 2002), but they have already been shown to be effective and efficient. Genetic algorithms have also been used to optimize GHG networks for single objectives (Rayner, 2004; Nickless et al., 2015).

Genetic algorithms evolve generations of a population of potential designs through a search space using notions such as survival-of-the-fittest and reproduction. Each loop of a genetic algorithm represents one generation, and at each generation four genetic operations are applied: *fitness assessment*, *reproduction*, *crossover* and *mutation*. The fitness step is an evaluation of the objectives at the design points for each member of the population. The members of the population are ranked according to their fitness scores. In reproduction, the members with the highest fitness rankings are given the highest probability of remaining in the population and surviving through subsequent generations. Crossover refers to the process of mixing characteristics of similarly ranked parents to produce offspring with potentially strong rankings. The mutation step adds randomness to the designs that are evolved.

For multiobjective problems, modern genetic algorithms also apply niche operators to promote diversity of the designs across the Pareto frontier. A genetic algorithm can therefore derive a diverse set of Pareto optimal solutions in a single optimization run, which is a great advantage over other methods that require multiple runs to characterize the multiobjective space. For our network design problem, we use the multi-objective genetic algorithm (MOGA) (Eddy and Lewis, 2001; Adams et al., 2010) to optimize performance and cost, and a single objective variation with the same genetic operators (SOGA – single objective genetic algorithm)

Table 1. Settings used in the genetic algorithms.

Setting	SOGA	MOGA
population_size	25	70
max_function_evaluations	6000	12 000
initialization_type	unique_random	unique_random
fitness_type	merit_function	domination_count
crossover_type	shuffle_random	shuffle_random
num_offspring	2	2
num_parents	2	3
crossover_rate	0.6	0.6
mutation_type	replace_uniform	replace_uniform
mutation_rate	0.3	0.3
replacement_type	elitist	below_limit $= 6$
niching_type	–	radial $= 0.15, 0.15$

Refer to Eddy and Lewis (2001) and Adams et al. (2010) for further details about these settings and other available options.

to optimize only the performance. Table 1 lists the MOGA and SOGA settings used for the network design. To our knowledge, this work represents the first application of a genetic algorithm to a multiobjective design of an atmospheric monitoring network.

4.2.4 Incremental optimization

As a benchmark for referencing the algorithmic performance of SOGA, we also employed the incremental optimization (IO) strategy described and benchmarked by Patra and Maksyutov (2002). These authors used IO to design a surface network for constraining CO_2 emissions. For their problem, they showed that IO outperformed another popular optimization method known as simulated annealing (Kirkpatrick et al., 1983) that was used to design a CO_2 network in earlier work (Rayner et al., 1996). More recently, IO was used to optimize single, scalar performance objectives for CO_2 measurement networks in Australia and South Africa (Ziehn et al., 2014; Nickless et al., 2015). In the latter case, the authors found that IO optimized the network more efficiently than a single objective genetic algorithm, which runs counter to our results in Sect. 5.1.

IO uses an intuitive, recursive approach to build up a network of observing stations. Starting with the first station, all of the candidate sites are evaluated and the station is placed at the location that optimizes objective f_1. This site is removed from the set of candidate sites, and then the next station is added at the location from the remaining sites that optimizes the objective. This process is repeated until each station in the network has been added. IO thus adds stations incrementally to the network by including them at locations that maximize the performance at each stage. This procedure requires

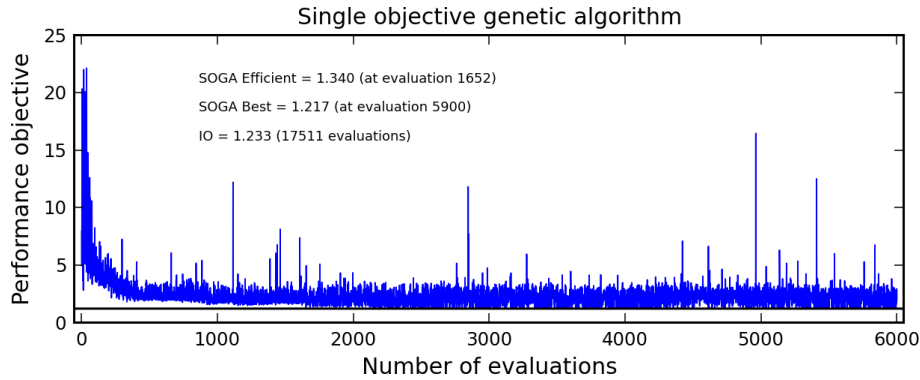$$\sum_{i=0}^{r-1} (s - i) \tag{22}$$

Figure 4. The figure displays the raw objective function evaluations during the evolution of a population of network designs using SOGA to optimize performance. The horizontal black line shows the SOGA Best case. The SOGA Best, SOGA Efficient, and IO cases are also displayed.

objective function evaluations for s candidate sites and r network stations. IO collapses the search in a large r-dimensional space into a series of trivial searches in one dimension, but the method does not account for potential synergistic benefits that can arise from adding two or more stations at the same time.

5 Network optimization results

5.1 Optimization of performance

SOGA is used to optimize only the performance of the HFC-134a observing network for a fixed cost (i.e., minimize f_1 while keeping f_2 constant). The daily measurement frequency is set at $4\,\mathrm{samples\,day}^{-1}$ for this experiment. Figure 4 displays the raw evaluations of the performance objective from a population of network designs over the evolution of the genetic algorithm. As shown, the algorithm clearly evolves the population toward an optimal solution. The population starts out with networks having performance objectives ranging from about 5 to over 20. Following initialization, there is a period of rapid improvement up through 300 objective function evaluations (and about 20 generations). Over this period, the weakest network designs are excluded, while the strongest designs are significantly improved. Between 300 and 1700 evaluations, the algorithm continues to improve the best designs, albeit more slowly, and still maintains a random search for potentially better designs (i.e., the spikes in the figure). Somewhere around evaluation 1700 (or 225 generations) there is a noticeable drop in the minimum f_1 as the genetic algorithm finds a design that is close to the best overall design. A minimum value of $f_1 = 1.217$ is achieved on generation 727 and objective function evaluation 5900 (referred to as "SOGA Best" below). The algorithm terminates after 740 generations and 6000 objective function evaluations because the max_function_evaluations limit setting is reached (see Ta-

ble 1). Note that a convergence criterion could be used instead to terminate the algorithm.

Although 6000 objective function evaluations may appear to be a large number, it is a tiny fraction of the number of design points that occupy the full six-dimensional search space (recall that z_7 is fixed for this experiment). A total of 6000 evaluations is also slightly more than twice the size of the search space for adding only a *single* station (i.e., 2921 candidate sites). Moreover, the algorithm found a reasonably optimal design ($f_1 = 1.340$) after only 1652 objective function evaluations (referred to as "SOGA Efficient" below).

To further put these results in perspective, we compare SOGA to the IO method described in Sect. 4.2.4. Applied to our problem, the IO strategy is indeed effective, yielding $f_1 = 1.233$, but it uses 17 511 evaluations to get there, which is almost 3 times as many evaluations as the total number shown in Fig. 4. However, SOGA and IO should not be compared only on the basis of the number of objective function evaluations. The time to compute the objective function f_1, described here as the "evaluation time," depends on the number of stations in the network because the sizes of $\boldsymbol{\xi}_\mathrm{m}$, $\boldsymbol{\xi}_\mathrm{o}$, and $\mathbf{X}_\mathrm{m}$, and hence the time to solve $\mathbf{S}_N$, vary with the network size. The "evaluation time" changes for IO because it adds stations one-by-one to the network, whereas it remains constant for SOGA because six stations are assessed during each iteration. We derive a linear relationship between the "evaluation time" and number of network stations, and use the relationship to estimate the cumulative "evaluation time" over the objective function evaluations for IO, SOGA Best, and SOGA Efficient. This analysis shows that the SOGA Best and SOGA Efficient cases are both more efficient than IO because the larger number of objective function evaluations for IO degrades algorithmic efficiency more than the increase in the "evaluation time" for larger networks in SOGA.

There is an additional factor besides the "evaluation time" that affects the efficiency of IO and SOGA. The "decision time" is the amount time it takes for the algorithm to decide

which station or stations to add to the network. For IO, the "decision time" is negligible and is based on a sort/search for the station with the smallest inversion variance at each stage. The "decision time" for SOGA, on the other hand, is tied to the four genetic operators (fitness assessment, reproduction, crossover and mutation) and varies from generation to generation because the population changes. We estimated an average "decision time" for SOGA and found that it is much smaller than the "evaluation time" and does not hinder the performance of the algorithm. We therefore conclude that, for our problem, SOGA is a more efficient algorithm for network design than IO. By our estimates, SOGA is about 2–7 times more efficient than IO, depending upon which design is used (i.e., SOGA Efficient versus SOGA Best). These results counter the findings of Nickless et al. (2015), whose analysis suggests that IO is more efficient than genetic algorithms. Further work is needed to compare SOGA and IO for network design under a variety of conditions (e.g., different tracers and larger networks), though we expect SOGA to scale well to larger networks with more candidate sites. By enabling smarter location encoding schemes (as described by the mapping in Eq. 16), we also expect to improve the search efficiency of SOGA relative to the current implementation used in Fig. 4.

Figure 5 displays three different HFC-134a observing networks resulting from the performance optimization. The figure shows the SOGA Best, SOGA Efficient, and IO network cases. The three networks have sites that overlap or that are in close proximity at four out of the six stations (in basins 1, 5 and 13, and near Los Angeles). For the two other stations, the SOGA Efficient and IO networks have overlapping sites in basins 5 and 8, while the SOGA Best case places the sites in basins 3 and 9. It is notable that the SOGA Efficient and IO networks are very similar, even though the latter requires more objective function evaluations to determine the locations.

Given the spatial distribution of HFC-134a emissions shown in Fig. 1, the positions of the stations in the three networks in Fig. 5 seem plausible. The three networks have stations surrounding or downwind of the three largest emitting regions in California (Southern California, the San Francisco Bay Area, and the Central Valley). The largest location difference occurs near the Bay Area and Central Valley, where the SOGA Best network appears to find better locations for constraining emissions from these regions. Recall that existing stations (white circles in Fig. 1) and measurement costs were not factored in this single objective experiment. We can therefore conclude that the best performing HFC-134a observing networks are not coincident with the assumed existing sites, implying that new measurement sites (with higher costs) can be developed to maximize performance.
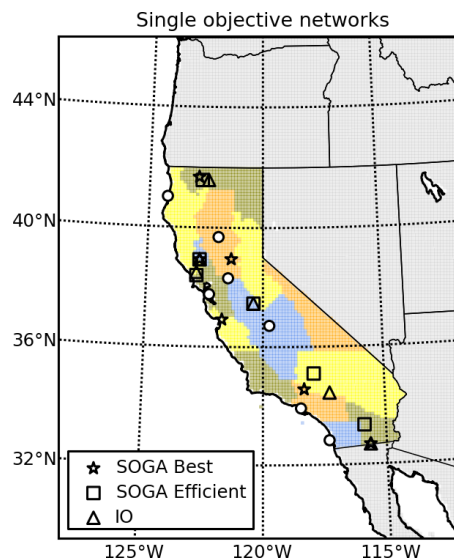

Figure 5. The figure shows the locations of observing stations in the SOGA Best case (stars), SOGA Efficient case (squares), and IO case (triangles). Reference locations of the seven existing observing sites are also shown (white circles).

5.2 Verification of emissions inversion

Because the HFC-134a observations are synthesized using Eqs. (1) and (2), the true weights for the emissions and the observation–model noise level are known. This information is used to verify the operation of the inversion algorithm. Figure 6 displays the posterior weights for the emissions estimated using the SOGA Best, SOGA Efficient, and IO networks. Considering the mean values and uncertainty ranges, and noting that covariances are excluded in the figure, the posterior weights are consistent with the emissions-scale factor of 0.7 applied in Eq. (1) for all but a few of the regions. The match is not expected to be perfect for all of the regions because some of them are greatly affected by non-Gaussian noise, are far from the stations in the networks, and have low levels of emissions. The inversion algorithm has a difficult time constraining the emissions in region 15, for example, because this region lies outside of California and has emissions that are not effectively transported to the network stations. The low-level emissions from regions 1, 2, 6, and 10 are also challenging for the inversion algorithm because they are located in remote portions of the state. Overall, however, the posterior weights are well estimated for the regions with the heaviest emissions, which provides verification that the algorithm is operating as desired.

The values of β inferred using evidence approximation can also be verified. The inverse square root of β represents the observation–model noise in the inversion algorithm and has units of ppt. The estimates of $\beta^{-1/2}$ for the SOGA Best, SOGA Efficient, and IO networks are 16.6, 16.2, and 16.1 ppt, respectively. These values are similar to each other
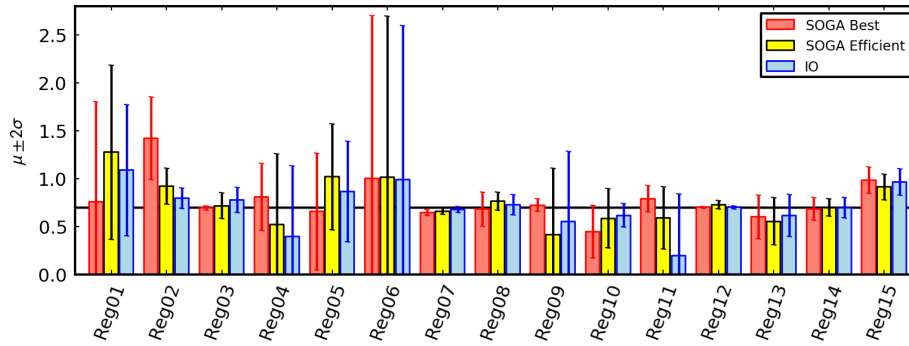
Figure 6. The figure shows the posterior weights for the emissions from the 15 regions for the SOGA Best, SOGA Efficient, and IO networks shown in Fig. 5. The posterior weights are presented as $\mu \pm 2\sigma$, which excludes off-diagonal covariance contributions that are important for some regions. The target value of 0.7 is shown by the horizontal black line.

and are strikingly similar to the noise amplitude of $\rho = 20$ ppt set in Eq. (2). The inferred values are slightly lower than 20 ppt because the conditions used to add noise in Eq. (2) truncate the negative values in the noise distribution and reduce the variance. From this comparison, we conclude that our inversion scheme successfully retrieves the observation–model noise from the data, and shows that this term does not have to be prescribed or specified, as is often done in other emissions inversion applications. For inversions with real observations, model errors generally dominate the noise and are difficult to estimate. Evidence approximation provides a way to account for these errors under approximate Gaussian assumptions.

5.3 Optimization of performance and cost

MOGA is used to jointly optimize the performance and cost of the HFC-134a observing network and to estimate the Pareto frontier between the two objectives. In the previous section, we showed that SOGA outperforms the IO optimization scheme (Patra and Maksyutov, 2002), both in terms of efficiency and effectiveness, for a single objective network design. This comparison suggests that the genetic algorithm will also perform well on the multiobjective problem. However, we cannot compare MOGA to IO in this section because the IO method is not designed for multiobjective network optimization.

The plots in Fig. 7 show the evolution of the performance and cost objectives in MOGA through 12 000 raw objective function evaluations. Even though SOGA and MOGA share many of the same algorithmic components, they solve different problems and therefore evolve their populations in different ways. Relative to the SOGA plots, the raw MOGA function evaluation plots appear noisier, have a periodic behavior, and do not easily show convergence. These are, however, expected features because MOGA is co-evolving the designs in two objective dimensions, trying to simultaneously optimize combinations of performance and cost. Because the performance and cost objectives are conflicting, these MOGA plots

do not show the same evolutionary changes as occurred in the SOGA experiment. One of the key differences is the periodic-like behavior in MOGA, which results from population members that are evolved to span desirable combinations of objectives through the niching operators. Close inspection of the cost objective plot indicates that the oscillations do not have a fixed period. The peak-to-peak spacing increases with function evaluations. This occurs because more evaluations are needed at later stages of the algorithm to improve the objectives, which provides an indicator of convergence.

Because it is difficult to ascertain convergence through the raw objective function evaluation plots, Fig. 8 shows the evolution of MOGA in terms of population generations. The figure specifically shows the members of the population with the lowest objective values at each generation from initialization through 149 generations. Population members that are not dominated by other members are carried forward through subsequent generations. From this figure, it is clear that MOGA is evolving networks to optimize performance and cost. As with the SOGA experiment, there is rapid improvement during the early phase of the optimization from initialization through about 10 generations. This period is followed by another period from about 10–60 generations with a slower rate of improvement. Cost is optimized more quickly than performance over this intermediate period because cost is based on a relatively simple expression, while performance is based on a complex atmospheric model. Around generation 60, the cost objective obtains its overall minimum value during the displayed evolution ($f_2 = 0.0094$). The performance objective, on the other hand, reaches its overall minimum value during the evolution on generation 116 ($f_1 = 1.126$).

Figure 8 only shows the evolution toward the extreme points of the Pareto frontier. To show the behavior along the whole frontier, Fig. 9 displays the populations of six-member observing networks for all of the generations in the two-dimensional objective space. Because non-dominated population members are retained from one generation to
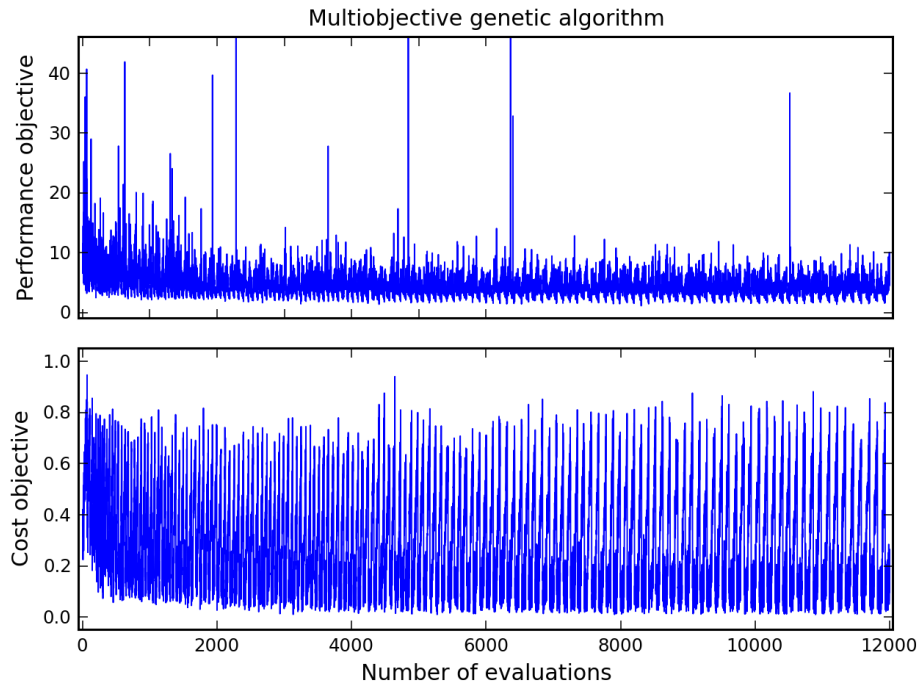
Figure 7. The figure displays the raw objective function evaluations during the evolution of a population of network designs using MOGA to optimize performance (upper panel) and cost (lower panel).
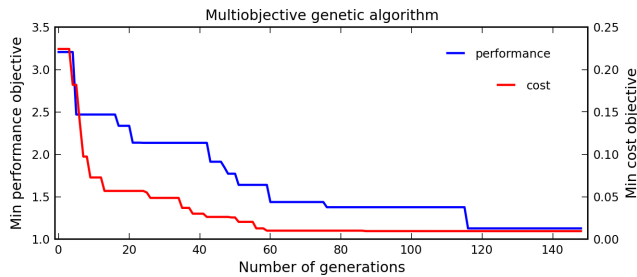

Figure 8. The figure displays the minimum value of the performance objective (blue line) and cost objective (red line) for each generation during the evolution of a population of network designs using MOGA.

the next, the figure contains more data points than the max_function_evaluations limit. The circles in the figure are color-coded by measurement frequency and have sizes based on their generation number. The smallest and largest circles correspond, respectively, to the earliest and latest generations. The sizes of the circles get progressively larger going from the upper right portion of the figure to the lower left. This provides a clear indication that the algorithm is evolving a set of potential solutions toward the Pareto frontier, which is represented by the approximate convex hull of large circles at the leading edge on the lower left side (see points A–G in the figure).

The tradeoffs between performance and cost in Fig. 9 are obvious. Optimizing performance leads to high cost designs,

and vice versa. The cost objective therefore plays a very strong role in determining monitoring locations in a multiobjective framework. The figure also shows a clear relationship between the sampling frequency and cost; low frequency solutions (blue circles) are less expensive than high frequency solutions (orange and red circles). The position along the Pareto frontier controls the expected returns in trading one objective for another. Networks with the poorest performance (e.g., points F and G) can be improved significantly with only moderate increases in cost. For example, networks near point G have monitoring stations that make one measurement per day and are located close to existing sites. By slightly re-positioning one or two of the stations, networks near point F achieve large gains in performance without incurring high costs. Further performance improvements, however, face steeper cost increases. For example, costs double between points E and D, and quadruple between points D and A. This sharp increase in cost occurs for two reasons. Minimizing the inversion errors requires (1) higher sampling frequencies and (2) the construction of new monitoring stations that are located far from existing sites. Using the measurement frequencies displayed in the figure, the *effective* number of new monitoring sites for each network can be determined by subtracting the operational cost from the total cost. For instance, networks near points A and B have operational costs of 0.4, and setup costs of about 0.35 and 0.18, respectively. Because each new, isolated site has a cost of about 0.083 (i.e., sites with $d(z_i) > d_{max}$), networks
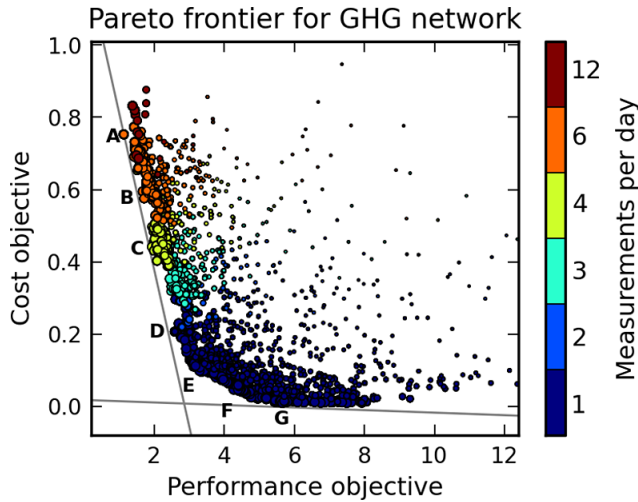
Figure 9. The figure displays the evolution the performance and cost objectives over generations of observing networks using MOGA. The stage of the evolution is denoted by circle size, with the earliest and latest generations corresponding to the smallest and largest circles, respectively. The measurement frequencies of the networks are color coded. Late generation points along the leading edge represent the approximate Pareto frontier, and points A–G are described in the text. The gray lines approximate the tangents to the objective minima, and their intersection defines the "utopia" point.

Figure 10. The figure shows the locations of observing stations in three networks that lie near the approximate Pareto frontier (see points A, B, and E in Fig. 9) using MOGA. Reference locations of the seven existing sites are also shown (white circles).

near points A and B create 4.2 and 2.2 effective new stations, respectively.

The station locations for three representative networks near points A, B and E along the Pareto frontier are shown in Fig. 10. It is important to note that these examples are only representative because networks that are adjacent in objective space can actually have stations that are far from each other in physical space. This occurs because the "fitness landscape" for this problem is extremely noisy (see Fig. 7), and the mapping from the design space to the objective space is not one to one (i.e., many designs can have nearly the same objectives). The point A and B cases are high cost, high performance examples, while the point E example lies close to the so-called "utopia" point, which is the point derived by intersecting the lines tangent to the objective extremes. As shown, network A has a high cost because it has only two stations that are nearby existing sites. Along with the high cost, network A has an associated high performance because it strategically places stations around large emitting regions but surprisingly uses only a single station in Southern California. Network B also has a relatively strong performance, but at a reduced cost because it deploys only two stations far from existing sites. The locations of the stations in networks A and B shown on the map are consistent with the approximate *effective* number of new sites estimated in the previous paragraph. The final example, network E, represents a design that attempts to achieve a balance between the cost and per-
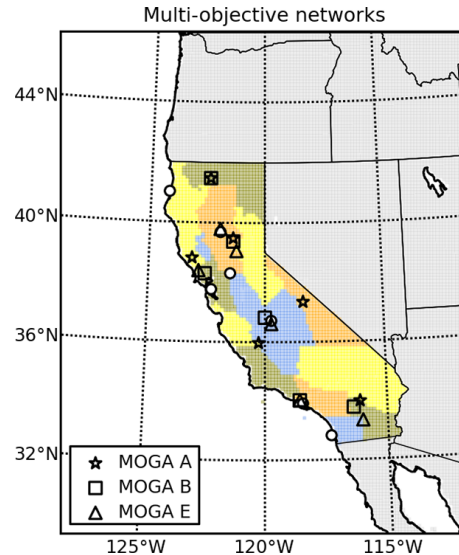
formance objectives by avoiding the steep portions of each. The resulting network has only one station far from an existing site (at the western edge of basin 13). It is also notable, but not unexpected, that none of these networks has a station near the existing sites at Trinidad Head and Scripps, California, which are used to monitor background GHGs through AGAGE (Prinn et al., 2000). These coastal sites are not well positioned to sample summertime HFC-134a emissions from California, which creates a challenge for doing top-down inversions with AGAGE data (Yver et al., 2011, 2013).

6 Summary and conclusions

In this report, we demonstrate the use of single objective and multiobjective genetic algorithms to design optimal observing networks to constrain GHG emissions through top-down inverse approaches. In particular, we use the algorithms to design a network of six stations to monitor HFC-134a emissions in California. The genetic algorithms search for station locations that optimize both the performance *and* cost of the network. When used to optimize only the performance of the observing network, the single objective genetic algorithm outcompetes an incremental optimization scheme that has previously been applied to CO_2. The genetic algorithm finds a better-performing network using fewer evaluations of the objective function. The performance-optimized stations are located relatively far from existing measurement sites, which indicates that current measurement networks could be improved to monitor HFC-134a or other GHGs with similar patterns of emissions.

Given a set of seven existing sites that could host observing stations at a minimal cost, the multiobjective genetic algorithm jointly optimizes the performance and cost of an HFC-134a observing network. The algorithm evolves different network configurations toward the Pareto frontier (i.e., the optimal combinations of the two objectives). The Pareto frontier is convex and clearly shows the tradeoffs between performance and cost. Low performing networks can be improved with minor increases in cost, but high performing networks require substantial increases in cost to achieve further improvements. The Pareto frontier thus provides a useful quantitative guide for decision makers to understand the tradeoffs in designing a GHG observing network. Because multiobjective genetic algorithms can easily accommodate additional, highly complex objectives that account for other GHGs and measurement modalities, we expect our method will provide a useful basis for designing practical GHG observing networks.

To better understand how the prototype GHG observing network could be extended to a real-world network design, we summarize below some of the key assumptions in our analysis. We have also released a data set of simulation time series to two public domain data repositories (Bache and Lichman, 2015; Lawrence Livermore National Laboratory Green Data Oasis, 2015). This data set can be used to test different inversion algorithms, optimization methods, cost functions, noise characteristics, and other assumptions that may impact the network design.

The structure of the noise used to generate the synthetic observations could affect the network. Although the noise differs from one location to another because different random seeds are used in Eq. (2), the noise has a constant amplitude and is spatially uncorrelated. These features are consistent with data that is independent and identically distributed, which is often a reasonable starting point for statistical analysis. In practice, however, GHG time series may be spatially correlated and have noise variations that scale with mixing ratio. By including spatially correlated noise in Eq. (2), we expect that the genetic algorithms would penalize stations that are close to each other because neighboring grid cells would experience similar fluctuations. However, the spatial correlation length scale is also expected to be relatively small (e.g., less than 10–20 km) because California has rough surface features and complex topography. The net effect of including spatially correlated noise on our analysis is therefore anticipated to be minor. By relaxing our constant noise amplitude assumption, on the other hand, we anticipate that the uncertainty in the inferred emissions of large emitting regions would increase, which would drive the optimization schemes to prefer stations near to those regions.

As a matter of convenience, we used the same measurement frequency at all of the stations in the network. Additional design variables could easily be introduced to optimize the location and frequency of each station, though the computational time to design the network would increase. We ex-

pect that such a change would result in a network with stations that collect measurements relatively more frequently in locations that are far from important sources (e.g., regions 1 and 6) than locations that are nearby (e.g., regions 7 and 12).

Last, we reiterate that the cost function used in the network design is idealistic. The form of the cost function is chosen to illustrate the notion of competing objectives (performance versus cost) and impart convexity to the Pareto frontier. Because we have more expertise on the performance aspects of network design than the cost side, it is difficult for us to extrapolate our results to situations involving realistic, detailed cost models. We invite researchers to use the publicly released data set to better explore the impacts of different cost decisions and models on network design.

References

Adams, B. M., Bohnhoff, W. J., Dalbey, K. R., Eddy, J. P., Eldred, M. S., Gay, D. M., Haskell, K., Hough, P. D., and Swiler, L. P.: DAKOTA: a Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis. Version 5.0 Users Manual, Tech. Rep. SAND2010-2183, Sandia National Laboratory, Albuquerque, New Mexico, USA, 2010.

Bache, K. and Lichman, M.: UCI Machine Learning Repository, available at: http://archive.ics.uci.edu/ml (last access: 12 May 2015), archived on 12 May 2015: http://www.webcitation.org/6YTndGkbi, 2015.

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., and Zhu, Z.: TransCom 3 inversion intercomparison: impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, Global Biogeochem. Cy., 20, GB1002, doi:10.1029/2004GB002439, 2006.

Barth, N. and Wunsch, C.: Oceanographic experiment design by simulated annealing, J. Phys. Oceanogr., 20, 1249–1263, 1990.

Bishop, C. M.: Pattern Recognition and Machine Learning (Information Science and Statistics), 1st Edn. 2006, corr. 2nd printing Edn., Springer-Verlag, New York, Secaucus, NJ, USA, 2007.

Carmichael, G. R., Sandu, A., Chai, T., Daescu, D. N., Constantinescu, E. M., and Tang, Y.: Predicting air quality: improvements through advanced methods to integrate models and measurements, J. Comp. Phys., 227, 3540–3571, 2008.

Carnevale, C., Finzi, G., Guariso, G. Pisoni, E., and Volta, M.: Surrogate models to compute optimal air quality planning policies at a regional scale, Environ. Modell. Softw., 34, 44–50, 2012.

Ciais, P., Dolman, A. J., Dargaville, R., Barrie, L., Bombelli, A., Butler, J., Canadell, P., and Moriyama, T.: GEO Carbon Strategy, Tech. rep., Geo Secretariat Geneva, FAO, available at: http://www.earthobservations.org/documents/sbas/cl/201006_geo_carbon_strategy_report.pdf (last access: 20 December 2014), 2010.

Ciais, P., Dolman, A. J., Bombelli, A., Duren, R., Peregon, A., Rayner, P. J., Miller, C., Gobron, N., Kinderman, G., Marland, G., Gruber, N., Chevallier, F., Andres, R. J., Balsamo, G., Bopp, L., Bréon, F.-M., Broquet, G., Dargaville, R., Battin, T. J., Borges, A., Bovensmann, H., Buchwitz, M., Butler, J., Canadell, J. G., Cook, R. B., DeFries, R., Engelen, R., Gurney, K. R., Heinze, C., Heimann, M., Held, A., Henry, M., Law, B., Luyssaert, S., Miller, J., Moriyama, T., Moulin, C., Myneni, R. B., Nussli, C., Obersteiner, M., Ojima, D., Pan, Y., Paris, J.-D., Piao, S. L., Poulter, B., Plummer, S., Quegan, S., Raymond, P., Reichstein, M., Rivier, L., Sabine, C., Schimel, D., Tarasova, O., Valentini, R., Wang, R., van der Werf, G., Wickland, D., Williams, M., and Zehner, C.: Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system, Biogeosciences, 11, 3547–3602, doi:10.5194/bg-11-3547-2014, 2014.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A Fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE T. Evolut. Comput., 6, 182–197, 2002.

Eddy, J. E. and Lewis, K.: Effective generation of Pareto sets using genetic programming, DETC2001/DAC-21094, in: Proceedings of ASME Design Engineering Technical Conference, ASME 2001 Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 9–12 September 2001, Pittsburgh, PA, USA, 2001.

EDGAR: Emission Database for Global Atmospheric Research (EDGAR), release version 4.1, European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL), available at: http://edgar.jrc.ec.europa.eu (last access: 20 December 2014), 2009.

Enting, I. G.: Green's function methods of tracer inversion, in: Inverse Methods in Global Biogeochem. Cy., edited by: Kasibhatla, P., Heimann, M., Rayner, P., Mahowald, N., Prinn, R. G., and Hartley, D. E., Vol. 114 of Geophys. Monogr. Ser., AGU, Washington, D.C., doi:10.1029/GM114p0019, 19–31, 2000.

Fischer, M. L. and Jeong, S.: Inverse Modeling to Verify California's Greenhouse Gas Emission Inventory, Tech. rep. ver. 1.3, Report for California Air Resources Board and the California Environmental Protection Agency, available at: http://www.arb.ca.gov/research/rsc/10-18-13/item5dfr09-348.pdf (last access: 20 December 2014), 2013.

Fonseca, C. M. and Fleming, P. J.: Genetic algorithms for multiobjective optimization: formulation, discussion and generalization, in: Proceedings of the Fifth International Conference on Genetic Algorithms, 17–21 July 1993, Urbana-Champaign, Illinois, USA, 416–423, 1993.

Gloor, M., Fan, S.-M., Pacala, S., and Sarmiento, J.: Optimal sampling of the atmosphere for the purpose of inverse modeling: a model study, Global Biogeochem. Cy., 14, 407–428, 2000.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmos. Environ., 39, 6957–6975, 2005.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Kowalczuk, E., Maki, T., Maksyutov, S., Peylin, P., Prather, M., Pak, B. C., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.-W.: TransCom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, Tellus B, 55, 555–579, 2003.

Huber, M.: Probabilistic framework for sensor management, Ph.D. thesis, Universitätsverlag Karlsruhe, Karlsruhe, 2009.

Jia, J., Chen, J., Chang, G., and Tan, Z.: Energy efficient coverage control in wireless sensor networks based on multi-objective genetic algorithm, Comput. Math. Appl., 57, 1756–1766, 2009.

Jonietz, K. K., Dimotakis, P. E., Rotman, D. A., and Walker, B. C.: A Greenhouse-Gas Information System: Monitoring and Validating Emissions Reporting and Mitigation, Tech. Rep. LA-UR-11-5512, Report for the US Department of Energy, Office of Science, prepared by Los Alamos National Laboratory, Los Alamos, New Mexico, USA, doi:10.2172/1033495, 2011.

Jourdan, L., Schuetze, O., Legrand, T., Talbi, E.-G., and Wojkiewicz, J. L.: An analysis of the effect of multiple layers in the multi-objective design of conducting polymer composites, Mater. Manuf. Process., 24, 350–357, 2009.

Judy, M. V., Ravichandran, K. S., and Murugesan, K.: A multi-objective evolutionary algorithm for protein structure prediction with immune operators, Comput. Methods Biomec., 12, 407–413, 2009.

Kasprzyk, J. R., Reed, P. M., Kirsch, B. R., and Characklis, G. W.: Managing population and drought risks using many-objective water portfolio planning under uncertainty, Water Resour. Res., 45, W12401, doi:10.1029/2009WR008121, 2009.

Kirkpatrick, S., Gelatt Jr., C. D., and Vecchi, M. P.: Optimization by simulated annealing, Science, 220, 671–680, 1983.

Kursawe, F.: A variant of evolution strategies for vector optimization, in: Parallel Problem Solving from Nature, edited by: Schwefel, H.-P. and Männer, R., Springer, Berlin, 193–197, 1991.

Lawrence Livermore National Laboratory Green Data Oasis, available through anonymous ftp at: ftp://gdo148.ucllnl.org/pub/, last access: 12 May 2015.

MacKay, D. J.: Bayesian Interpolation, Neural Comput., 4, 415–447, 1992.

Masutani, M., Woollen, J. S., Lord, S. J., Emmitt, G. D., Kleespies, T. J., Wood, S. A., Greco, S., Sun, H., Terry, J., Kapoor, V., Treadon, R., and Campana, K. A.: Observing system simulation experiments at the national centers for environmental prediction, J. Geophys. Res.-Atmos., 115, D07101, doi:10.1029/2009JD012528, 2010.

Mauger, G. S., Bumbaco, K. A., Hakim, G. J., and Mote, P. W.: Optimal design of a climatological network: beyond practical considerations, Geoscientific Instrumentation, Meth. Data Syst., 2, 199–212, doi:10.5194/gi-2-199-2013, 2013.

Morss, R. E., Emanuel, K. A., and Snyder, C.: Idealized adaptive observation strategies for improving numerical weather prediction, J. Atmos. Sci., 58, 210–232, 2001.

Morss, R. E., Miller, K. A., and Vasil, M. S.: A systematic economic approach to evaluating public investment in observations for weather forecasting, Mon. Weather Rev., 133, 374–388, 2005.

National Research Council: Verifying Greenhouse Gas Emissions: Methods to Support International Climate Agreements, The National Academies Press, Washington, D.C., USA, 2010.

Nickless, A., Ziehn, T., Rayner, P., Scholes, R., and Engelbrecht, F.: Greenhouse gas network design using backward Lagrangian particle dispersion modelling; Part 2: Sensitivity analyses and South African test case, Atmospheric Chemistry and Physics, 15, 2051–2069, doi:10.5194/acp-15-2051-2015, 2015.

Nisbet, E. and Weiss, R.: Top-down versus bottom-up, Science, 328, 1241–1243, 2010.

Pareto, V.: Cours D'Economie Politique, Rouge, Lausanne, 1896.

Patra, P. K. and Maksyutov, S.: Incremental approach to the optimal network design for CO_2 surface source inversion, Geophys. Res. Lett., 29, 1459, doi:10.1029/2001GL013943, 2002.

Prinn, R. G.: Measurement equation for trace chemicals in fluids and solution of its inverse, in: Inverse Methods in Global Biogeochem. Cy., edited by: Kasibhatla, P., Heimann, M., Rayner, P., Mahowald, N., Prinn, R. G., and Hartley, D. E., vol. 114 of Geophys. Monogr. Ser., AGU, Washington, D.C., doi:10.1029/GM114p0003, 3–18, 2000.

Prinn, R. G., Weiss, R. F., Fraser, P. J., Simmonds, P. G., Cunnold, D. M., Alyea, F. N., O'Doherty, S., Salameh, P., Miller, B. R., Huang, J., Wang, R. H. J., Hartley, D. E., Harth, C., Steele, L. P., Sturrock, G., Midgely, P. M., and Mcculloch, A.: A history of chemically and radiatively important gases in air deduced from ALE/GAGE/AGAGE, J. Geophys. Res., 115, 17751–17792, 2000.

Prinn, R., Heimbach, P., Rigby, M., Dutkiewicz, S., Melillo, J. M., Reilly, J. M., Kicklighter, D. W., and Waugh, C.: A Strategy for a Global Observing System for Verification of National Greenhouse Gas Emissions, Tech. Rep. Report 200, MIT Joint Program on the Science and Policy of Global Change, available at: http://globalchange.mit.edu/files/document/MITJPSPGC_Rpt200.pdf (last access: 20 December 2014), 2011.

Rayner, P. J.: Optimizing CO_2 observing networks in the presence of model error: results from TransCom 3, Atmos. Chem. Phys., 4, 413–421, doi:10.5194/acp-4-413-2004, 2004.

Rayner, P. J., Enting, I. G. and Trudinger, C. M.: Optimizing the CO_2 observing network for constraining sources and sinks, Tellus B, 48, 433–444, doi:10.3402/tellusb.v48i4.15924, 1996.

Sarigiannis, D. and Saisana, M.: Multi-objective optimization of air quality monitoring, Environ. Monit. Assess., 136, 87–99, doi:10.1007/s10661-007-9725-z, 2008.

Schaffer, J. D.: Multiple objective optimization with vector evaluated genetic algorithms, in: Proceedings of an International Conference on Genetic Algorithms and Their Applications, edited by: Grefenstette, J. J., Texas Instruments and NCARAI, 24–26 July 1985, Pittsburgh, PA, USA, 93–100, 1985.

Sergeyev, Y. D., Strongin, R. G., and Lera, D.: Introduction to Global Optimization Exploiting Space-Filling Curves, Springer, New York, NY, USA, 2013.

Skamarock, W. C. and Klemp, J. B.: A time-split nonhydrostatic atmospheric model for research and NWP applications, J. Comp. Phys., 227, 3465–3485, 2007.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.: A description of the Advanced Research WRF Version 2, Tech. Rep. NCAR Tech Notes-468+STR, NCAR, available at: http://www2.mmm.ucar.edu/wrf/users/docs/arw_v2_070111.pdf (last access: 20 December 2014), 2005.

Stuart, A. L., Aksoy, A., Zhang, F., and Nielsen-Gammon, J. W.: Ensemble-based data assimilation and targeted observation of a chemical tracer in a sea breeze model, Atmos. Environ., 41, 3082–3094, 2007.

Thompson, R. L., Gerbig, C., and Rödenbeck, C.: A Bayesian inversion estimate of N_2O emissions for western and central Europe and the assessment of aggregation errors, Atmos. Chem. Phys., 11, 3443–3458, doi:10.5194/acp-11-3443-2011, 2011.

Trujillo-Ventura, A. and Hugh Ellis, J.: Multiobjective air pollution monitoring network design, Atmos. Environ. A, 25, 469–479, 1991.

Weiss, R. F. and Prinn, R. G.: Quantifying greenhouse-gas emissions from atmospheric measurements: a critical reality check for climate legislation, Philos. T. Roy. Soc. A, 369, 1925–1942, doi:10.1098/rsta.2011.0006, 2011.

Yver, C., Lucas, D. D., Graven, H. D., Cameron-Smith, P. J., Keeling, R. F., and Weiss, R. F.: A framework for improving top-down estimates of halocarbon emissions through synthetic multi-tracer inversions, A13J-05, AGU Fall Meeting, 5–9 December 2011, San Francisco, CA, USA, 2011.

Yver, C. E., Graven, H. D., Lucas, D. D., Cameron-Smith, P. J., Keeling, R. F., and Weiss, R. F.: Evaluating transport in the WRF model along the California coast, Atmos. Chem. Phys., 13, 1837–1852, doi:10.5194/acp-13-1837-2013, 2013.

Ziehn, T., Nickless, A., Rayner, P. J., Law, R. M., Roff, G., and Fraser, P.: Greenhouse gas network design using backward Lagrangian particle dispersion modelling; Part 1: Methodology and Australian test case, Atmo. Chem. Phys., 14, 9363–9378, doi:10.5194/acp-14-9363-2014, 2014.

Zitzler, E. and Thiele, L.: Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach, IEEE T. Evolut. Comput., 4, 257–271, 1999.

Zitzler, E., Laumanns, M., and Thiele, L.: SPEA2: Improving the strength pareto evolutionary algorithm for multiobjective optimization, in: Evolutionary Methods for Design, Optimisation and Control with Application to Industrial Problems, edited by: Giannakoglou, K., EUROGEN 2001, 19–21 September 2001, Athens, Greece, 95–100, 2002.