1-1-2012

Atmospheric tomography: a bayesian inversion technique for determining the rate and location of fugitive emissions

Ruhi Humphries
University of Wollongong, rsh615@uow.edu.au

Charles Jenkins
CSIRO

Ray Leuning
CSIRO Marine and Atmospheric Research, Australia

Steve Zegelin
CSIRO

David Griffith
University of Wollongong, griffith@uow.edu.au

See next page for additional authors

Atmospheric tomography: a bayesian inversion technique for determining the rate and location of fugitive emissions

Abstract

A Bayesian inversion technique to determine the location and strength of trace gas emissions from a point source in open air is presented. It was tested using atmospheric measurements of N2O and CO2 released at known rates from a source located within an array of eight evenly spaced sampling points on a 20 m radius circle. The analysis requires knowledge of concentration enhancement downwind of the source and the normalized, three-dimensional distribution (shape) of concentration in the dispersion plume. The influence of varying background concentrations of ~1% for N2O and ~10% for CO2 was removed by subtracting upwind concentrations from those downwind of the source to yield only concentration enhancements. Continuous measurements of turbulent wind and temperature statistics were used to model the dispersion plume. The analysis localized the source to within 0.8 m of the true position and the emission rates were determined to better than 3% accuracy. This technique will be useful in assurance monitoring for geological storage of CO2 and for applications requiring knowledge of the location and rate of fugitive emissions

Keywords

technique, determining, emissions, inversion, fugitive, tomography, location, rate, bayesian, atmospheric, GeoQUEST

Disciplines

Life Sciences | Physical Sciences and Mathematics | Social and Behavioral Sciences

Publication Details

Humphries, R., Jenkins, C., Leuning, R., Zegelin, S., Griffith, D., Caldow, C., Berko, H. & Feitz, A. (2012). Atmospheric tomography: a bayesian inversion technique for determining the rate and location of fugitive emissions. Environmental Science and Technology, 46 (3), 1739-1746.

Authors

Ruhi Humphries, Charles Jenkins, Ray Leuning, Steve Zegelin, David Griffith, Christopher Caldow, Henry Berko, and Andrew Feitz

Atmospheric Tomography: A Bayesian inversion technique for determining the rate and location of fugitive emissions.

Ruhi Humphries[1], Charles Jenkins[2,3]*, Ray Leuning[2,4], Steve Zegelin[4], David Griffith[1], Christopher Caldow[1], Henry Berko[5] and Andrew Feitz[5]

[1] Centre for Atmospheric Chemistry, University of Wollongong, Wollongong, NSW, 2522, Australia

[2] Cooperative Research Centre for Greenhouse Gas Technologies (CO2CRC), GPO Box 463, Canberra, ACT, 2601, Australia

[3]CSIRO Energy Transformed Flagship, PO Box 3023, Canberra, ACT, 2601, Australia.

[4] CSIRO Marine and Atmospheric Research, PO Box 3023, Canberra, ACT, 2601, Australia

[5]Geoscience Australia, GPO Box 378, Canberra, ACT, 2601, Australia

Abstract

A Bayesian inversion technique to determine the location and strength of trace gas emissions from a point source in open air is presented. It was tested using atmospheric measurements of N_2O and CO_2 released at known rates from a source located within an array of eight evenly spaced sampling points on a 20 m radius circle. The analysis requires knowledge of concentration enhancement downwind of the source and the normalized, three-dimensional distribution (shape) of concentration in the dispersion plume. The influence of varying background concentrations of ~1% for N_2O and ~10% for CO_2 was removed by subtracting upwind concentrations from those downwind of the source to yield only concentration enhancements. Continuous measurements of turbulent wind and temperature statistics were used to model the dispersion plume. The analysis localized the source to within 0.8 m of the true position and the emission rates were determined to better than 3% accuracy. This technique will be useful in assurance monitoring for geological storage of CO_2 and for applications requiring knowledge of the location and rate of fugitive emissions.

1 Introduction

Carbon capture and storage (CCS) in an emerging technology for reducing CO_2 emissions to the atmosphere [1]. Because of the associated energy and carbon costs in CCS, even a small rate of leakage from geological storage over a prolonged period could result in more CO_2 being emitted in the long-term than would otherwise occur [1-3]. Reliable monitoring techniques must therefore be developed to locate and quantify leakages in order to fix leaks, confirm the long-term benefit of CCS to mitigating climate change and to underpin credits needed for carbon pricing schemes.

Many techniques are available for monitoring geologically stored CO_2 above and below ground [4] but direct atmospheric measurements of leakage have the advantage of focusing on the key purpose of CCS. A concise review of inversion analyses of concentration data to recover source strength and location is given by Shankar Rao [5]. Bayesian methods have been used extensively in large-scale atmospheric modeling, and have also been advocated and used for inversions on smaller scales with data captured over short periods in steady winds. A comprehensive summary of Bayesian inversion and its advantages is given by Yee and Flesch [6], with numerous other expansive studies undertaken [7-11]. Of particular importance are the abilities of Bayesian methods to incorporate prior information, to independently solve for the probability distribution of individual parameters, to quantify uncertainties from measurements and models, and to generalize to quite complex cases [8]. Bayesian inversion is conceptually quite distinct from parameter estimation of assumed source distributions (e.g. [12]), although in simple cases, can appear mathematically similar.

In the current paper, Bayesian model inversion techniques are combined with knowledge of the, three-dimensional distribution (shape) of concentration in the dispersion plume to provide both the position and strength of a source using concentrations measured at points surrounding a suspected source. Unlike previous studies which used tracers with minimally varying backgrounds [6, 7, 9, 10, 13]**,** this study investigates small, hypothetical CCS leaks [3] which involve increases in CO_2 concentrations of only a few parts per million (ppm) above a typical diurnal range of 370 to 500 ppm [14, 15]. The influence of this background variability can be reduced considerably using an array of sensors that encircle the suspected leak location. Some sensors will then be upwind of the source irrespective of wind direction, thereby allowing the variable background concentration to be subtracted from the downwind measurement. Because such leaks are likely to be small and long-term [1], monitoring can be performed over long periods, allowing errors in background estimations to average away. Hence, perturbations which may be less than the variability of the background can be detected. Multiple wind directions are exploited to localize sources and determine their strength making the analysis analogous to tomographic imaging. The large data set associated with this long-term monitoring scheme also facilitates selection of favorable periods, such as day-time, when background concentrations are relatively steady.

Although Bayes' theorem can be applied to quite complex cases, this paper will describe the simplest and most accurate application where the leak is spatially small and temporally constant. The use of Bayes' theorem to determine the position and strength of emission sources is presented in detail elsewhere (e.g. [6, 7]) and only a brief outline is given here. The paper presents results of a small-scale field experiment designed to test the ability of this approach to determine both the position and source strengths of both CO_2 and a tracer (N_2O), released at known rates over several months under a range of meteorological conditions. The potential utility of atmospheric tomography for monitoring emissions from CO_2 geological storage sites is also discussed.

2 Materials and Methods

2.1 Field Site

Measurements were made from 1 July to 30 October 2010 at the CSIRO Ginninderra Experiment Station near Canberra, Australia. The site, shown in Figure 1, was a flat, recently ploughed field surrounded by others used for sheep grazing and cropping. A 6 m high earth-fill dam wall was 150 m north-west of the release point. A caravan and shed housing the instrumentation were located ~35 m west-north-west of the release point. Average temperature during the four-month campaign was 8°C (range -3°C to 25°C). Mean rainfall was 72 mm $month^{-1}$ and north-west winds occurred > 60% of the time.

Figure 1: Schematic representation of the experimental site. The numbered yellow dots indicate air intake positions, Source 1 is the location of both the CO_2 and N_2O release from the centre, and Source 2 is where the N_2O was moved after some time for the blind test.

2.2 Emission Sources

Mass flow controllers controlled the simultaneous release of N_2O and CO_2, piped via Dekabon® tubing to a mixing chamber, 0.30 m above ground. N_2O was released at 1.070 ± 0.04 g min^{-1}, the rate required to get at least 10 % (~30 ppb) enhancement at downwind sampling points in most atmospheric conditions. CO_2 was released at 56.65 ± 0.8 g min^{-1}, chosen to simulate a leak barely detectable above background variability in unstable, well-mixed conditions when background concentrations approximated a normal distribution with a standard deviation of 3.5 ppm. Figure 1 shows the two locations of the source during the campaign: Source 1 at the centre of the circle (known); and Source 2, which was unknown prior to the inverse analysis.

2.3 Concentration Measurements

Figure 1 shows the eight sampling points positioned every 45° on a 20 m radius. Air from intakes at 1.5 m above ground was pumped continuously through equal lengths of conjoined polyethylene and Dekabon® tubing (to minimize adsorption or migration of atmospheric constituents across tube walls) that was insulated and kept ~1 m above the ground to prevent condensation of water in the lines. An auxiliary switching manifold that allowed continuous flushing and autonomous switching of all 8 sampling lines was used to sample air into a Fourier Transform Infrared (FTIR) Spectrometer [16] for analysis of CO_2 and N_2O mixing ratios. All sampling lines were filtered at sampling points and before lines entered instrumentation (AF20-02-C, SMC Pneumatics, Sydney, NSW, Australia) then dried using Nafion® membrane and magnesium perchlorate prior to entering the FTIR. All 8 lines were analyzed sequentially within a 30 minute period, allowing enough time for FTIR precision of 0.025 ppm for CO_2 and 0.1 ppb for N_2O to be attained, but quick enough to

characterize changing atmospheric conditions. Air from radially-opposite intakes was analyzed sequentially to minimize the delay between upwind and downwind measurements.

2.4 Micrometeorology Measurements

Temperature and the x, y, and z components of the wind vector were measured at 20 Hz using a three-dimensional sonic anemometer (Type HS, Gill Instruments Ltd., Lymington, UK) mounted at 1.42 m. These data provided the Monin-Obukhov stability parameter (MO length, L, [17]), friction velocity, u^*, mean horizontal wind speed, U, and wind direction, θ. These quantities are required to model the shape and direction of the dispersion plume downwind of the source. The MO length provides a measure of atmospheric stability that controls horizontal and vertical mixing in the surface boundary layer, while friction velocity is a measure of the flux density of momentum to the surface. Using these turbulence statistics, the shape of the down-wind concentration plume can be calculated by a stochastic Lagrangian method [18, 19]. Because the MO theory for dispersion is unreliable under light wind conditions, both FTIR and micrometeorological data were removed when $u^* \leq 0.15$ m s^{-1} [20].

2.5 Data Analysis: Bayesian Tomography

Estimating the position and strength of the source from a sensor network is an ill-posed inverse problem, which does not have a unique solution because of noise in the data and inadequacies in the dispersion model. As discussed in detail by Keats et al. [7] and by Yee and Flesch [6], the most probable solution may be determined using Bayes' Theorem to determine probability distributions for the locations and source strengths of fugitive emissions given experimental data, D, (background-subtracted concentrations, meteorological data and a dispersion model) and prior information (likely number of sources and their approximate location). The Bayesian method has the advantage of being able to handle a wide range of assumptions in its applications. In the current study, the simplest case is adopted.

For the simplest case of a single, continuous and steady source, Bayes' Theorem states [6, 7]

$$P\left(s, x_0, y_0 \mid D\right) = \frac{P\left(D \mid s, x_0, y_0\right) P(s, x_0, y_0)}{P(D)} \qquad (1)$$

where $P(s,x_0,y_0 \mid D)$ is the *posterior probability* of a source with strength s located at $(x_0\ ,\ y_0)$ given the data D; $P(D \mid s,x_0,y_0)$ is the *likelihood* of D given s, x_0 and y_0, and $P(s,x_0,y_0)$ is the *prior probability* of s, x_0 and y_0 that is assigned before data are obtained. The normalizing factor, $P(D)$ is the *evidence*. It is not needed in our application but could be used to test the plausibility of competing models, e.g. a point vs. line source.

The likelihood can be calculated if the probability distribution of errors in the data is known. In this application, the prior captures the preexisting knowledge that the source is most likely located

within the sensor array. Both the prior and the likelihood can be expressed analytically and a function can be formed for the posterior probability, $P(s,x_0,y_0 \mid D)$.

2.5.1 Plume Functions

A Lagrangian stochastic (LS) model [18], as implemented in the WindTrax software (Thunder Beach Scientific, Nanaimo, Canada), was used to calculate plume shapes for the range of stability classes encountered during the field campaign. Run in forward mode, 10^5 'particles' were released and their unique paths tracked to develop a plume shape for each stability class. Model concentrations were sampled at a height of 1.5 m, every 0.5 m over a 50 m x 50 m grid.

The computed plume shapes calculated using the LS model are noisy because of the finite number of particles released. They are also time-consuming to calculate and it was computationally more efficient to fit the following interpolation function, based on the theoretical cross-wind integrated function described by Hsieh *et al.* [21], to the 2-dimensional plume shapes at 1.5 m calculated using the LS model:

$$f\left(x,y,x_0,y_0,s\right)=f\left(x-x_0,y-y_0,s\right)=\frac{a\,s}{\left(x-x_0\right)^{\alpha}}e^{-\frac{d}{(x-x_0)^{\zeta}}-(y-y_0)^2\frac{\left(b+(x-x_0)^{\beta}\right)^2}{c(x-x_0)^{\gamma}}} \qquad (2)$$

where x and y are distances downwind and crosswind of the source at (x_0, y_0) respectively, s is the source strength, and the parameters a, b, c, d, α, β, γ, ζ were calculated separately for each defined stability class. Plume functions can be rotated to lie along the wind direction using a standard rotation matrix applied to the displacement vector of the argument of the plume equation. This enabled the measured wind direction for any particular period to be utilized in calculations.

2.5.2 Background Concentrations

A two-stage iteration was used to calculate background concentrations. The first iteration used measurements on the upwind air intake of the opposing member of the pair most closely aligned with the wind direction. This value was subtracted from measurements of concentration in the downwind intake and used with the inverse analysis to give the first estimate of the location and strength of the source. These results, combined with appropriate wind direction and modeled plumes, were used to identify those air intakes unaffected by the release. The identified unaffected air intake measurements were averaged to provide a second estimate of background concentration and an improved estimate of source position and strength. This second iteration enabled extraction of up to 7 perturbations within each 30 minute measurement period, dependent on stability conditions.

2.5.3 Calculating the Likelihood, $P(D|s, x_0, y_0)$

Suppose $f\ (x, y, x_0, y_0, \theta, L)$ is a function (equation (2)) giving the mixing ratio, for a unit source strength, at position (x, y), when the source is at (x_0, y_0), the wind direction is θ and the Monin-Obukhov length is L. This function can be used to predict the concentration at each sensor, given the

measured values of θ and L and multiplying by the source strength, s. For the i^{th} measurement, if the error distribution is Gaussian with a standard deviation, σ_i, the probability, P_i, for measuring the observed concentration, c (x_i, y_i, θ_i, L_i), is [22]

$$P_i \propto e^{-\frac{\left(c(x_i,y_i,\theta_i,L_i)-s\,f(x_i,y_i,x_0,y_0,\theta_i,L_i)\right)^2}{2\sigma_i^2}} \tag{3}$$

The probability of the N measurements is $P = \Pi\, P_i$. If σ_i, is constant, then we define

$$\chi^2 = \frac{1}{2\sigma^2}\sum_i^N \left(c\left(x_i,y_i,\theta_i,L_i\right)-s\,f(x_i,y_i,x_0,y_0,\theta_i,L_i)\right)^2 \tag{4}$$

the likelihood becomes

$$P = \Pi P_i \propto e^{-\chi^2} \tag{5}$$

2.5.4 Estimating the Prior, $P(s, x_0, y_0)$

The prior expresses what is known before any experimental data are obtained. It decreases in importance as the number of experimental measurements increases. It is assumed that the source is located somewhere inside our array of sensors, and is spatially small and weak. If s, x_0, and y_0 are statistically independent of each other,

$$P\left(s,x_0,y_0\right) = P\left(s\right)P\left(x_0\right)P(y_0) \tag{6}$$

By assumption, the probability of finding the source in the area within the sensor array is finite, and drops quickly to zero outside this area. $P(x_0)$ and $P(y_0)$ can be taken to be Gaussian around (x_0, y_0) such that,

$$P\left(x_0\right) = \sum_i^8 e^{-\frac{(x_i-x_0)^2}{2w^2}} \tag{7}$$

where x_i is the x coordinate of sensor i, and w is a constant indicating the prior uncertainty in the location, set to 20 m for this experiment. A similar equation can be written for $P(y_0)$.

Since the probability of a scaled variable is being considered, the most agnostic prior that can be adopted is commonly known as Jeffreys' prior [23, 24]:

$$P\left(s\right) \propto \frac{1}{s} \tag{8}$$

This prior suits this problem well because it favours small leaks, the target of this technique. The model will produce a poor fit if the actual leak is large or variable, indicating the prior may have to be generalised. A complete Bayesian analysis would use several competing models for source and plume geometry to determine their posterior probability. Here the focus is on the simple situation of one plausible model.

2.5.5 Marginalization

Marginalization is an important step in the analysis which allows independent calculation of either location or source strength. Marginalization involves integrating the probability distribution over a subset of unknown variables in order to calculate a probability that is a function only of the remaining variable(s). The probability distribution for location is calculated by elimination of the source strength:

$$P(x_0, y_0 \mid D) = \int_{s_l}^{s_u} P\left(s, x_0, y_0 \mid D\right) ds \tag{9}$$

where s_u and s_l are upper and lower limits of the range of plausible values of source strength respectively. The source strength probability distribution is calculated similarly:

$$P(s \mid D) = \int_{x_{0_l}}^{x_{0_u}} \int_{y_{0_l}}^{y_{0_u}} P\left(s, x_0, y_0 \mid D\right) dy_0 dx_0 \tag{10}$$

where x_{0_u}, x_{0_l}, y_{0_u} and y_{0_l} are the upper and lower limits of the range of plausible values of the source location coordinates. The integrand of the above integrals has the form

$$P(s, x_0, y_0 \mid D) = \frac{1}{sP(D)} g\left(x_0, y_0\right) e^{-\chi^2} \tag{11}$$

where g(x_0, y_0) is a product of equation (7) for both x_0 and y_0, $P(D)$ is the evidence, and χ^2 is a function of s, x_0 and y_0 (equation (4)). This function contains N terms in χ^2 alone, each with exponential functions meaning that integrations had to be performed numerically. Determination of the most likely values and uncertainties of s, x_0 and y_0 were done by inspection of the probability distributions of each variable, which were calculated and plotted over the integral domain using Wolfram Mathematica software (Wolfram Research, Inc., Champaign, USA).

3 Results

3.1 Background Characterization

Atmospheric N_2O concentrations recorded for five weeks prior to N_2O release showed that the FTIR analyzer (1σ repeatability ~0.1 ppb for N_2O) could resolve the low background variability in N_2O (2.2 ppb short-term, 8 ppb long-term). Variations in N_2O concentrations seen in Figure 2 are caused by the changing activity of N_2O emitting soil bacteria (longer term variation) and emitted N_2O mixing into a surface boundary of variable depth during the diurnal cycle of atmospheric stability (shorter term variation). This same interaction between boundary layer depth with local sinks and sources (daytime photosynthesis and nighttime respiration for CO_2 respectively) causes background CO_2 concentrations to vary from 380 – 500 ppm (Figure 2).

Detection of small, short term increases in concentrations above background levels is required as the release plume passes over a sensor. Background concentrations of CO_2 vary on similar timescales to these perturbations making the Bayesian analysis for CO_2 more difficult than for N_2O.

Examination of the no-release data obtained over several weeks showed that errors in the background-subtracted concentrations decrease as $N^{-1/2}$ as expected for uncorrelated fluctuations (N is the number of measurements). If drifts in the instrument calibration are small or can be measured, the combination of averaging and Bayesian model-fitting should allow the effective use of data where the concentration perturbations due to sources are well below the instantaneous noise level.

Figure 2: Background variability of N_2O (top) and CO_2 (bottom). N_2O is significantly more stable compared to CO_2. This high relative stability is caused by weaker sources and sinks of N_2O.

3.2 Atmospheric Tomography

3.2.1 Simulated Data

The analysis scheme was first tested using simulated data as input. An array of wind direction (θ) and stability conditions (L) was generated at random within suitable ranges and the forward model was used to calculate concentrations at all sensor locations from sources at various positions. A noise component was added from a normal distribution with $\mu = 0$ ppm and $\sigma = 20$ ppm. This procedure generated realistic simulated data which were then used as input to the Bayesian inversion.

Analyses of the simulated datasets showed no systematic errors in the code. The source strength was estimated accurately with a precision determined solely by the number of data points used. The analyses located the sources to within 1 m, regardless of their position within the sensor array (data not shown). This corresponds to an uncertainty of 5% of the radius of the sensor array. Results were similar even when the source was placed outside the array (e.g. 30 m, -30 m).

3.2.2 N_2O at Origin

N_2O was released at a rate of 1.070 ± 0.04 g min^{-1} at the centre of the ring of sensors (referred to hereafter as the origin) between 12 August and 12 October 2010. This resulted in a ~30 ppb increase in mixing ratio at sensors 20 m downwind from the source under well-mixed daytime conditions and > 300 ppb under stable, nighttime conditions (Figure 2, Supporting Information).

Figure 3 shows that the logarithmic posterior probability distribution for the source location has a narrow ridge along the prominent north-westerly wind direction (78% of the 1569 measurement periods consisted of plumes directed over sensor 3). The peak is taken as the most likely source location at (0.0 m, 0.0 m) with peak shape represented by an ellipse with major and minor semi-axes of 0.9 m and 0.4 m (2σ) respectively, rotated 60^{o} east of north. Using a smaller dataset (65% of data removed) with a more even distribution of wind directions gave a more symmetrical probability distribution for the source location, with a narrower peak with major and minor semi-axes 0.1 m each, still rotated 60° east of north (Figure 3 in Supporting Information).

The accuracy and precision in calculating the source strength increases with more data, independent of wind direction. Using the full data set, the source strength was determined (Figure 3) to be 1.103 ± 0.06 g min^{-1} (μ ± 2σ), which was within 3.1% of the correct value.

Figure 3: Probability distributions calculated using the full set of experimental data collected while N_2O was being released at 1.070 ± 0.04 g min^{-1} from the origin. The location probability distribution (log scale) shows the effect of sensor locations and the dominant wind direction on the distribution. The probability distribution for source strength, with a peak at 1.103 ± 0.06 g min^{-1}, shows excellent accuracy and precision (red curve) comparable to the true release (purple bar).

3.2.3 N_2O at Unknown Location

The N_2O source was repositioned to a location initially unknown to the analysis team from 12 - 30 October while the N_2O release rate remained unchanged. Filtering out ~50% of the data to give an even distribution of wind directions, the analysis placed the source at (6.2 m, 0.5 m), just 0.8 m from the actual location at (6.4 m, 1.3 m), with an uncertainty described by elliptical peak axes (2σ): major 3.7 m; minor 1.0 m; rotated 60° east of north (Figure 4, Supporting Information). The error in source location is 4% when normalized by the radius of the instrument array. Source strength calculated using all available data was 1.12 ± 0.1 g min^{-1}, within 4.7% of the correct value. Due to the much smaller data set (18 days, 438 measurement periods), results of this analysis were less accurate than for release at the origin (61 days, 1569 measurement periods).

3.2.4 CO_2 at Origin

Release of CO_2 occurred between 10 September and 30 October 2010 at a rate of 56.65 ± 0.8 g min^{-1}. Peaks in CO_2 concentration during the release period are seen in Figure 4 to be clearly discernible, particularly when background concentrations were comparatively stable under well-mixed atmospheric conditions. These enhancements were just 1% of background concentrations. Atmospheric mixing was generally suppressed at night and resulted in high and variable concentrations and low u^* values. Much of the data for these periods were not used in the analysis.

Figure 4: Raw CO_2 mixing ratios from all 8 sampling locations measured over two consecutive nights during CO_2 release at 56.65 ± 0.8 g min^{-1}. Perturbations in concentrations are discernible from background variations during well mixed conditions in the daytime and on the night of 18 September. Concentrations vary greatly during stable overnight conditions (e.g. on the night of 19 September), making it difficult to discern perturbations in the data due to the source. The inset shows the variation in CO_2 mixing ratios measured sequentially at each of the 8 inlets (solid line) and the first iteration of background-subtracted values (connected dots). The latter are generally < 1ppm except for the sensor downwind of the source.

Good localization of the source was achieved using the 50 day (1000 measurement periods) data set, shown in Figure 5. The distribution in the logarithmic posterior probability plot is similar to

that for the N_2O source. Using a dataset with ~45% of data removed for even weighting of wind directions, the peak of the probability distribution for the source at (-0.5 m, 0.5 m) was within 0.8 m (4%) of the true position (0 m, 0 m), with uncertainty described by elliptical peak axes (2σ): major 1.1 m; minor 0.6 m; rotated 60° east of north. The estimated source strength was 54.9 ± 4 g min^{-1} (μ ± 2σ), statistically equal to the true value. Compared to results for N_2O, the larger errors for CO_2 source location and strength were due to the smaller concentration perturbations relative to background (CO_2 ≈ 1%, N_2O ≈ 36%), the smaller number of measurements, and the difference in wind direction data caused by different timing of respective releases.

Figure 5: Probability distributions for experimental data collected while CO_2 was being released at 56.65 ± 0.8 g min^{-1} from the origin. The location distribution was calculated using a reduced, more evenly weighted data set, while the source strength distribution was calculated using the full data set available for analysis.

3.2.5 Determining the required number of measurements

The number of measurements necessary for accurate localization of the source depends strongly on having a variety of wind directions, while for source strength, the most important factor is having a large number of measurements under a range of stability conditions. The time required for convergence depends on the required accuracy of a particular application. Figure 5 in the Supporting Information shows convergence to within 5% of the correct source strength with ~1000 measurements (at a rate of up to 14 per hour when $u^* > 0.15$ m s^{-1} in our field experiment).

3.2.6 Importance of characterizing dispersion

The effect of model errors was tested by creating simulated data with a range of stability conditions, but analyzing using only one plume. Results showed that localization remained just as accurate, although became less precise, while for source strength, both accuracy and precision were affected by up to 30%. This result suggests that errors in dispersion models will provide complications, however the method employed here is robust enough to deal with it, producing useful, albeit lower quality results.

4 Discussion

The tomographic technique used in this study has been shown to determine both source strength and location well on a small scale, using high-precision instrumentation to measure concentration perturbations. Pumps and tubing were used to sample air for analysis in this field experiment but this is unlikely to be practical for applications at larger scales. At these scales, a network of independent sensors will need to be deployed, likely lowering the resolution of the technique because of inherent inter-sensor variability.

Simple linear scaling of experimental results combined with Lagrangian stochastic modeling under various atmospheric stabilities suggests that for a sensor located 1 km downwind, a point source

emission of ~5 kg min^{-1} would be required in all but the most turbulent conditions to increase CO_2 concentrations by the required 1% above background (in the most unstable conditions, the required emission rate increases by a factor of ~3). This equates to a detectable emission of 2.6 kt CO_2 yr^{-1}. Enting et al. [3] calculate that for CCS to be beneficial in the long-term, a leak rate below 0.01% per year is required. This equates to a leak rate of 1000 t yr^{-1}, assuming a modestly-sized industrial storage site of 10 Gt. The required measurement time depends on required accuracy, the source strength, atmospheric stability, topography and the distances between sources and detectors.

More complex situations such as multiple sources, or the optimum location of extra detectors, can be handled in the Bayesian framework [8, 23]. Complex spatial sources, including multiple point, line and area sources [5-7] can be analyzed by the tomographic method by acquiring data from enough distinct wind directions for disambiguation. In the current analysis, temporally varying sources are not resolved but rather the average rate is calculated. Time resolved rates could be calculated by solving for source strengths within time sub-sets of data. More complex terrain will introduce more complex dispersion which, as discussed in Section 3.2.6, appears likely to increase the uncertainty in source strength, but may affect the determination of location less. Geological storage will probably mostly occur in sedimentary basins which often have low relief, so terrain may not be as much of a problem as it first seems. In real-world situations, background concentrations may also be more variable because of the greater heterogeneity of sources and sinks.

In addition to measurements of total CO_2, natural ($^{13}CO_2$, $^{14}CO_2$, CH_4,) and man-made (SF_6) tracers in the geologically stored CO_2 will assist in the detection, attribution and quantification of any emissions [7, 25, 26]. Atmospheric measurements alone are unlikely to have to bear the entire burden of the proof of a leak though. With early knowledge of the approximate location of a leak, additional measurements such as soil CO_2 flux and concentration measurements could be made to confirm predictions from atmospheric measurements.

Although qualitative short term (~1 week) localization is possible, high accuracy, quantitative localization is largely dependent on having data from a variety of wind directions, meaning that long term monitoring applications of this technique are ideal as they allow time for the variety of wind directions to be collected as well as the detection of perturbations of species in environments with highly variable backgrounds (e.g. CO_2). Possible areas of application include industrial emission verification (see Table 1 in Supporting Information) investigation of GHG emissions from groundwater bores, air pollution monitoring, and localizing fugitive emissions from gas wells, coal mines, pipelines and other oil and gas operations [27].

5 Acknowledgements

We thank Martin Riggenbach and Graham Kettlewell for technical support. We also thank Phil Dunbar and CSIRO Plant Industry for maintenance of the site and use of the Ginninderra Experiment

Station. We acknowledge the funding provided by the University of Wollongong, the Commonwealth of Australia to CSIRO, the Cooperative Research Centre for Greenhouse Gas Technologies and to Geoscience Australia. RH acknowledges support from the Centre for Atmospheric Chemistry at the University of Wollongong. AF and HB publish with the permission of the CEO, Geoscience Australia.

Supporting Information Available. This information is available free of charge via the Internet at http//pubs/acs/org.

6 References

1. Metz, B.; Davidson, O.; de Coninck, H.; Manuela Loos, H.; Meyer, L. *IPCC Special Report on Carbon Dioxide Capture and Storage* Intergovernmental Panel on Climate Change: 2005; p 443.
2. Haugan, P. M.; Joos, F., Metrics to assess the mitigation of global warming by carbon capture and storage in the ocean and in geological reservoirs. *Geophys. Res. Lett.* **2004,** *31*, (18), L18202.
3. Enting, I. G.; Etheridge, D. M.; Fielding, M. J., A perturbation analysis of the climate benefit from geosequestration of carbon dioxide. *International Journal of Greenhouse Gas Control* **2008,** *2*, (3), 289-296.
4. Benson, S.; Cook, P.; Anderson, J.; Bachu, S.; Nimir, H. B.; Basu, B.; Bradshaw, J.; Deguchi, G.; Gale, J.; Goerne, G. v.; Heidug, W.; Holloway, S.; Kamal, R.; Keith, D.; Lloyd, P.; Rocha, P.; Senior, B.; Thomson, J.; Torp, T.; Wildenborg, T.; Wilson, M.; Zarlenga, F.; Zhou, D.; Celia, M.; Gunter, B.; King, J. E.; Lindeberg, E.; Lombardi, S.; Oldenburg, C.; Pruess, K.; Rigg, A.; Stevens, S.; Wilson, E.; Whittaker, S., Underground Geological Storage. In *IPCC Special Report on Carbon Dioxide Capture and Storage*, Intergovernmental Panel on Climate Change: 2005.
5. Shankar Rao, K., Source estimation methods for atmospheric dispersion. *Atmospheric Environment* **2007,** *41*, (33), 6964-6973.
6. Yee, E.; Flesch, T. K., Inference of emission rates from multiple sources using Bayesian probability theory. *Journal of Environmental Monitoring* **2010,** *12*, (3), 622-634.
7. Keats, A.; Yee, E.; Lien, F.-S., Bayesian inference for source determination with applications to a complex urban environment. *Atmospheric Environment* **2007,** *41*, (3), 465-479.
8. Yee, E., Theory for Reconstruction of an Unknown Number of Contaminant Sources using Probabilistic Inference. *Boundary-Layer Meteorology* **2008,** *127*, (3), 359-394.
9. Yee, E.; Lien, F.-S.; Keats, A.; D'Amours, R., Bayesian inversion of concentration data: Source reconstruction in the adjoint representation of atmospheric diffusion. *Journal of Wind Engineering and Industrial Aerodynamics* **2008,** *96*, (10-11), 1805-1816.
10. Yee, E., Probability Theory as Logic: Data Assimilation for Multiple Source Reconstruction. *Pure and Applied Geophysics* **2011**, 1-19.
11. Jenkins, C.; Leuning, R.; Loh, Z., Atmospheric tomography to locate CO2 leakage at storage sites. *Energy Procedia* **2011,** *4*, 3502-3509.
12. Drescher, A. C.; Gadgil, A. J.; Price, P. N.; Nazaroff, W. W., Novel approach for tomographic reconstruction of gas concentration distributions in air: Use of smooth basis functions and simulated annealing. *Atmospheric Environment* **1996,** *30*, (6), 929-940.
13. Yee, E., Bayesian probabilistic approach for inverse source determination from limited and noisy chemical or biological sensor concentration measurements. *Proc. SPIE* **2007,** *6554*, (1), 65540W.
14. Leuning, R.; Etheridge, D.; Luhar, A.; Dunse, B., Atmospheric monitoring and verification technologies for CO_2 geosequestration. *International Journal of Greenhouse Gas Control* **2008,** *2*, (3), 401-414.

15. Loh, Z.; Leuning, R.; Zegelin, S.; Etheridge, D.; Bai, M.; Naylor, T.; Griffith, D., Testing Lagrangian atmospheric dispersion modelling to monitor CO_2 and CH_4 leakage from geosequestration. *Atmospheric Environment* **2009,** *43*, (16), 2602-2611.
16. Griffith, D.; Deutscher, N.; Krummel, P.; Fraser, P.; van der Schoot, M.; Allison, C., The UoW FTIR Trace Gas Analyser: comparison with LOFLO, AGAGE and tank measurements at Cape Grim and GASLAB. *Baseline Atmospheric Program (Australia) 2007-2008* **2011**.
17. Stull, R. B., *An Introduction to Boundary Layer Meteorology*. Springer: 1988; Vol. 13.
18. Thomson, D. J., Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *Journal of Fluid Mechanics* **1987,** *180*, 529-556.
19. Flesch, T. K. W., J. D.; Yee, E., Backward-Time Lagrangian Stochastic Dispersion Models and Their Application to Estimate Gaseous Emissions. *Journal of Applied Meteorology* **1995,** *34*, 12.
20. Flesch, T. K.; Wilson, J. D.; Harper, L. A.; Crenna, B. P.; Sharpe, R. R., Deducing Ground-to-Air Emissions from Observed Trace Gas Concentrations: A Field Trial. *Journal of Applied Meteorology* **2004,** *43*, 16.
21. Hsieh, C.-I.; Katul, G.; Chi, T.-w., An approximate analytical model for footprint estimation of scalar fluxes in thermally stratified atmospheric flows. *Advances in Water Resources* **2000,** *23*, (7), 765-772.
22. Bevington, P. R., *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill Book Company: 1969.
23. Keats, A.; Yee, E.; Lien, F. S., Information-driven receptor placement for contaminant source determination. *Environmental Modelling & Software* **2010,** *25*, (9), 1000-1013.
24. Jeffreys, H., An Invariant Form for the Prior Probability in Estimation Problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **1946,** *186*, (1007), 8.
25. Etheridge, D.; Luhar, A.; Loh, Z.; Leuning, R.; Spencer, D.; Steele, P.; Zegelin, S.; Allison, C.; Krummel, P.; Leist, M.; van der Schoot, M., Atmospheric monitoring of the CO2CRC Otway Project and lessons for large scale CO2 storage projects. *Energy Procedia* **2011,** *4*, 3666-3675.
26. Keeling, R. F.; Manning, A. C.; Dubey, M. K., The atmospheric signature of carbon capture and storage. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2011,** *369*, (1943), 2113-2132.
27. Picard, D., Fugitive Emissions from Oil and Natural Gas Activities. *IPCC: Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories* **2000**.