



Atmospheric CO₂ observations and models suggest strong carbon uptake by forests in New Zealand

Kay Steinkamp¹, Sara E. Mikaloff Fletcher¹, Gordon Brailsford¹, Dan Smale¹, Stuart Moore¹, Elizabeth D. Keller², W. Troy Baisden², Hitoshi Mukai³, and Britton B. Stephens⁴

¹National Institute of Water and Atmospheric Research, Wellington, New Zealand

²GNS Science, Lower Hutt, New Zealand

³National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan

⁴National Center for Atmospheric Research, Boulder, Colorado, USA

Correspondence to: Kay Steinkamp (kay.steinkamp@gmail.com)

Received: 22 March 2016 – Published in Atmos. Chem. Phys. Discuss.: 13 April 2016

Revised: 5 December 2016 – Accepted: 8 December 2016 – Published: 2 January 2017

Abstract. A regional atmospheric inversion method has been developed to determine the spatial and temporal distribution of CO₂ sinks and sources across New Zealand for 2011–2013. This approach infers net air–sea and air–land CO₂ fluxes from measurement records, using back-trajectory simulations from the Numerical Atmospheric dispersion Modelling Environment (NAME) Lagrangian dispersion model, driven by meteorology from the New Zealand Limited Area Model (NZLAM) weather prediction model. The inversion uses in situ measurements from two fixed sites, Baring Head on the southern tip of New Zealand's North Island (41.408° S, 174.871° E) and Lauder from the central South Island (45.038° S, 169.684° E), and ship board data from monthly cruises between Japan, New Zealand, and Australia. A range of scenarios is used to assess the sensitivity of the inversion method to underlying assumptions and to ensure robustness of the results. The results indicate a strong seasonal cycle in terrestrial land fluxes from the South Island of New Zealand, especially in western regions covered by indigenous forest, suggesting higher photosynthetic and respiratory activity than is evident in the current a priori land process model. On the annual scale, the terrestrial biosphere in New Zealand is estimated to be a net CO₂ sink, removing 98 (±37) Tg CO₂ yr^{−1} from the atmosphere on average during 2011–2013. This sink is much larger than the reported 27 Tg CO₂ yr^{−1} from the national inventory for the same time period. The difference can be partially reconciled when factors related to forest and agricultural management and exports, fossil fuel emission estimates, hydrologic fluxes,

and soil carbon change are considered, but some differences are likely to remain. Baseline uncertainty, model transport uncertainty, and limited sensitivity to the northern half of the North Island are the main contributors to flux uncertainty.

1 Introduction

The exchange of carbon between the atmosphere and the Earth's oceans and terrestrial biospheres plays a crucial role in climate projections (IPCC, 2013; Friedlingstein et al., 2014). Predicting the future trajectories of atmospheric CO₂, temperature, and precipitation requires a solid understanding of the magnitude of fluxes, their regional distribution, and how and why they vary on seasonal to decadal timescales. National greenhouse gas budgets are especially important in light of current policies regarding climate change, such as the annual reporting requirements under the United Nations Framework Convention on Climate Change (UNFCCC).

New Zealand's National Inventory Report (NIR) is compiled by the Ministry for the Environment (MfE), and published annually. For 2013, the NIR puts New Zealand's total CO₂ emissions at about 35 Tg CO₂ yr^{−1} and it is estimated that the land-use and forestry sector acted as a sink for carbon by removing three-quarters of that (27 Tg CO₂ yr^{−1}) from the atmosphere (MfE, 2015). These estimates are based on measurements from a network of forest plots throughout New Zealand, regularly updated land-use maps, and models. Inventory methods give precise estimates of carbon uptake

and release of the locally present vegetation type, but they are often difficult to scale up to regional or country scales due to heterogeneous biome composition (Ciais et al., 2010). Independent methods are needed to verify the reported carbon sink.

On very large, i.e. continental or global, scales both prognostic land models and inverse atmospheric models have been used (Gurney et al., 2004; Mikaloff Fletcher et al., 2007; Gruber et al., 2009; Le Quéré et al., 2013; Steinkamp and Gruber, 2013; Friedlingstein et al., 2014). Global inversions are a valuable top-down tool to estimate large-scale sinks and sources by combining observations from a global network with atmospheric circulation models. An inverse model interprets the observations to yield an optimized carbon flux distribution that is most consistent with the atmospheric data. In this approach, atmospheric model simulations relate fluxes at the surface to concentration changes at the observation sites. However, the number of available observation sites and the model resolution are usually insufficient to constrain the exchange of trace gases like CO₂ on smaller, i.e. regional to country, scales.

To address those scales, regional atmospheric greenhouse gas inversions (Lin et al., 2003; Stohl et al., 2009; Bergamaschi et al., 2010; Manning et al., 2011), including CO₂ inversions (Gerbig et al., 2003; Matross et al., 2006; Lauvaux et al., 2008), have been developed and used to estimate the carbon budgets of regions like Europe and the USA as well as individual nations. They are complementary to bottom-up inventories and provide a means to verify national inventories. Like their global counterparts, regional inversion methods combine CO₂ observations from surface sites with modelled atmospheric circulation to derive the distribution of sinks and sources over an area of Earth's surface (the inversion domain). In a regional inversion, however, the sites are not distributed globally and the domain is typically the size of a country or continent, which poses some additional challenges (Manning, 2011). An accurate model of background concentrations or sinks and sources from outside the inversion domain is required, i.e. a baseline. There is also a need for adequate spatial and temporal resolution for both the estimated fluxes and the circulation model to account for topographic effects and local emission gradients and hotspots. Many regional inversions use continuous in situ observations from one to a few measurement stations to sample the whole domain over the course of days to a few weeks. They use air from a background sector to construct a baseline time series (Manning et al., 2011; Uglietti et al., 2011). Due to chemical inertness of CO₂, atmospheric loss processes can be neglected once the gas has entered the domain, making this approach viable as long as the measurement station is positioned so that background air can be observed for significant fractions of time. It is generally of advantage to use multiple stations, which are sensitive to a larger surface area and allow for a better interpretation of spatial gradients in atmospheric CO₂.

In their inversion study, Stohl et al. (2009) estimate emissions for three HFC and HCFC greenhouse gases on national to global scales for 2005–2007. Their approach uses the FLEXPART Lagrangian model to describe the recent air history arriving at nine observation stations distributed globally. They use a priori emission maps and estimate both the baseline and the regional emissions as part of the inverse modelling. Manning et al. (2011) use 20 years of in situ CH₄ and N₂O observations from a single station, Mace Head, on the west coast of Ireland. Mace Head regularly receives air from the mid-latitude North Atlantic as well as from the UK and continental Europe, which allows them to estimate both the baseline and terrestrial emissions. Their emission estimates for the UK have been used to complement those reported to the UNFCCC for the period 1990–2007. Matross et al. (2006) derive regional-scale CO₂ flux estimates for summer 2004 in the north-eastern United States and southern Québec using the STILT Lagrangian model in conjunction with aircraft and tower observations.

Here, we present the first regional inversion for New Zealand, which leverages the country's unique characteristics. New Zealand is an isolated country surrounded by approximately 2000 km of ocean on all sides. This simplifies the construction of an accurate baseline model, as CO₂ signals from other land masses, especially from Australia and the Northern Hemisphere, will be significantly diluted and become part of the baseline before reaching the country. The expected national carbon sink, which is estimated by the inversion, is a large fraction of the fossil fuel emissions (about three-quarters according to the NIR) which are prescribed. In addition, New Zealand has multiple atmospheric CO₂ measurement sites across a relatively small country.

Our inverse model utilizes in situ observations from two observing stations in New Zealand for 2011–2013 and ship board measurements from a regular transect between Australia, New Zealand, and Japan (conducted by Japan's National Institute for Environmental Studies, NIES). The Numerical Atmospheric dispersion Modelling Environment (NAME III) Lagrangian model (Jones et al., 2007) was combined with meteorological output from the New Zealand Limited Area Model (Davies et al., 2005) at ~ 12 km resolution (NZLAM-12), to model the pathway of air arriving at the stations. Land model simulations from Biome-BGC (Thornton et al., 2005) are used as a priori estimates. We compare our results to New Zealand's NIR (MfE, 2015), point out differences and implications, and discuss the regional distribution and seasonal cycle of CO₂ sinks and sources across the country.

2 Observations

We use in situ measurements of CO₂ from two intercalibrated stations in New Zealand (Fig. 1). Baring Head (BHD) is located on the south coast of the North Island, while

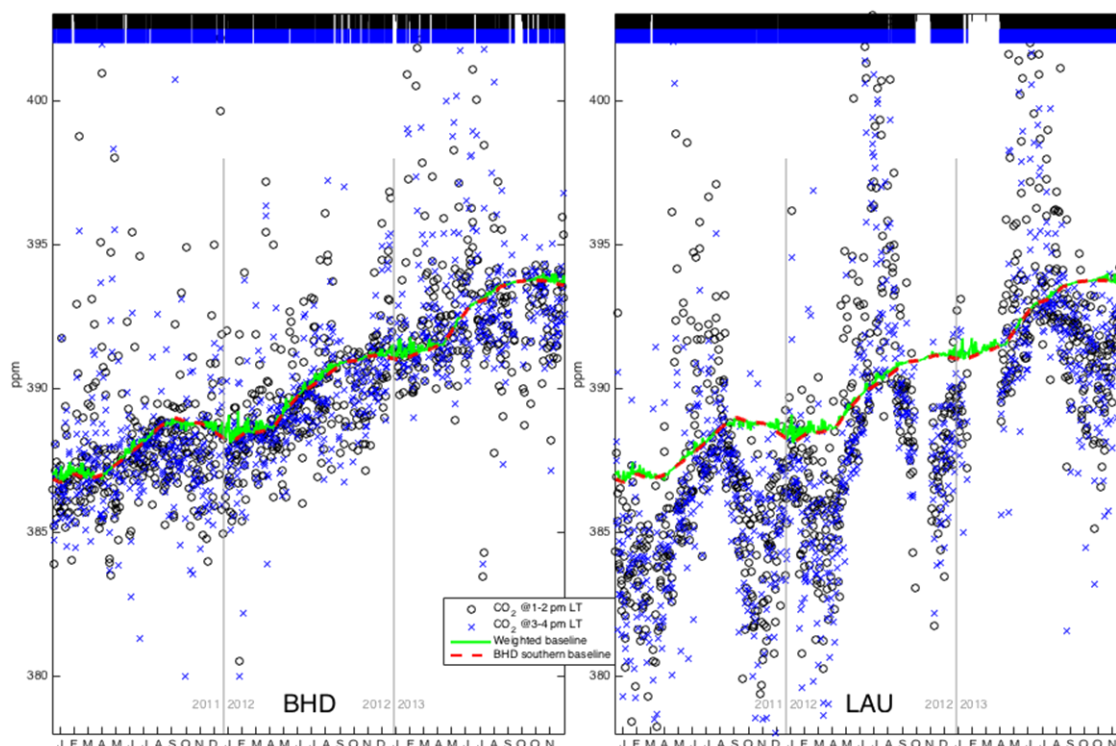


Figure 1. CO₂ observations from BHD and LAU, twice-daily at 13:00–14:00 and 15:00–16:00 LT, through 2011–2013. The BHD southern baseline is shown along with the weighted baseline used in the inversion. Gaps in the coloured bars at the top indicate days when no observations are available.

Lauder (LAU) is located in the central South Island (Fig. 2). Both stations are sensitive to different source regions of CO₂ and complement each other to allow for a comprehensive regional coverage spanning the Southern Ocean, Tasman Sea, the South Island, and – to a lesser degree – the North Island and the subtropical South Pacific.

The instruments used at LAU and BHD are operated with reference gases traceable to the World Meteorological Organisations mole fraction scale as maintained by the Central Calibration Laboratory (CCL) at the US National Oceanic and Atmospheric Administration (NOAA). Both instruments share common data processing code, improving data inter-comparability between the sites (Brailsford et al., 2012). In addition, fine-scale instrumental biases were assessed by using a suite of four transfer standard tanks with trace gas concentrations defined by the CCL. These instrument-specific multiplicative scalings were applied to the processed hourly data before the inversion. For typical ambient mole fractions of CO₂ (i.e. 380–410 ppm) these adjustments were generally less than 0.07 and 0.1 ppm for BHD and LAU, respectively.

Some of the elements of this study are prepared to include data from a third station, Maunga Kākaramēa/Rainbow Mountain, located in the northern half of the North Island (Fig. 2). For example, the region definitions in Sect. 5.2 include a local Maunga Kākaramēa/Rainbow Mountain region.

Because the calibration and quality control processes are equivalent to BHD and LAU, the station can be integrated seamlessly into the network. It is ideally located to extend the sensitivity of the inverse model into the north. However, it is not incorporated in this study because continuous CO₂ measurements from Maunga Kākaramēa/Rainbow Mountain were not yet available for the full 2011–2013 period.

For this study the hourly mean CO₂ records from BHD and LAU covering 1 January 2011 to 31 December 2013 are used. Both stations measure in situ with near-continuous observations throughout the day. The observations are strongly influenced by local signals at night and under certain meteorological conditions, e.g. a shallow planetary boundary layer (PBL) or low wind speed (Stephens et al., 2013). Measurements at times with deep, well-mixed PBL are better suited for inversion modelling as they are sensitive to sinks and sources from a wider region and not subject to localized processes or complex atmospheric structure. Similarly, because the vertical resolution of the meteorological model NZLAM (Sect. 3) is not fine enough to resolve the exact height of a station inlet, conditions with a well-mixed PBL are preferable (see also Sect. 7.4). Afternoon observations are, on average, well suited. Analysis of the diurnal cycle of CO₂ concentrations at the various inlet heights of both BHD and LAU shows least variability with altitude in the 13:00–

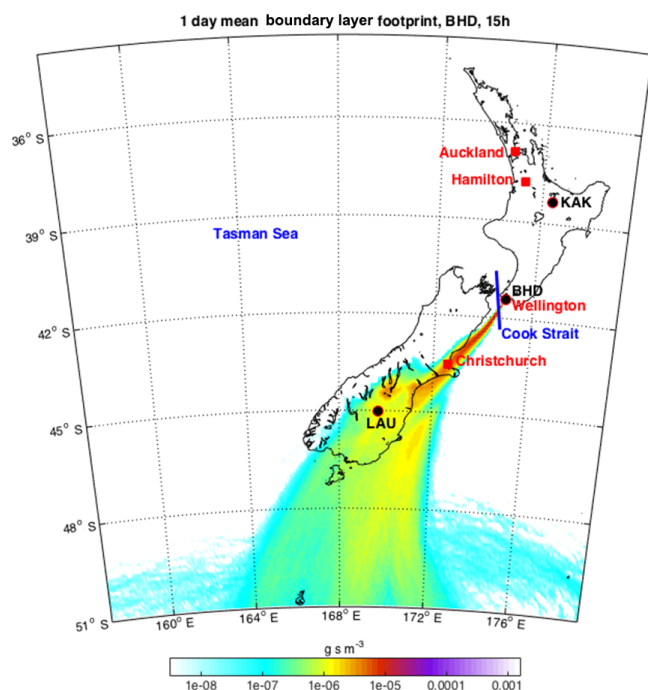


Figure 2. Air history map for 10000 particles released at BHD during 15:00–16:00 LT on 19 May 2012 using the NAME III model. The particle back trajectories show a southerly event locally at the station, though not a baseline event as the air crosses parts of South Island. Also shown are the locations of LAU and Kākaramea/Rainbow Mountain (KAK) stations.

16:00 afternoon hours. For the inversion, two hourly average data points per day are selected in the afternoon, 13:00–14:00 local time (LT) and 15:00–16:00 LT. Local time represents NZST (NZST = UTC + 12 h) in winter and NZDT (NZDT = UTC + 13 h) in summer. One potential risk of using only afternoon data is the potential for misrepresentation of fluxes with strong diurnal variations. Potential biases introduced by our sampling strategy are investigated and quantified in Sect. 7.4.

For both stations, 1 standard deviation of a 5 min measurement interval is taken as random data uncertainty for the hourly mean. This uncertainty is generally much greater than the measurement precision, as it reflects real atmospheric variability, and is instead intended to capture representativeness errors such as different temporal resolutions of model and data or the model failing to represent the specific conditions at an individual location.

2.1 Baring Head

BHD station (Lowe et al., 1979) is located at 41.408° S, 174.871° E, 85 m a.m.s.l., approximately 10 km south-east of the Wellington urban area (Fig. 2), close to the edge of a south-facing coastal cliff. The surrounding land is sparsely populated and has primarily been used for low density live-

stock farming. Wind speeds at the site regularly exceed 10 m s^{-1} , reducing the influence of local sources. The wind directions are primarily bimodal with the dominant wind from the north and the secondary direction from the south (Stephens et al., 2013). Southerly air arriving at the site has often been travelling over the Southern Ocean for at least 4 days, sometimes weeks, without any contact to terrestrial sinks or sources of CO₂. The station is ideally situated to determine baseline levels of atmospheric CO₂ at latitudes up to 70° S during these conditions. At other times, BHD measures air that has recently travelled over Australia or New Zealand, carrying a terrestrial signal of CO₂ sinks and sources.

We use hourly averaged CO₂ data from the non-dispersive infrared (NDIR) analyser (Ultramat 3, Siemens) in situ observations during the 13:00–14:00 and 15:00–16:00 time windows (Fig. 1) from a 10 m air inlet height. For more details on measurements, calibration, and data processing, we refer to Brailsford et al. (2012); Stephens et al. (2013).

2.2 Lauder

LAU station is located at 45.038° S, 169.684° E, 370 m a.m.s.l., in a broad river valley on the South Island, approximately 35 km north of the township of Alexandra (population 5000). The surrounding land is sparsely populated and largely used for low density livestock farming and seasonal cropping. The local wind direction is predominantly ranging from north-westerly to south-westerly. To the west lies a valley system in mountainous terrain, behind which the north–south running mountain range of the Southern Alps divides the island into a western coastal strip with a humid maritime climate and the eastern part with a more continental climate and relatively clear unclouded skies.

At LAU, CO₂ has been measured in situ using a dual cell NDIR analyser (LI-7000, LI-COR Inc.) since 2008 from a 10 m air inlet height. Unlike the BHD measurements, where detailed descriptions and analyses are given by Brailsford et al. (2012) and Stephens et al. (2013), the LAU measurements have not been published before. Therefore, we provide an extended description of the LAU in situ CO₂ measurement system in Appendix A.

3 Model simulations

We use NAME III (Jones et al., 2007), a Lagrangian dispersion model developed by the UK Met Office driven by three-dimensional meteorological fields precomputed by a numerical weather prediction (NWP) model. While initially developed more than 2 decades ago as an emergency response particle tool for nuclear outfall, NAME has since evolved into a general purpose dispersion model that is being used from local scales (a few hundred metres) to mesoscales and global scales. Atmospheric turbulence is simulated using a random

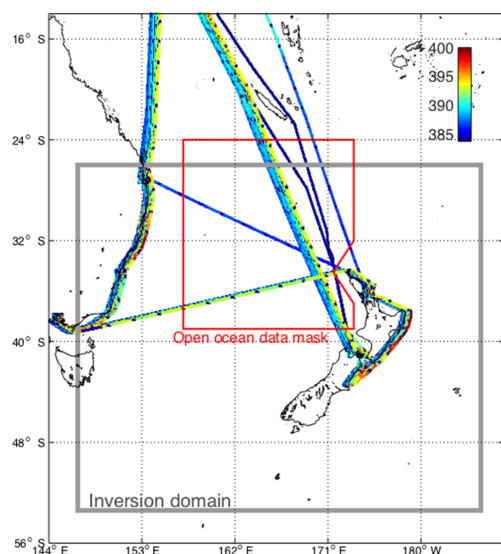


Figure 3. Voyages of the Ship of Opportunity *Trans Future 5*. Ship tracks have been slightly spread longitudinally to allow one to differentiate individual cruises (more recent cruises are to the right). Colours give the in situ CO₂ concentration. Measurements made inside the open-ocean mask were used in the northern baseline analysis. The inversion domain is outlined in grey.

walk technique (Morrison and Webster, 2005). CO₂ is modelled as an inert gas due to its long lifetime in the atmosphere (> 300 years), far exceeding the 4-day periods used for the back trajectories.

In this work, meteorology from the NZLAM-12 NWP model was used to drive NAME. NZLAM-12 is a local configuration of the UK Met Office Unified Model (Davies et al., 2005) with a horizontal resolution of ~ 12 km and 70 vertical levels up to a ceiling height of 80 km. The meteorology is a sequence of short, 6 h forecasts with hourly output that are produced from successive NZLAM simulations and cover the period 2011–2013. At the beginning of each simulation the available meteorological observations are assimilated into the model to match the state of the NZLAM atmosphere to the measured atmosphere. NAME uses the boundary layer depth (BLD) from NZLAM and applies a minimum and maximum BLD of 50 and 4000 m, respectively. The maximum height in NAME is 30 km, corresponding to the first 59 levels of NZLAM. Both NZLAM and NAME cover a domain ranging from 146.8 to 185.8° E in longitude and from 53.4 to 26.0° S in latitude (inversion domain is shown in Fig. 3).

3.1 Air history and station footprints

The NAME model is run in backward mode to analyse the history of the air travelling towards BHD and LAU over the preceding 4 days. Model particles are released from both stations during a period of 1 h, twice per day in 2011–2013,

at 13:00–14:00 and 15:00–16:00 LT. A simulation period of 4 days was found sufficiently long to allow all particles to leave the domain during most meteorological conditions, except during extended periods of very low wind speed. An air history map has been calculated for each release (Fig. 2). We use model output that represents the 4-day integrated air concentration (also called dosage, unit g s m^{-3}) inside each grid box on a regular $0.1^\circ \times 0.1^\circ$ grid, designed to be very similar to the ~ 12 km grid of NZLAM-12. During each release, the dispersion of 10 000 particles is modelled and every particle registered within the boundary layer at a given time contributes to the dosage of the respective grid cell. Particles are simulated in three dimensions and do not disappear when leaving the boundary layer as long as they remain below the maximum model height of 30 km. Particles can leave the boundary layer temporarily and descend back into it at a later time, in which case they would again contribute to the dosage. An example of this can be seen in Fig. 2, where many particles leave the boundary layer just south of the South Island and later (from the point of view of the backward simulation) descend again, visible as a weaker dosage (less strong colours) for some stretch of the map.

Average footprints for the BHD and LAU stations were computed by summing the dosage maps for each day and release period in 2011–2013 and normalizing them such that the domain integral equals one (Fig. 4). These footprints represent the average sensitivity of a station to surface fluxes (sinks and sources) of CO₂ at a specific location. They have also been used to inform the partitioning of the inversion domain into a set of regions for which weekly surface fluxes are calculated (Sect. 5.2).

3.2 Transport matrix

For particle transport, the mass flux during each 1 h release period is $1 \text{ g CO}_2 \text{ s}^{-1}$, amounting to a total emission of 3600 g CO_2 over the period (0.36 g CO_2 is assigned to each particle). The flux strength is an arbitrary choice and does not affect the transport results due to the implied linearity of transport. A transport matrix \mathbf{T}_g (unit s m^{-1}) is formed by dividing the dosage by the total emitted mass and multiplying by the area (m^2) of each surface grid cell. Thus, each element of \mathbf{T}_g describes the atmospheric transport of a continuous emission of $1 \text{ g CO}_2 \text{ m}^{-2} \text{ s}^{-1}$ from a given grid cell over the previous 4 days and subsequent contribution to the air concentration at the receptor (BHD or LAU) during each 1 h period. With \mathbf{x}_g being a vector containing all grid cells and \mathbf{c} a vector containing the concentration (unit $\text{g CO}_2 \text{ m}^{-3}$) for all 1 h periods, this is written as $\mathbf{T}_g \mathbf{x}_g = \mathbf{c}$. Given \mathbf{T}_g and the measured concentrations \mathbf{c} , the aim is to solve for the CO₂ fluxes \mathbf{x}_g using a Bayesian inversion, i.e. a statistical model that balances information from measurements with a priori knowledge about the fluxes (Sect. 6).

Instead of solving on the grid scale, the fluxes in \mathbf{x}_g are pre-aggregated into a set of regional fluxes (Sect. 5.2), \mathbf{x} , and

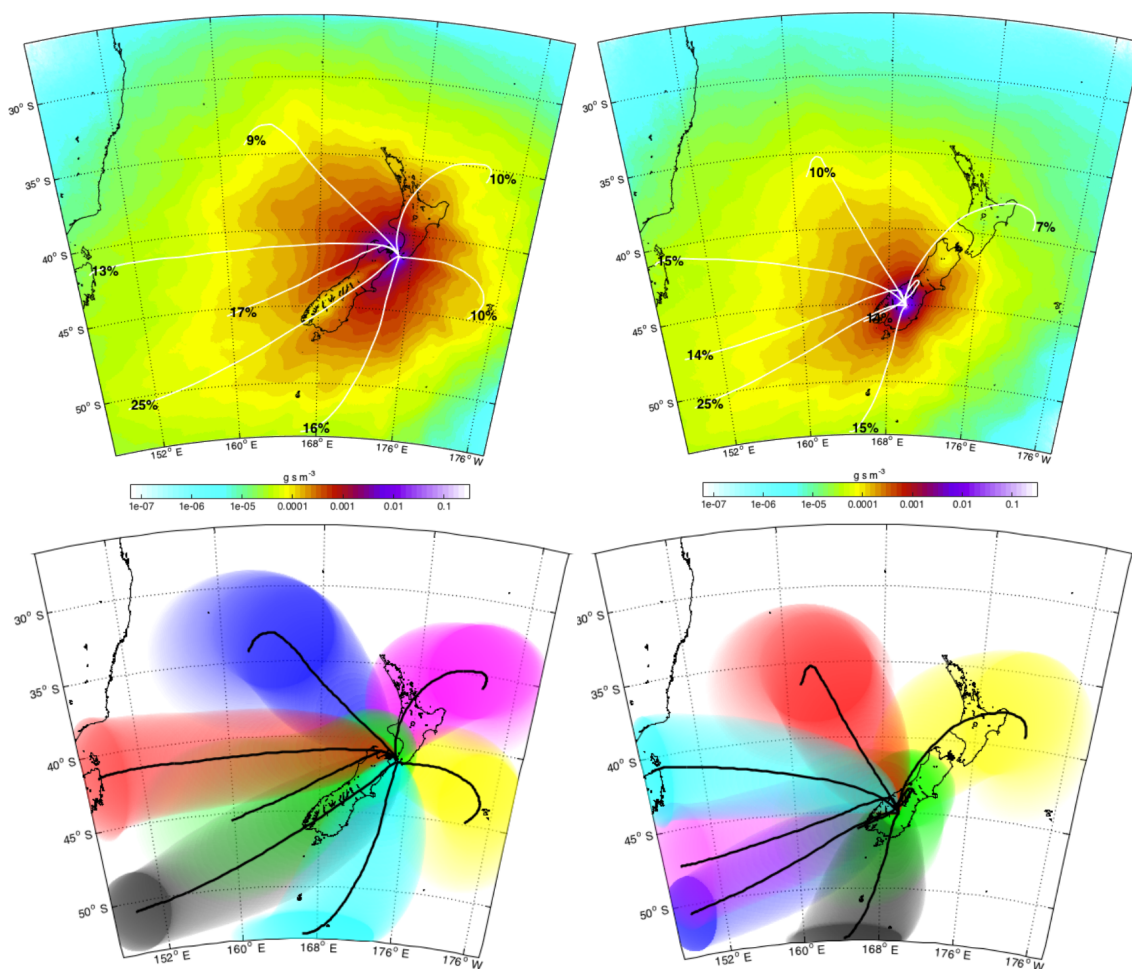


Figure 4. Top panel: 2011–2013 mean footprints for BHD and LAU stations, based on twice-daily air history maps at 13:00–14:00 and 15:00–16:00 LT. Clusters of 4-day back trajectories are overlain. Percentages give the sizes of clusters, i.e. the probability that a particle released on a random day has followed that pathway. Lower panel: major atmospheric transport pathways for both stations from cluster analysis of the back trajectories from twice-daily particle release. Shades represent the geographical spread of each pathway (1 standard deviation from cluster centroid in latitude–longitude).

a transport matrix \mathbf{T} created by aggregating grid cells in \mathbf{T}_g to reflect the regions in \mathbf{x} :

$$\mathbf{T}\mathbf{x} = \mathbf{c}. \quad (1)$$

The Bayesian inversion developed here solves for \mathbf{x} . In addition, a priori flux maps are taken into account for the terrestrial and oceanic portions of the domain (Sect. 4).

4 A priori CO₂ flux maps

The Bayesian approach in this study uses spatially distributed information about CO₂ sinks and sources as first-guess, or a priori, fluxes for terrestrial and oceanic regions. These fluxes are optimized by the inversion using the constraints imposed by the CO₂ measurements at the stations (Sect. 6). Fossil fuel and cement emissions are accounted for as well, though unlike the natural fluxes they are prescribed and not optimized

by the inversion. Here we describe the datasets and flux maps used, while their incorporation in the inversion is described in Sect. 5.

4.1 Terrestrial

First-guess land-to-air CO₂ fluxes from the biosphere for every month in 2011–2013 are obtained from the Biome-BGC model (Thornton et al., 2005). Biome-BGC (v4.2 final release) is an ecosystem process model that estimates the storage and flux of carbon, nitrogen, and water (Thornton et al., 2002). The model has been extensively tested and validated for North American and European ecosystems and in addition was recently extended and applied to New Zealand managed pasture systems (Keller et al., 2014). The adaptation of the Biome-BGC model to New Zealand by Keller et al. (2014) is used in this study to estimate net ecosystem pro-

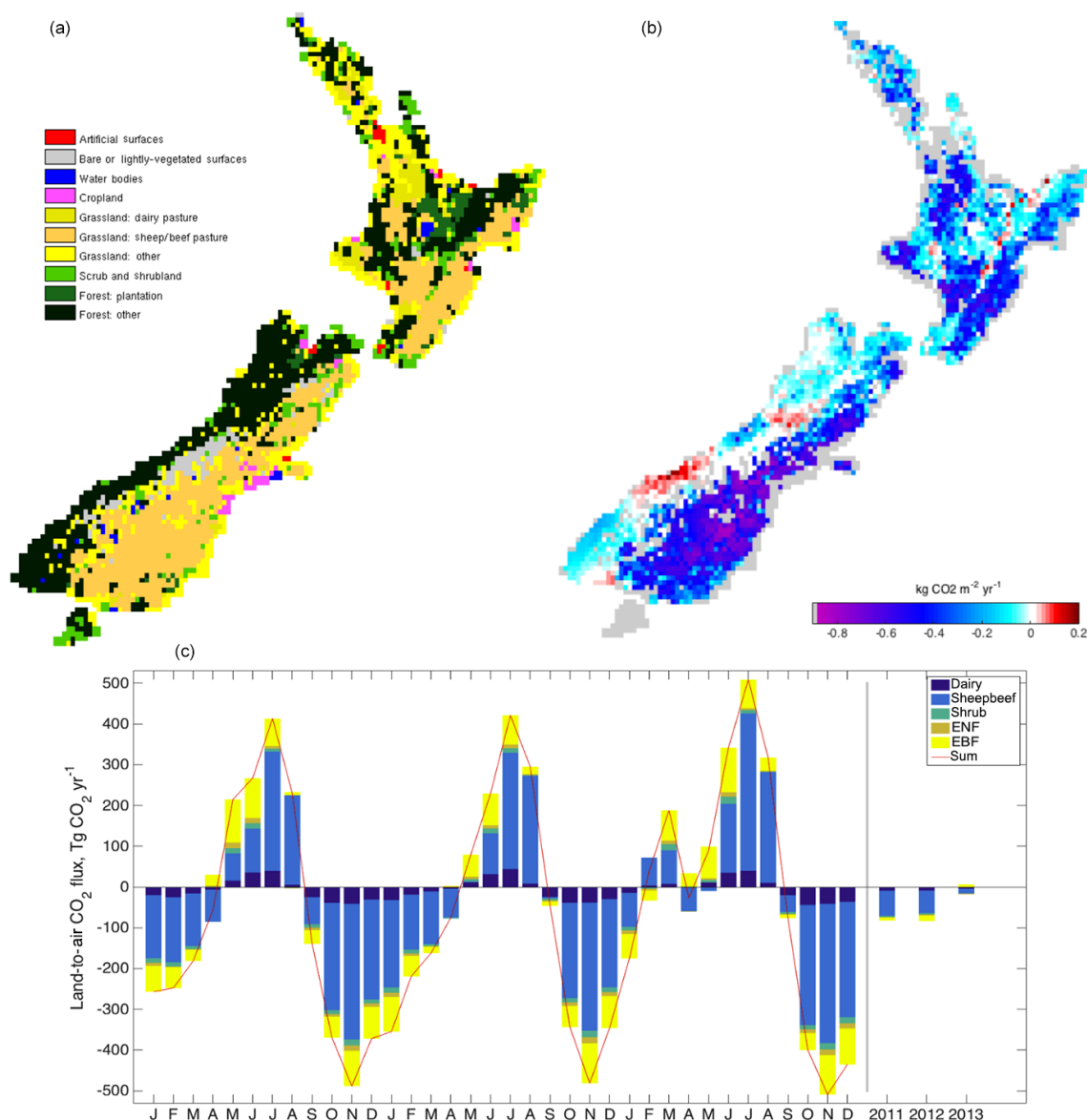


Figure 5. (a) Land-cover and land-use (LCLU) map with 10 categories, based on the Land Cover Database (LCDB3) and the Land Use in Rural New Zealand (LURNZ) basemap. (b) A priori CO₂ flux distribution, averaged over 2011–2013, from modelled NEP using Biome-BGC with the LCLU categories. Both maps have been regridded to the NAME model grid. (c) Monthly and annual contributions from each biome to the overall a priori CO₂ flux.

duction (NEP) for five biomes across New Zealand: dairy pasture, sheep and beef pasture, shrub, evergreen broadleaf forest (EBF), and evergreen needleleaf forest (ENF). The model is driven by daily weather data from the NIWA virtual climate station network (VCSN). VCSN data include numerous meteorological parameters on a regular (~ 5 km) grid covering the whole of New Zealand (Tait et al., 2006). The data are based on the spatial interpolation of actual data observations made at climate stations located around the country. Soil attributes are incorporated from the fundamental soil layers database (Landcare Research, 2015).

Biome-BGC produces NEP maps for each biome covering the whole country; i.e. it does not make assumptions about the actual distribution of biomes. In order to partition the country into biomes approximating the five categories available in Biome-BGC and then mask and sum the NEP contributions from each biome, we produced a land-cover and land-use (LCLU) map. The LCLU map uses 10 categories based on a combination of the land-cover database (LCDB) for New Zealand (Shepherd and Newsome, 2009; Dymond et al., 2012) and the Land-Use in Rural New Zealand (LURNZ) model (Hendy et al., 2007; Timar, 2011; Kerr et al., 2012).

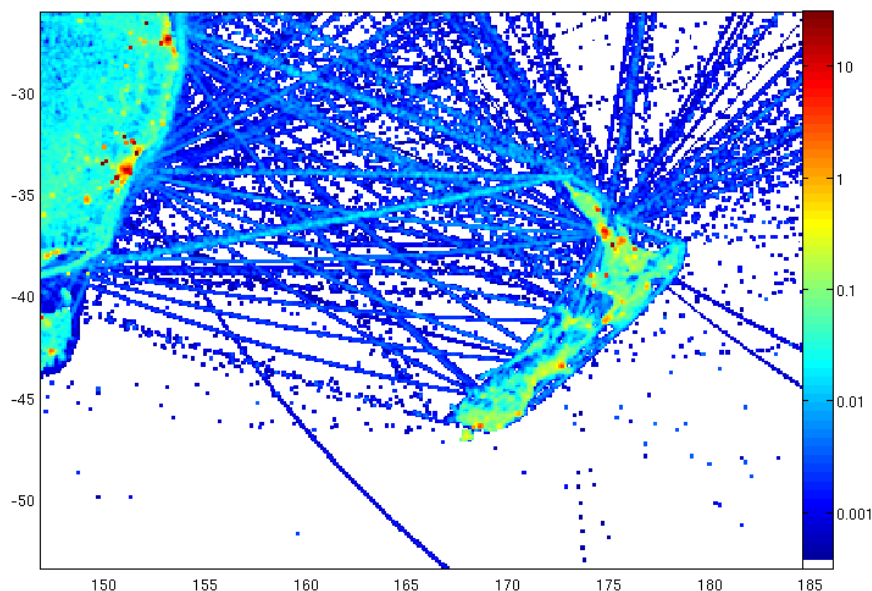


Figure 6. Gridded fossil emissions of CO₂ across the model domain and averaged over 2011–2013, in kg CO₂ m^{−2} yr^{−1}. Emissions are based on EDGAR v4.2, with 2011–2013 estimates extrapolated from the 2000–2010 trend in global total emissions.

The New Zealand LCDB is a thematic classification of land-cover and land-use categories, created using satellite imagery and covering all of mainland New Zealand. The dataset is polygon based and designed to be compatible in scale and accuracy with Land Information New Zealand's 1 : 50 000 topographic database. For the purpose of this study, the distribution of 33 land-cover types from LCDB version 3 for 2008/2009 was used and rasterized on a 5 km × 5 km grid.

LURNZ is a dynamic partial equilibrium model that simulates changes in private rural land use over time and space. It focuses on four key land uses – dairy, sheep and beef, forestry (plantations), and scrub and shrubland. While the model's primary focus is on simulating future changes in land use under scenario projections of commodity prices in one of the four sectors, it also provides a baseline of actual land use in 2008. This 2008 basemap is used in this study to match the Biome-BGC dairy and sheep and beef pasture biomes; however, LURNZ does not include native forests.

To account for all biomes in Biome-BGC, the LURNZ 2008 basemap and the LCDB 2008/2009 land-cover map are combined as follows. First the 33 LCDB categories are aggregated into 7 – forest, scrub and shrubland, grassland, cropland, water bodies, bare or lightly vegetated surfaces, and artificial surfaces. The forest and grassland categories are then sub-divided into plantations, “other forests”, dairy pasture, sheep and beef pasture, and “other grasslands” using LURNZ, which results in the LCLU map in Fig. 5a.

The ENF biome is assumed to be well represented by the plantation forest category, with plantations consisting primarily of pine trees, and the EBF biome is assumed to be better represented by the “other forests” category. The cat-

egories of artificial surfaces, bare and lightly vegetated surfaces, and water bodies are assigned a zero flux, i.e. no exchange of CO₂ with the atmosphere. No flux estimates are made for cropland and “other grasslands”; this does not affect results significantly, because these categories represent only a small portion of the total land area. Figure 5b shows the 2011–2013 mean a priori land-to-air CO₂ flux as estimated by matching the LCLU and Biome-BGC biomes in this manner and summing their contributions to the overall NEP. The monthly and annual contributions are shown in Fig. 5c. Weekly first-guess CO₂ flux maps are obtained by simple interpolation of the monthly estimates throughout 2011–2013.

An uncertainty estimate is computed for the a priori CO₂ flux from each grid cell. Based on Keller et al. (2014) and personal communication with the authors, we assign 10 % of the flux as uncertainty for pasture land. For forests, we assign 10 % everywhere except in the Canterbury and Otago regions in the South Island, where 56 and 36 % are used, respectively. These are conservative estimates based on a comparison of the Biome-BGC ENF modelled live stem carbon with the national exotic forest regional yield tables (MPI, 2012). The Canterbury and Otago regions were assigned larger uncertainties to reflect the larger discrepancy between the Biome-BGC model and the MPI yields in these regions. The uncertainty is taken into account by the Bayesian optimization (Sect. 5). We note that the EBF biome has not been calibrated or tested under New Zealand conditions and might contain additional uncertainty.

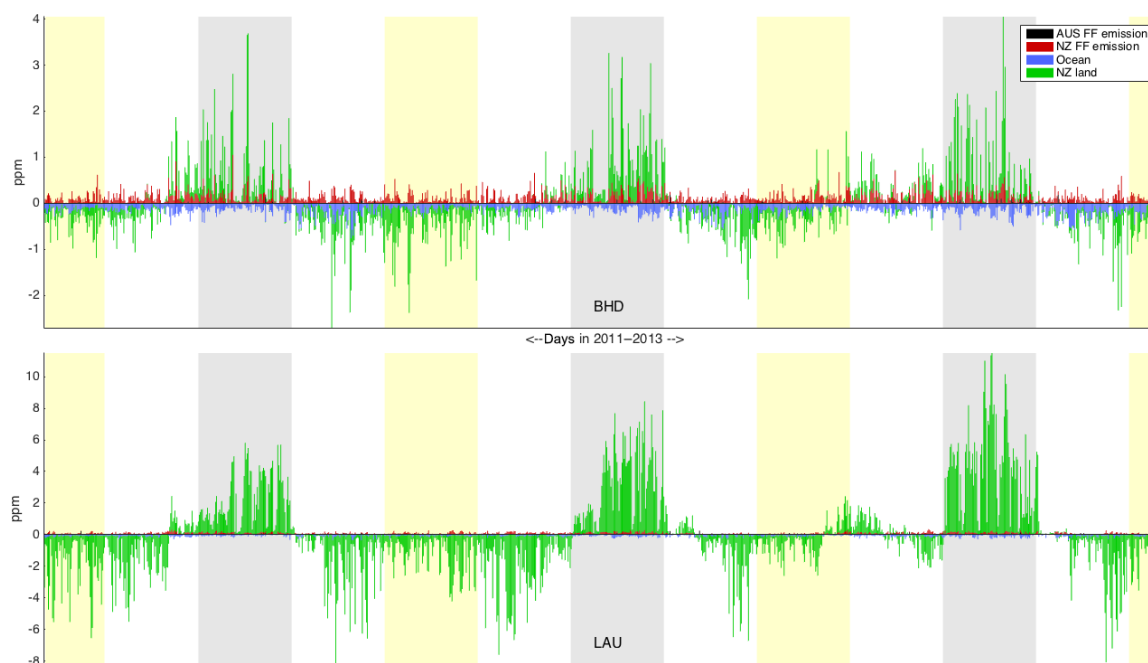


Figure 7. Simulated contributions to the observed CO₂ anomaly, i.e. concentration minus baseline, at BHD (top) and LAU (bottom) for each day in 2011–2013 averaged over both 13:00–14:00 and 15:00–16:00 LT release periods. Contributions are calculated using NAME transport matrices with EDGAR v4.2 fossil fuel emissions, prior oceanic CO₂ flux from the Takahashi *p*CO₂ dataset, and prior terrestrial CO₂ flux from the Biome-BGC model. Note that scales vary due to stronger anomalies and seasonal amplitude at LAU.

4.2 Oceanic

First-guess air–sea CO₂ fluxes are calculated based on a global dataset of surface ocean *p*CO₂ (Takahashi et al., 2009a). The dataset contains approximately 4.5 million measurements of surface water partial pressure of CO₂ (*p*CO₂) obtained over the global oceans during 1968–2008, approximately 90 000 of which were taken inside the model domain of this study. A monthly climatology on a global 4 × 5 grid was derived by Takahashi et al. (2009b), which also includes an estimate of air–sea CO₂ flux derived from the difference of surface ocean and atmospheric CO₂ and a gas exchange rate following Wanninkhof (1992).

An uncertainty estimate for the a priori ocean fluxes is computed as the root mean square of two components reflecting the uncertain gas exchange rate and the spatiotemporal coverage of measurements inside grid cells. For the first component, we recalculated the CO₂ flux using each of seven additional gas transfer models (Ho et al., 2006; Sweeney et al., 2007) and used 1 standard deviation from the eight-model mean as uncertainty. The second component applies an uncertainty to grid-scale fluxes inversely proportional to the number of measurements taken inside them for a given month of the climatology (Steinkamp and Gruber, 2013). As with the terrestrial CO₂ flux prior, the uncertainty is accounted for in the Bayesian inversion.

4.3 Fossil emissions

A gridded map of CO₂ emissions is derived from the Emission Database for Global Atmospheric Research (EDGAR) version 4.2 (EC-JRC/PBL, 2011). EDGAR contains global emission inventories for greenhouse gases and air pollutants from sectors including energy, industrial processes, solvents and other product use, agriculture, land-use change and forestry, and waste. Annual emissions are available on a 0.1° × 0.1° grid over the globe up to the year 2010. Emissions for the 2011–2013 time period were approximated by extrapolation using the trend in global total emissions over 2000–2010. The spatial distribution was assumed unchanged from 2010 (Fig. 6). Total emissions for the New Zealand mainland are 47.8 Tg CO₂ yr^{−1} in 2011–2013.

Fossil CO₂ emissions are not optimized by the inversion; instead, their contribution to the CO₂ signal at both stations (Fig. 7) is subtracted from the actual measurements beforehand. That contribution is calculated using the transport matrix from NAME, i.e. applying Eq. (1) with *x* containing the emissions from every grid cell and week in 2011–2013. The vector *c* then contains CO₂ concentrations for the twice-daily release periods at both stations that are caused by the emissions. To convert concentrations into mole fractions (ppm) the atmospheric pressure and temperature from the NAME model are used, which were interpolated to the BHD and LAU site coordinates from the NZLAM-12 temperature and pressure fields.

5 Regional flux estimation

5.1 Baseline analysis

Any regional CO₂ inversion can only estimate sinks and sources within the boundaries of the model domain. Sink and source processes from outside the domain become part of the CO₂ background concentrations (i.e. baseline) seen by the regional inversion at the boundary. Therefore, an accurate description of this baseline is needed. A common approach is the background-sector method (Manning et al., 2011; Uglietti et al., 2011), where air is classified as baseline if it originates from a certain wind sector and fulfils site-specific meteorological criteria. A continuous baseline is constructed using a gap-filling technique, which is subtracted from all other measurements before the inversion. The inverse model then interprets the remaining variability in the data, or anomalies, to find the optimal distribution of sinks and sources within the model domain.

The background-sector method has been applied to the BHD CO₂ record by Brailsford et al. (2012) and Stephens et al. (2013). They use steady background CO₂ mole fractions during southerly wind conditions at BHD and apply a multi-step filter to the BHD record to obtain a CO₂ baseline representative of a large region over the Southern Ocean. In short, the filter selects measurements during extended periods of southerly winds at the site, during which a maximum standard deviation of 0.1 ppm is achieved. Additional meteorological conditions must be fulfilled to preclude the influence of local sources and to ensure the air has not passed over the South Island before arriving at BHD. After filtering the data for baseline conditions, a continuous baseline is constructed using the seasonal time series decomposition by Loess (STL) algorithm (Cleveland et al., 1990), which can be sampled hourly, i.e. during the 13:00–14:00 and 15:00–16:00 LT release periods. The baseline derived from the BHD record is shown in red in Fig. 1 and will be called the southern baseline.

One disadvantage of a background-sector approach based on a single site is that it may not capture variability in background concentrations from different wind conditions, in particular along the latitudinal axis with its gradient in atmospheric CO₂, which could lead to errors in the flux estimates within the domain. In our case, BHD's background sector is ideally situated to obtain a CO₂ baseline representative of a large region over the Southern Ocean. However, observations made during northerly events are not always well described using this baseline, as the air often originates from the northern Tasman Sea or the subtropical South Pacific and carries a contribution from northern hemispheric CO₂ (Sect. 6, Fig. 4).

To alleviate this we augment the southern baseline with a second baseline from ship data representative of the northern sector. This northern baseline is based on in situ CO₂ observations using a NDIR analyser on board the *Trans Future 5* (TF5), a ship of opportunity that cruised the triangle

Japan–Australia–New Zealand about once a month during the period 2011–2013 (Chierici et al., 2006). We filter the ship track data (Fig. 3) to select open-ocean measurements and mask out observations taken close to the land, especially near the Australian east coast as it is located upwind during average south-westerly conditions and hosts large urban centres with significant CO₂ emissions. These selected data are then latitudinally averaged between 26 and 27° S to produce a baseline representative of the northern edge (26° S) of the inversion domain. A continuous baseline is constructed using the same STL routine as for the southern baseline.

For both baselines, uncertainty estimates are formed based on the monthly standard deviations of the in situ data as well as differences between measurements and the STL smoothed curve. A more detailed description of the construction of both baselines is provided in Appendix B. A caveat of this concept of baseline is that synoptic patterns of CO₂ can travel over large distances and cause variations in the background concentrations at the regional boundary, which are not accounted for by the baseline derived here. Given New Zealand's geographic isolation and location in the Southern Hemisphere, the strength of synoptic variations is limited compared to regions where inflow is characterized by strong variability due to terrestrial signals from other areas. While an explicit treatment of synoptic variability is beyond the scope of this study, a sensitivity experiment is described in Sect. 5.4 to address potential biases in the baseline.

A combined CO₂ baseline is constructed that takes into account where the modelled trajectories originated for any given data point. The daily NAME station footprints for the 13:00–14:00 and 15:00–16:00 LT windows are integrated along the southern and northern edges of the domain to determine the relative fraction of back trajectories leaving the domain to the south and north. These fractions are then used to weight the two baselines and create a baseline associated with each of the twice-daily data points. Uncertainties are weighted in the same way. The combined baseline is shown in green in Fig. 1. For the plot the 13:00–14:00 and 15:00–16:00 LT weighted baselines were averaged, as they are visually almost indistinguishable, but the individual baselines are used in the inversion.

5.2 Regional partitioning

CO₂ fluxes are estimated for every week in 2011–2013 and for 25 geographic regions distributed across the inversion domain (Fig. 8). The within-region pattern of the fluxes is prescribed using the a priori flux maps for New Zealand and the surrounding oceans, while the inversion estimates regional totals. The definition of the regional boundaries was guided by several factors, including the distance from the measurement stations, the gradient of the station footprint, local orography, and fossil emission hotspots. For land regions in New Zealand, additional factors include land-cover and land-use

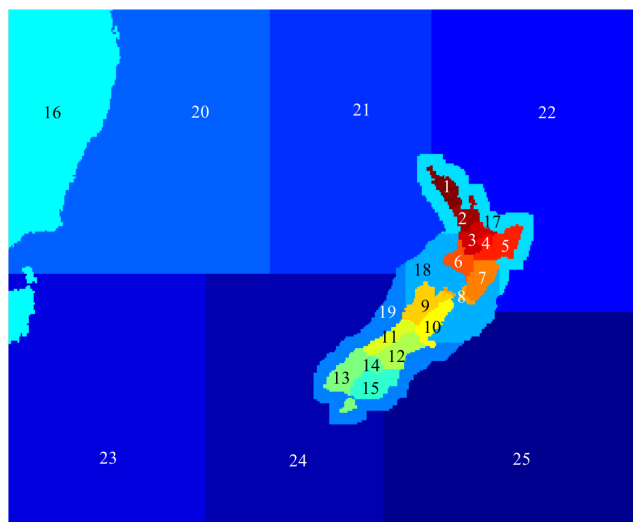


Figure 8. Regional partitioning and indices in the inversion.

types as well as the expected (first-guess) CO₂ flux distribution.

Due to its large distance from the stations, the portion of Australian land inside the inversion domain is represented by a single region (no. 16). No a priori information about natural CO₂ fluxes from Australia is assumed, i.e. they are set to zero with a very large uncertainty of 1000 Tg CO₂ yr⁻¹, so that the inversion is free to adjust them. This is based on an analysis suggesting generally low sensitivity of CO₂ measurements at our stations in New Zealand to fluxes from the Australian region (Sect. 6, Fig. 7). The analysis uses fossil emissions from the Australian region (Sect. 4.3) and investigates whether these emissions leave a significant imprint on measured CO₂ at BHD and LAU. Except for a few days this imprint is negligible (Sect. 6).

The portions of the Southern Ocean, South Pacific and Tasman Sea that are inside the model domain were divided into six open-ocean and three coastal regions (17–25). Large open-ocean regions are chosen to make their regionally integrated contribution to the CO₂ signal become discernible at the stations, although these contributions remain much smaller than those from the land regions in New Zealand. Three northern and three southern open-ocean regions allow for the difference in ocean biogeochemistry between Southern Ocean and subtropical Pacific waters, with guidance from patterns of surface ocean *p*CO₂ from the a priori map. The coastal ocean regions were included to separate the open ocean from the land explicitly and to account for their stronger influence on the measured CO₂ due to their relative proximity to the stations compared to the open ocean. The coastal ocean was defined as the combination of a 60 km coastal band around New Zealand and the portion of ocean with a mean 2011–2013 BHD footprint value above a fixed threshold. The threshold was chosen such that the integrated

CO₂ signal at BHD from coastal regions is 25 % of that from all ocean areas.

It is generally important to separate regions that exhibit a strong variability in sensitivity, as otherwise these within-region gradients can skew the regional totals estimated by the inversion towards the most sensitive areas inside the region. For land regions in New Zealand, we used the spatial gradients of the 2011–2013 footprints as an estimate for this variability, similar to the coastal regions, except that for the land, we use the combined footprint of BHD and LAU (Fig. 9) and also account for additional factors.

New Zealand was divided into 15 land regions (1–15) as follows. Three small regions around BHD, LAU and Maunga Kākaramea/Rainbow Mountain were defined, which have the largest contributions to the CO₂ signal at the respective stations. This separation of the highly influential local regions from the rest of the country follows the same rationale as the separation of the coastal from the open ocean, with the aim to prevent the inversion from allocating local signals to regions further upwind. A separate region around the urban areas of Auckland and Hamilton was defined to capture the strong fossil emissions there. The remaining regions were defined with the aim to minimize the footprint variability, the expected flux variability inside each region, and the number of land-cover and land-use types within each region, while accounting for topographic features.

The resulting regional partitioning is shown in Fig. 8. A major feature is the role of the Southern Alps as a dividing range between the humid west coast of the South Island containing large patches of native forest, and the dryer regions in the central and eastern parts, where pasture land is predominant. On the North Island, the axial mountain ranges divide the land into east and west as well, but the distribution of forests and pasture is more complex. BHD and LAU have relatively low sensitivity to the northern half of the North Island, which results in large uncertainties after the inversion for individual regions. However, regionally aggregated results are well constrained in that part of the country.

5.3 Inversion methodology

The aim of the inverse method is to estimate a net CO₂ flux from every region and for every week between 2011 and 2013 using a Bayesian approach (Gurney et al., 2004; Taranola, 2005; Steinkamp and Gruber, 2013). The approach assimilates information from the twice-daily observations from both stations (the “data”) and accounts for a priori fluxes (the “prior”) and contributions from fossil emissions.

The data time series is constructed by subtracting the baseline from the station measurements. The modelled CO₂ signal from fossil emissions is also subtracted. The resulting time series represents the part of the observed CO₂ signal that cannot be explained by background concentrations or fossil emissions and is therefore due to the net effect of sinks and sources of CO₂ over the ocean and land portions inside the

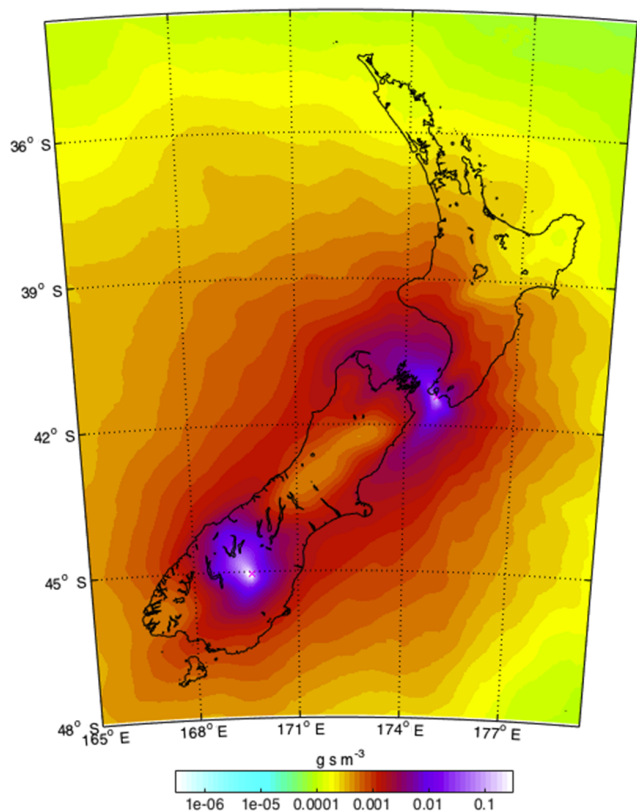


Figure 9. Combined 2011–2013 footprint for both 13:00–14:00 and 15:00–16:00 LT release periods for BHD and LAU around New Zealand.

model domain. The data for each station and each 1 h period are written as a vector \mathbf{d} .

The uncertainty applied to each data point is calculated as the quadrature sum of the baseline uncertainty (Sect. 5.1) and the CO₂ data uncertainty (Sect. 2). A minimum uncertainty of 0.4 ppm is assumed to account for uncertainties in the transport model as well as possible errors in the fossil fuel emission estimates. The final uncertainty is multiplied by 2.9, a value based on a goodness of fit analysis of the inverse model (reduced chi-squared statistic, as described below). The resulting uncertainty is taken as the root mean square (quadrature) of both components and has a minimum value of 1.16 ppm. The mean uncertainty is 1.91 ppm and 95 % of the values are within the 1.16 to 4.56 ppm range. The square of the uncertainty populates the main diagonal of the data covariance matrix \mathbf{C}_d . We assume no correlations between pairs of data points, so all off-diagonal elements of \mathbf{C}_d are set to zero.

The regional prior (denoted \mathbf{x}_0) is obtained by integrating the weekly a priori terrestrial and oceanic flux maps over each region except the Australian region. The prior uncertainty is similarly obtained by aggregating the grid-scale uncertainty estimates. Since the within-region flux patterns remain fixed, we assume full spatial correlation when propa-

gating grid-scale uncertainties to the regional scale. For land regions we added (via root mean square) an additional uncertainty component of 50 % of the seasonal flux amplitude. This allows the inversion to shift the seasonal cycle more freely; otherwise, the seasonal turning points – i.e. the switch between net CO₂ uptake in the summer months and net release in the winter – would essentially be fixed as the flux is near zero and the grid-scale uncertainty estimates for the Biome-BGC model are proportional to the flux strength. The diagonal prior covariance matrix \mathbf{C}_0 contains the regional uncertainty.

The regional prior is linked to the data vector as in Eq. (1), except \mathbf{x} now contains fluxes on the regional instead of grid scale, and the transport matrix \mathbf{T} links regional total fluxes to the data time series with baseline and fossil signal subtracted.

The inversion process seeks an optimal solution to the transport equation by balancing the data and prior constraints (Tarantola, 2005), i.e. by minimizing a Bayesian cost function J with respect to \mathbf{x} :

$$\begin{aligned}
 J &= \frac{1}{2}(\mathbf{T}\mathbf{x} - \mathbf{d})^T \mathbf{C}_d^{-1} (\mathbf{T}\mathbf{x} - \mathbf{d}) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{C}_0^{-1} (\mathbf{x} - \mathbf{x}_0) \\
 &+ \frac{1}{2}(\mathbf{S}\mathbf{x})^T \mathbf{C}_s^{-1} (\mathbf{S}\mathbf{x}) \\
 &= \frac{1}{2}(\tilde{\mathbf{T}}\mathbf{x} - \tilde{\mathbf{d}})^T \tilde{\mathbf{C}}_d^{-1} (\tilde{\mathbf{T}}\mathbf{x} - \tilde{\mathbf{d}}) \\
 &+ \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{C}_0^{-1} (\mathbf{x} - \mathbf{x}_0). \quad (2)
 \end{aligned}$$

The first term in the equation evaluates the deviation of the modelled time series from the data, with each data point weighted with the inverse uncertainty. The second term evaluates the deviation of the optimized regional fluxes from the prior fluxes. The last term is a Gaussian smoother, which limits changes in week-to-week fluxes. The operator \mathbf{S} forms a vector whose elements correspond to the difference of each flux in \mathbf{x} and the flux of the following week. The diagonal matrix \mathbf{C}_s contains the squares of values representing the strength of the smoother. We chose 5 kg CO₂ m⁻² yr⁻¹ for every grid cell, which is more than 10 times larger than the largest flux from any grid cell of the a priori flux maps (Fig. 5); hence the smoother is very weak. In fact the smoother was designed to have a negligible effect on estimated CO₂ fluxes and not interfere with the prior and data constraints. Its role is merely to favour solutions with small week-to-week changes in cases where a second solution with much larger week-to-week changes would result in a very similar cost, J . Note that the smoothing factors for individual regions differ slightly due to varying surface areas. Due to their similar mathematical forms, the smoothing term can be absorbed in the data term in Eq. (2) by appending \mathbf{S} to \mathbf{T} (forming $\tilde{\mathbf{T}}$), \mathbf{C}_s to \mathbf{C}_d (forming $\tilde{\mathbf{C}}_d$), and a zero vector of appropriate length to \mathbf{d} (forming $\tilde{\mathbf{d}}$). The reduced chi-squared statistic $\chi^2 = 2J/n$ is used to assess the fit of the inverse model to the observations (other examples of how this statistic can be used are described in e.g. Gurney et al., 2004;

Baker et al., 2006), where n is the number of observations. Computing the data uncertainty as described above ensures $\chi^2 \approx 1$, which means that the extent of the match between observations and the model as well as between the a priori and a posteriori sources are in accord with their respective uncertainties.

The cost function in Eq. (2) is minimized analytically (Enting, 2002; Tarantola, 2005) to yield a posteriori fluxes \mathbf{x} and covariance matrix \mathbf{C} :

$$\mathbf{x} = \mathbf{C} \left(\tilde{\mathbf{T}}^T \tilde{\mathbf{C}}_d^{-1} \tilde{\mathbf{d}} + \mathbf{C}_0^{-1} \mathbf{x}_0 \right)$$

$$\mathbf{C} = \left(\tilde{\mathbf{T}}^T \tilde{\mathbf{C}}_d^{-1} \tilde{\mathbf{T}} + \mathbf{C}_0^{-1} \right)^{-1}. \quad (3)$$

The square root of the diagonal elements of \mathbf{C} is reported as the uncertainty estimate for the a posteriori fluxes.

5.4 Sensitivity scenarios

Considerable effort has been undertaken to ensure e.g. the high quality of available observations, the inter-comparability of measurements from BHD and LAU, and the use of a state-of-the-art land process model to provide meaningful first-guess estimates. However, there are a number of potential sources for error that cannot be accounted for explicitly but could have a significant influence on estimated land fluxes. These include (i) the CO₂ baseline, (ii) the flux distribution within each region, and (iii) the ocean prior fluxes.

Sensitivity scenarios were designed to quantify each of these potential errors, as described below. The results of these scenarios and analysis of potential bias due to diurnal variability and atmospheric transport model error are discussed in Sect. 7.4.

- i. The inverse method assumes that air entering the domain is accurately characterized by the baseline CO₂ time series. While random noise in the baseline concentration is accounted for (Sect. 5.1), there remains the possibility of systematic errors. A positive (negative) bias in the baseline would cause the inversion to estimate a total CO₂ flux that is depressed (elevated), in order to explain the measurements at the stations. Sensitivity runs were conducted with both positive and negative biases. The baseline mixing ratio is first decreased, then increased, by 1 standard deviation, i.e. its uncertainty.
- ii. The inverse model optimizes total fluxes across large spatial regions, but the geographic distribution of the CO₂ fluxes within each region is not adjusted by the inversion but prescribed based on the prior flux distribution. This can lead to biases in the estimated flux if the region is being unevenly sampled. That is, if a specific observation is sensitive to only a small area inside the region, then the flux estimate for the entire region will

be biased towards that area, which may not be representative for the region. This is why we took the geographic distribution of biomes into account when defining the regions. The number of different biomes was minimized and isolated patches of biomes avoided inside each region. However, the region definition remained subjective, so we included a sensitivity case where the within-region flux pattern is flat; i.e. the flux is constant region-wide. Not all potential biases are removed this way, as that would require solving the inverse problem at a much higher resolution, but it gives an indication of the influence of a particular choice of pattern.

- iii. Estimates of terrestrial CO₂ fluxes in New Zealand are influenced by the ocean flux prior through atmospheric transport. After entering the model domain at baseline levels, the air travels inevitably over a large stretch of ocean and will arrive at the New Zealand coast carrying an oceanic signal in its CO₂ concentration. Errors in the oceanic prior flux estimates can result in inaccurate terrestrial CO₂ flux estimates. In a sensitivity test, the oceanic prior uncertainty was raised to 10⁸ Tg CO₂ yr⁻¹, effectively excluding the ocean prior.

6 Analysis of New Zealand's in situ CO₂ observing sites

We conducted a clustering analysis using NAME III to characterize the catchment areas of the BHD and LAU stations. The clustering was performed using a convergent k means procedure, which is based on Kidson (1994), but adjusted slightly to allow a larger number of trajectories to be clustered, i.e. by using a smaller number of random seeds. This significantly boosts the computation at the expense of likelihood to find the global minimum, however, the reduced number of seeds appeared large enough to come sufficiently close, as repeated computations with randomly different subsets all produced very similar results.

A set of 1000 trajectories was used between 15:00 and 16:00 LT for every day in 2011–2013, resulting in approximately 1 million trajectories for each station. The number of clusters was set to seven, because this number maximized the distinctness of clusters with respect to each other as obtained from their silhouette values. Cluster centroids and sizes are overlain on the station footprints in Fig. 4, together with the geographical width of the clusters.

In addition to the clustering analysis, we applied Eq. (1) to the a priori flux maps for every day in 2011–2013. This allows us to calculate the imprint of Australian and New Zealand fossil emissions as well as oceanic and New Zealand terrestrial sinks and sources on the CO₂ concentration measured at BHD and LAU (Fig. 7).

Generally, CO₂ measurements at BHD are most sensitive to sinks and sources in the Southern Ocean (south of 55° S), the Tasman Sea, and the South Island. Australia and the

North Island influence BHD CO₂ to a lesser extent. Observations at LAU are strongly influenced by local to regional terrestrial sinks and sources of CO₂, enabling the station to see air from a large portion of the southern South Island. Further site-specific details are given in the subsections below.

The low sensitivity to Australia means it is infeasible to infer Australian CO₂ fluxes with our observational network most of the time. However, this underscores the isolation of New Zealand, where air is received that largely contains background concentrations from the vast body of surrounding ocean. This allows us to estimate terrestrial fluxes in New Zealand with high sensitivity and little disturbing influence from continental sources.

6.1 Baring Head

In 2011–2013, BHD sampled air that has travelled from the Southern Ocean 41 % of the time along two cluster pathways (Fig. 4), which correspond to southerly wind conditions at the site. The more southerly cluster of the two (16 %) contains trajectories that mostly have not seen land over at least 4 days and will carry Southern Ocean baseline CO₂. Trajectories in the more westerly cluster (25 %) have travelled across most of the South Island after originating in the Southern Ocean and will carry a signal of the terrestrial sinks and sources of CO₂ there. Another 17 % of trajectories are originating from the south-west but are associated with slower wind speeds, so that within the 4-day time frame of the back trajectories they have not yet left the domain. They correspond to a local northerly wind at BHD associated with a common synoptic pattern involving an anticyclone over the Tasman Sea. The Southern Alps on the South Island strongly influence the south-westerly air flow and deflect it northward along the west coast and then through Cook Strait, where it is channelled into a northerly flow by local topography. Trajectories arriving from Australia and the Tasman Sea occur 13 and 9 % of the time, respectively; 10 % of the trajectories have crossed large parts of the North Island before arriving at the station.

The application of Eq. (1) to the a priori flux maps shows that there are only 4 days in 2011–2013 when a discernible (> 0.1 ppm) signal from Australian fossil fuel emissions within the inversion domain was received at BHD. The signal was always much smaller than 1.16 ppm, the minimum overall uncertainty assumed in the modelling system.

During the winter and summer seasons, the New Zealand land area is the main contributor to the BHD data series, with a seasonal pattern matching the respiration and growing cycles. Assumed aseasonal fossil emissions from New Zealand (mostly from the nearby city of Wellington) and seasonal oceanic fluxes also play an important role at BHD.

6.2 Lauder

For LAU, there are two southernmost clusters representing a combined 40 % percent of trajectories. These are very similar in size to the corresponding clusters for BHD and have identical source areas in the Southern Ocean. While similar to BHD's southern cluster, the air from the LAU southern cluster would have travelled over a considerable stretch of land before arriving at LAU. 14 % of the time, the air being sampled belongs to a south-western cluster, which originates in the Tasman Sea. In addition, there is a western cluster containing 15 % of trajectories that has crossed South Australia and the Tasman Sea as well as two northern clusters representing air with mixed origin from the northern Tasman Sea or the North Island. About 14 % of trajectories are contained in a slow cluster whose origin is not very far from LAU. These cases correspond to slow winds at the site and indicate that the measurements are highly impacted by local and regional sources as the air has been travelling over nearby land for the preceding 4 days.

The application of Eq. (1) to the a priori flux maps shows that LAU station is dominated by terrestrial fluxes from New Zealand (particularly from the South Island), with only minor contributions from the ocean, reflecting its location further inland and shielded from the predominant westerly winds by the Southern Alps. The seasonal amplitude in the CO₂ signal at LAU is about twice as large as at BHD, due to the more continental climate and more pronounced growing seasons in the central South Island. Similar to BHD, there are only 5 days in 2011–2013 when a larger than 0.1 ppm signal from Australian fossil fuel emissions within the inversion domain was received at LAU.

7 Flux results and discussion

7.1 Seasonal cycle

The inversion finds a much stronger seasonal cycle in terrestrial CO₂ fluxes than the Biome-BGC model simulations used as a prior (Fig. 5c), especially associated with enhanced CO₂ uptake during the growing season in (austral) summer (Fig. 10). There is very good agreement in the phasing of the seasons with the land process model during all 3 years, which is particularly encouraging in light of the weak constraints on the phasing applied through the prior (Sect. 5.3). This strong seasonality is robust within the estimated a posteriori uncertainty range and across the sensitivity cases. Uncertainties for weekly fluxes were reduced significantly compared to the prior, even when the range of sensitivity cases is added as extra uncertainty.

The enhanced seasonal amplitude is assigned to the South Island almost exclusively, with much stronger uptake during the growing season compared with carbon uptake in Biome-BGC. The uncertainties associated with South Island fluxes

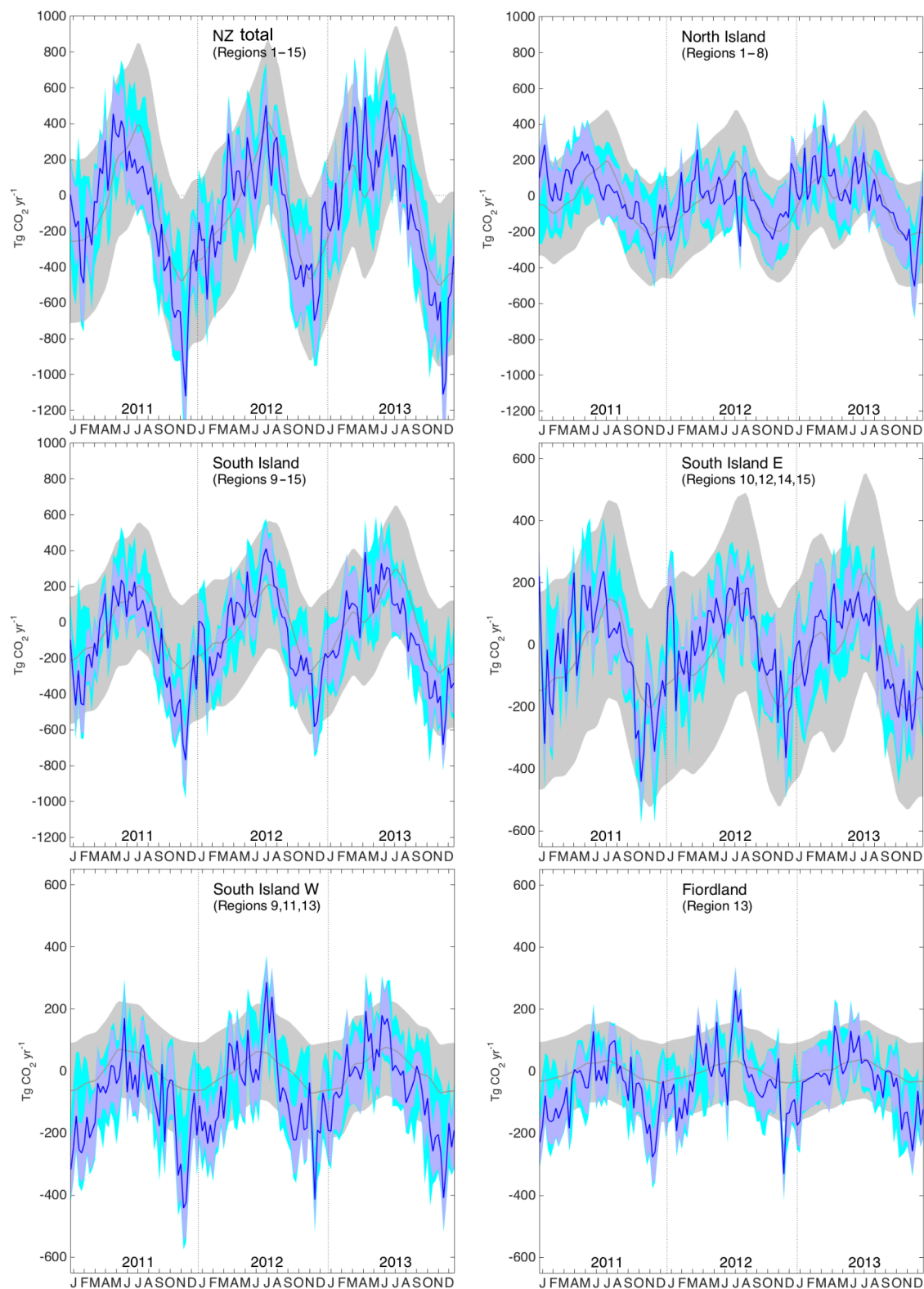


Figure 10. Weekly CO₂ fluxes in 2011–2013 from selected regions, in Tg CO₂ yr^{−1}. Prior flux estimates are shown in grey, and the inversion results are shown in blue. Shaded areas represent flux uncertainty (1 σ). The cyan shade represents the extra uncertainty obtained from the sensitivity cases (Sect. 7.4). Note that there is a one-off change in scale of the flux axis for sub-island scale regions. A positive flux indicates a net release of CO₂ to the atmosphere, while a negative value indicates uptake by the land biosphere. Regional indices in parentheses correspond to Fig. 8.

are generally smaller than on the country scale because of the high sensitivity of the LAU and BHD stations to fluxes from much of the South Island. In contrast, the North Island is estimated to have a weaker seasonal cycle, in good agreement with the prior, which can be attributed to widespread areas of summer soil water deficits, and the more marine climate there, i.e. weaker seasonal temperature variations and milder winters. Uncertainties for North Island fluxes, especially from the northern half of the North Island, are generally larger due to the lower sensitivity of the stations to that area. While northerly breezes are very common at BHD (Fig. 4), they often correspond to a situation where south-westerly air was deflected by the Southern Alps and channelled by local topography to turn into a northerly at the station. Air that has travelled across the North Island and picked up its terrestrial CO₂ signal is therefore less often sampled at BHD than local wind direction would suggest. At LAU, North Island air can be sampled only about 8 % of the time, based on the NAME cluster analysis. In a future study, the sensitivity to North Island fluxes can be greatly enhanced by CO₂ observations at the recently established Maunga Kākaramēa/Rainbow Mountain station in the central North Island.

When separating the South Island into parts east and west of the Southern Alps, it becomes apparent that most of the enhanced seasonal cycle occurs, in fact, in the west, despite the slightly smaller surface area (86 173 km² compared to 88 348 km²). Along the west coast, the inversion estimates the seasonal amplitude to be more than twice as large as suggested by the prior. Tracing the cause further to the individual regions reveals that Fiordland (region 13) is the strongest contributor to the signal. Fiordland is extremely sparsely populated and covered to a large extent by indigenous temperate rainforest with southern beeches (*Nothofagus* sp.), fern trees, and shrub. When forming the prior flux map, these forests were categorized as EBF and the respective module from Biome-BGC used. However, the EBF module had not been optimized for New Zealand forests, so it is possible that the Fiordland forests are not well described by that category. The inversion suggests much stronger photosynthetic and respiratory activity in these forests than the prior model.

7.2 Response to the 2012/2013 drought

In this section, we assess the ability of the inversion to detect the carbon flux response to drought. The austral summer of 2013 was characterized by unusually high temperatures and low precipitation over much of New Zealand (Turner, 2013; Blunden and Arndt, 2014), with sustained periods of severe drought in February–March 2013. The North Island and the west of the South Island were the most strongly affected regions (Porteous and Mullan, 2013). The Biome-BGC model is driven by detailed, reanalysed weather data and clearly shows a positive flux anomaly, i.e. loss of CO₂

to the atmosphere, due to enhanced respiration and inhibited growth during that period (Fig. 5c). The inversion sees this event in the observations as well, suggesting even more CO₂ release than Biome-BGC across the South Island. Unfortunately, a prolonged data gap in the LAU time series during that period caused by a lack of field standards (Fig. 1) leads to weaker constraints from the atmospheric CO₂ data and therefore larger uncertainty in the flux estimates in the South Island.

A signal of excess CO₂ release in February–March 2013 is seen by the inversion across the North Island, too (Fig. 10). The limited coverage of some areas in the North Island, especially in the north and east (Fig. 9), leads to high annual mean flux uncertainty for individual regions and prevents a robust analysis as to which regions responded the most strongly to the drought. Eddy-covariance data from a dairy pasture site in the north-western North Island during a ~ 100 day drought in 2008 found a temporary loss of CO₂ to the atmosphere, but the ecosystem recovered to become a net sink of CO₂ for the year (Mudge et al., 2011).

7.3 Annual fluxes

The geographic air–land flux distribution averaged over 2011–2013 is shown in Fig. 11, including flux gradients on the sub-regional level that were prescribed in the inversion. A comparison to the a priori distribution in Fig. 5b shows larger areas acting as a net carbon source. These include the central and north-eastern parts of the South Island, which roughly correspond to the Canterbury region and mostly contain pasture land, in particular sheep and beef pasture (Fig. 5a). The inversion does not, however, resolve ecosystem processes, but merely estimates net air–land fluxes, so it is not possible to make a causal link between pasture and a net CO₂ source. A counterexample is the south-east of the South Island, which also contains large areas of pasture but is estimated to be a net carbon sink. In general, the inversion assigns much more of the total land sink to forested areas than the Biome-BGC prior. This is particularly apparent along the western South Island, as well as in the eastern half of the North Island. The strong flux gradients seen in region no. 3 (north-western central North Island) are likely to be the result of the very heterogeneous composition of LCLU types there, combined with BHD and LAU having low sensitivity in the region (Fig. 9), rather than a real signal. The inclusion of an additional station with high sensitivity to the northern North Island, such as Maunga Kākaramēa/Rainbow Mountain, would be needed to improve flux estimates there.

In the inset of Fig. 11, we compare annual mean results from the inversion with bottom-up estimates from the National Inventory Report (MfE, 2015), or NIR. The inversion suggests a much larger net CO₂ sink across the country compared to the NIR. Particularly in the forest of the south-western South Island, the inversion suggests both stronger photosynthetic and respiratory activity than the prior model,

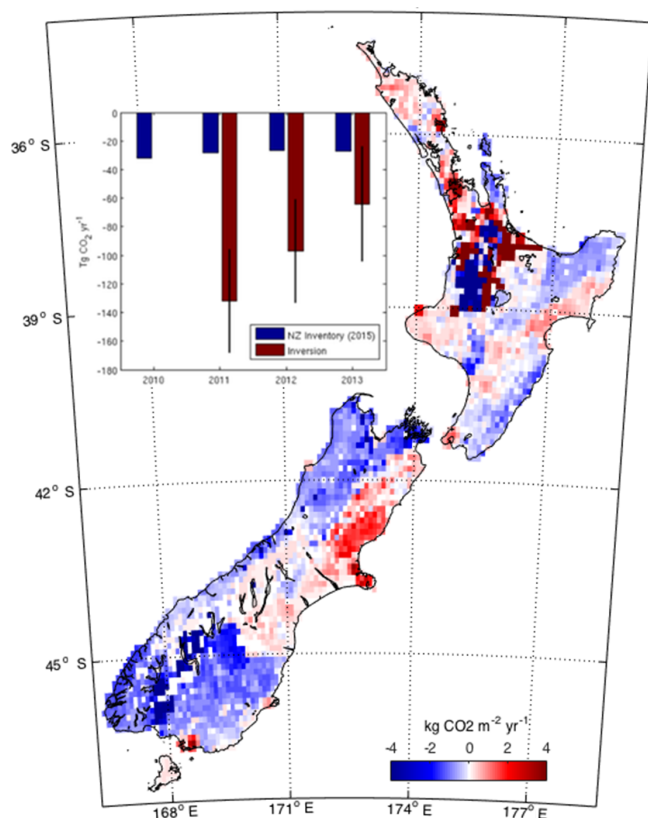


Figure 11. Geographic distribution of land-to-air CO₂ flux, averaged over 2011–2013. Blue and red regions indicate net carbon uptake and release, respectively. Per-area ocean fluxes are too small to show on this scale. Fossil fuel emissions are included and reach up to 20 kg CO₂ m⁻² yr⁻¹ in a few grid cells (Auckland area). The colour scale is capped to focus on natural fluxes. Inset: annual mean results compared to the National Greenhouse Gas Inventory Report.

with the overall balance towards a larger CO₂ sink over the course of a year. For example, Fiordland appears to take up between 22 and 68 Tg CO₂ each year in 2011–2013 (Table 1), which corresponds to per-area uptake rates of 614 and 1899 g CO₂ m⁻² yr⁻¹, respectively. By comparison, the Biome-BGC estimates range from 0 to 3 Tg CO₂ yr⁻¹.

The NIR estimates do not come with an overall uncertainty, but based on their reporting of typical uncertainty for individual ecosystems, and personal communication, an approximate figure of 50 % of the flux value was identified. This implies statistical significance for the difference in annual sink estimates, except in 2013, when both estimates agree within their uncertainty range. Without ocean prior or by assuming a baseline bias of 0.1 ppm, the differences are reduced by up to a half (Sect. 7.4) but do not disappear. How can these differences be explained? There are a number of possible scenarios, which we explore in the following.

The accounting of fossil emissions differs between the NIR and the inversion. The EDGAR emissions of 47.8 Tg CO₂ yr⁻¹ prescribed in the inversion contain ele-

ments of land-use change and agriculture. The NIR gives total emissions of 34.6 Tg CO₂ yr⁻¹ for 2013. The difference of about 13 Tg CO₂ yr⁻¹ would appear in the inversion as an additional sink of equal size.

The inversion and NIR estimates are not directly comparable due to differences in the top-down vs. bottom-up viewpoints. While the inversion sees the overall net CO₂ exchange between the atmosphere and the land, the NIR estimate represents the so-called LULUCF sector; i.e. it includes contributions from land use, land-use change, and forestry. In the LULUCF model, it is assumed that CO₂ emissions from harvested wood products occur at the location of the tree, a process particularly important to forest plantations located in the central North Island and in the north of the South Island (Fig. 5a). However, about 70 % of the biomass from forest harvesting is exported before major processing, i.e. in the form of logs, sawn timber, or manufactured wood products (Pike, 2014). Most of the CO₂ release associated with harvesting will subsequently occur far away from New Zealand, e.g. in China, with a 34 % share of New Zealand's forestry exports in 2012. In a regional inversion these emissions cannot be seen, leaving a larger net sink. Emissions from harvested wood products are reported in the NIR at 10.3 Tg CO₂ yr⁻¹ in 2013, translating into about 7 Tg CO₂ yr⁻¹ that cannot be seen by the inversion when assuming a 70 % export rate. No emissions are reported for earlier years because the harvested wood products category was introduced for the first time in the 2015 report. The 2013 estimate is likely to be an upper bound for the years 2011 and 2012 because the volume of harvested wood products has increased steadily since 2009 (MfE, 2015). Other possible discrepancies between the NIR methodology and the net CO₂ fluxes for forests include the variance in the timing of root carbon emission following tree mortality (Kirschbaum et al., 2013). Large sinks observed but not accounted for in NIR can result from applying steady state assumptions to natural or pre-1990 forests when they are accumulating carbon in biomass during recovery from past disturbance, with potential rates of biomass accumulation by native species reaching 700–900 g CO₂ m⁻² yr⁻¹ (Trotter et al., 2005).

For pastoral agriculture, a more complex set of differences applies to the intercomparison of inversion results, a priori process-based model results and the NIR methodology. Similar to forestry, agricultural exports (e.g. milk, meat, and wool) equated to 340 g CO₂ m⁻² yr⁻¹ for a dairy pasture (Mudge et al., 2011), and the 165 Gg of nitrogen estimated as exported in produce (Parfitt et al., 2006) will equate to an apparent net CO₂ uptake of 5.8 Tg CO₂ yr⁻¹ across New Zealand. The second-most important gap between methodologies results from the NIR calculation that 6.3–6.5 % of the energy content of pasture consumed by ruminants is converted to CH₄ emissions. These CH₄ emissions represent a carbon flux to the atmosphere not observable as CO₂ and therefore require separate quantification. They have been calculated as 79 g CO₂ m⁻² yr⁻¹ in a dairy pasture (Mudge et

al., 2011) and the carbon content of the NIR's 1137 Gg of CH₄ emissions equates to an unobserved 3.1 Tg CO₂ yr⁻¹. Several additional terms, including leaching of dissolved carbon forms, and imports of feed and fertiliser can also provide important corrections between NEP seen by inversions and eddy covariance and net ecosystem carbon balance (NECB) (Mudge et al., 2011).

The NIR methodology also does not account for above or below-ground grassland biomass nor does it account for soil carbon changes. The process-based model Biome-BGC potentially accounts for both these flux terms, but not in relation to intensive management. Therefore, both biomass and soil carbon must be considered to explain additional CO₂ uptake or loss by pastures that might be seen by the inversion, but not by NIR or Biome-BGC.

Biomass carbon is relatively small in New Zealand pastures (Tate et al., 1997) but can be a significant component of seasonal net exchange (Mudge et al., 2011; Rutledge et al., 2015; Hunt et al., 2016) as described in Sect. 7.1. Repeated measurements of soil profiles suggest that soil carbon changes can also be significant but uncertain due to limited sites available for resampling. For example, a recent analysis of all sites available nationally suggests that sites on flat pasture are losing soil carbon at rates of $\sim 170 \text{ g CO}_2 \text{ m}^{-2} \text{ yr}^{-1}$, while sites in hill country are gaining $\sim 770 \text{ g CO}_2 \text{ m}^{-2} \text{ yr}^{-1}$ (Schipper et al., 2014). In addition to large areas of grazed pastures on both islands, significant areas of tussock grasslands on the South Island could be gaining biomass and soil carbon as they recover from historic overgrazing (Tate et al., 1997). The extensive area of grasslands on both islands could result in large net CO₂ exchange fluxes usefully observed by inversion studies. New Zealand's first process-based studies of net national ecosystem carbon balance suggested large uncertainties in grasslands (Tate et al., 2000) and later suggested grasslands were approximately carbon neutral in 2001 (Trotter et al., 2004). Eddy-covariance studies remain limited in coverage across New Zealand but tend to suggest potential for large negative NEP and near neutral NECB. Rutledge et al. (2015) updated and extended the Mudge et al. (2011) results to 4 years, yielding average NEP of $600 \pm 180 \text{ g CO}_2 \text{ m}^{-2} \text{ yr}^{-1}$ and NECB of $220 \pm 200 \text{ g CO}_2 \text{ m}^{-2} \text{ yr}^{-1}$. Hunt et al. (2016) also report eddy-covariance carbon budgets for an irrigated intensively grazed dairy pasture and an unirrigated winter-grazed pasture in Canterbury on the South Island's east coast. Over 1 year, the unirrigated pasture was carbon neutral ($\pm 80 \text{ g CO}_2 \text{ m}^{-2} \text{ yr}^{-1}$), while the intensively managed and irrigated pasture displayed NEP of $1500 \pm 140 \text{ g CO}_2 \text{ m}^{-2} \text{ yr}^{-1}$ and NECB of $380 (\pm 150) \text{ g CO}_2 \text{ m}^{-2} \text{ yr}^{-1}$. The forest and grassland studies described above suggest that large, negative flux anomalies estimated by the inversion may be plausible when extrapolated across the large areas of these LCLU categories.

Additional real land carbon balance terms may also contribute to large, negative flux anomalies that differ from

NIR and process-based models such as Biome-BGC. These terms include areas of organic soil accumulation in wet forests and bogs, typical of west-coast environments where the largest negative flux anomalies are observed. Campbell et al. (2014) used eddy covariance to find net ecosystem exchange (NEE) of $800\text{--}900 \text{ g CO}_2 \text{ m}^{-2} \text{ yr}^{-1}$ with a strong seasonal cycle. Erosion and deposition can also create a net carbon sink that may be unusually significant at active tectonic margins such as New Zealand (Tate et al., 2000; Baisden and Manning, 2011). Small catchment and site studies have estimated rates of net pasture soil carbon sequestration due to erosion and burial, accounting for upland soil carbon recovery, of $220 \text{ g CO}_2 \text{ m}^{-2} \text{ yr}^{-1}$ (Page et al., 2004) and $370 \text{ g CO}_2 \text{ m}^{-2} \text{ yr}^{-1}$ (Parfitt et al., 2013). Scott et al. (2006) have estimated the national delivery of eroded carbon to the coast as $11 \pm 4 \text{ Tg CO}_2 \text{ yr}^{-1}$ and suggested that much of this carbon is likely to be buried and replaced in uplands. Dymond (2010) attempts to more fully and dynamically account for erosion, burial, and replacement, suggesting a range of $4\text{--}20 \text{ Tg CO}_2 \text{ yr}^{-1}$. Both studies suggest the largest erosion-induced CO₂ sinks occur in the Southern Alps in the west of the South Island, where the Lauder station allows observation of a strong sink, as well as in the North Island's east coast. These estimates may partly be included in the hill country soil carbon accumulation estimated by Schipper et al. (2014). Smith et al. (2015) suggest that fiords may also create a strong carbon sink, with about 18 Mt of organic carbon being buried in fiord sediments globally each year, yielding a rate of $198 \text{ g CO}_2 \text{ m}^{-2} \text{ yr}^{-1}$. Thus, a number of plausible mechanisms have been documented in the literature to explain that the CO₂ sink seen by the inversion is stronger than estimated in the NIR.

7.4 Uncertainty and bias assessment

In addition to regional uncertainty, the posterior covariance matrix from the Bayesian optimization also contains weekly error correlations between regional flux estimates, which are used to inform flux uncertainties. Strong negative correlations between two regions would indicate that the inversion has difficulties distinguishing their individual flux components with the available data, while their sum is better constrained. Similarly, positive correlations indicate that the difference of flux components is constrained better than each individually. An analysis of all regional error correlations reveals that both negative and positive correlations are present; however, only 0.13 % of all pairwise correlations have an absolute value greater than 0.1. Hence, with the available data, the inversion appears able to resolve weekly fluxes at the chosen regional level.

An analysis of the mismatch of modelled CO₂ (i.e. the CO₂ time series obtained by propagating the posterior flux through the transport model) and observed CO₂ reveals differences between the BHD and LAU stations (Fig. 12). At BHD, the mismatch distributions are very similar for the

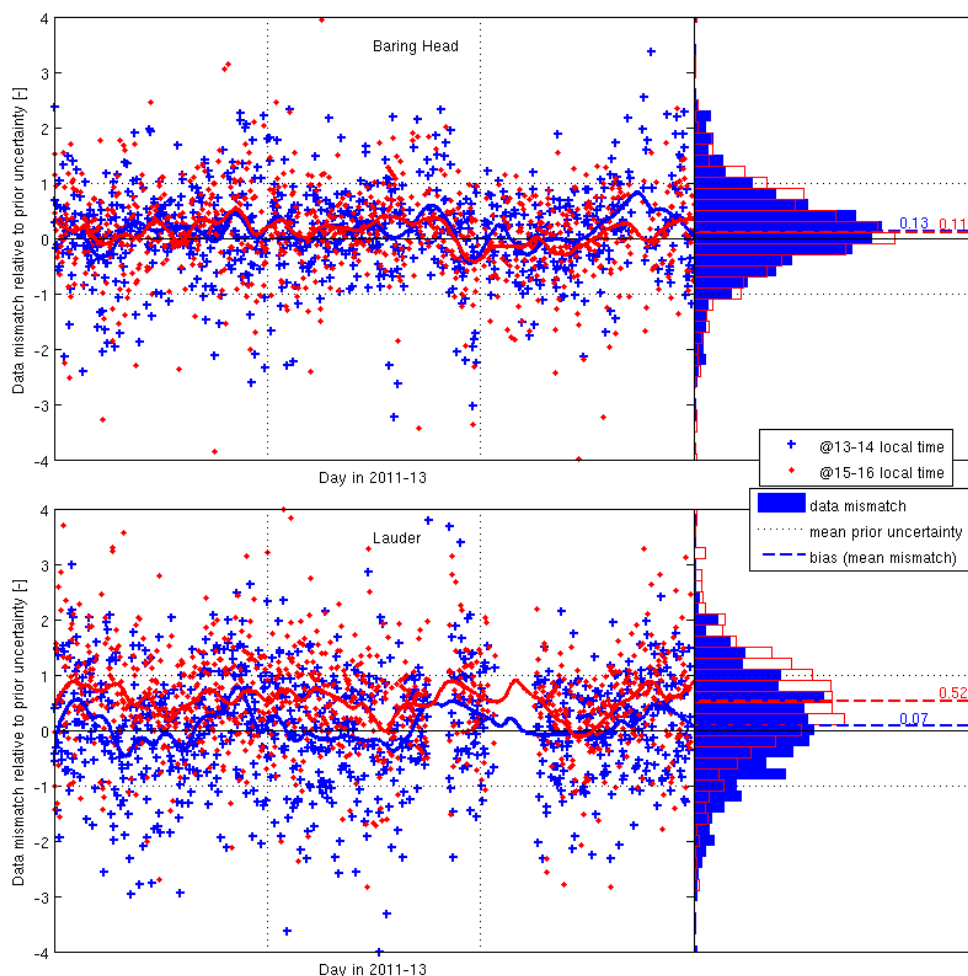


Figure 12. Mismatch (residuals) of modelled vs. observed CO₂ in multiples of the prior data uncertainty at Baring Head (top) and Lauder (bottom). Vertical dotted lines separate the years 2011, 2012, and 2013. Solid lines represent a Loess fit with a 3-month window. Horizontal dotted lines mark the prior uncertainty (1σ). The left column shows scatter plots for every day and 1 h release period; the right column shows the mismatch distribution over 2011–2013. Dashed lines with numbers give the bias.

13:00–14:00 and 15:00–16:00 time series, have a positive bias (modelled CO₂ is higher than observations) of 11–13 % of the prior data uncertainty, and show no discernible temporal pattern (smoothed, thick lines in the figure). At LAU, the mismatch distribution is similar for the 13:00–14:00 time series, with an even smaller bias of 7 %, but for the 15:00–16:00 time series there is a much larger bias of 52 % of the data uncertainty. This means the inversion has difficulties reproducing the low CO₂ concentrations in the LAU 15:00–16:00 observational record. The temporal evolution indicates an alternating pattern of small mismatch during the (austral) winter and larger mismatch during summer.

In an attempt to explain this mismatch, we looked at two possible sources of bias: (1) a misrepresentation of the PBL depth in the NZLAM model and (2) diurnal variability in the fluxes that is not captured because we use afternoon data only and do not have diurnal variability in our a priori fluxes.

7.4.1 Biases in the modelled planetary boundary layer

At a site like LAU, strong solar radiation during a clear summer day might lead to a sudden deepening of the PBL in the afternoon between the two release periods, which could prove difficult for NZLAM to model accurately. Consequently, if the modelled PBL is deeper than in reality, any signal from surface fluxes would be mixed in a larger volume of air and their effect on modelled CO₂ concentrations would be attenuated. This would require the inversion to estimate stronger surface fluxes in order to explain measured CO₂ concentrations. This would lead to an overestimation of CO₂ uptake during summer. The opposite is true if the modelled PBL is too shallow and the measured concentrations can be explained with less CO₂ uptake. We compared the model PBL depth at 15:00–16:00 to radiosonde measurements made at LAU (Fig. 13). The Heffter method (Heffter,

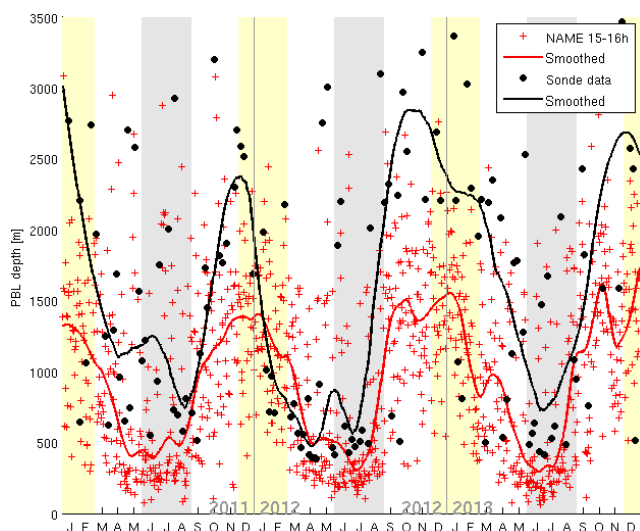


Figure 13. Comparison of boundary layer depth at LAU in NAME at 15:00–16:00 LT and radiosonde observations made at the site (Heffter method). The seasonal cycle has been made more visual using a robust Loess smoother. Summer periods are highlighted in yellow, winter periods in grey.

1980) was used to compute PBL height from the radiosonde data. The comparison suggests that the boundary layer is too shallow in the model during summer, thus enhancing the effect of CO₂ uptake on the modelled concentrations. Thus, smaller fluxes should be required to match the uptake signals observed at LAU, and therefore this process does not explain the observed mismatch. An equivalent analysis with the 13:00–14:00 LT PBL data suggests a similar discrepancy, yet these data can be explained by the inversion, further reducing the likelihood that the PBL representation in the model causes the observed mismatch. However, this analysis has caveats, because the radiosonde dataset is preliminary, and only few measurements were taken during the right times of day (13:00–14:00 and 15:00–16:00 LT, respectively).

7.4.2 Diurnal variability

Our method estimates surface fluxes on a weekly scale based on afternoon measurements and does not resolve the diurnal cycle of the CO₂ land natural fluxes. The diurnal cycle is particularly pronounced in summer, with strong uptake during the afternoon. This could cause the inversion to overestimate the weekly sink and thus bias the annual sink towards more uptake, in an attempt to match the low afternoon observations that are being assimilated. This issue is of most concern near the stations, because varying atmospheric transport integrates flux signals from more distant regions across a larger range of times. We would therefore expect this effect to be most pronounced in the local regions surrounding each station. However, based on a synthetic data experiment described below, we find that this is not the case; the sink in

Table 1. Annual mean CO₂ flux for selected aggregated regions, in Tg CO₂ yr^{−1}. A negative sign indicates uptake by the land. In parentheses the a posteriori uncertainty (1 σ) is shown (excluding the sensitivity cases).

Region	2011	2012	2013
NZ total	−132 (36)	−97 (36)	−64 (40)
North Island	18 (28)	−40 (28)	−1 (30)
North	5 (25)	−10 (25)	7 (25)
South	13 (17)	−30 (17)	−8 (19)
South Island	−149 (22)	−56 (23)	−63 (28)
East	−37 (17)	9 (18)	−10 (23)
West	−113 (17)	−65 (16)	−52 (17)
Fiordland	−68 (13)	−22 (12)	−31 (14)

the two local regions is actually slightly suppressed and does not reveal any discernible seasonal pattern.

We conducted a synthetic data experiment in order to assess the potential bias caused by diurnal variations not represented in our inversion framework. A synthetic dataset that included the atmospheric signature of diurnal variability was generated using hourly footprints from the transport model (instead of the 4-day mean footprints described in Sect. 3.1) together with hourly Biome-BGC flux maps for 2012. These flux maps were prepared by imposing a simplified diurnal cycle on daily, gridded Biome-BGC outputs of gross primary production (GPP), net primary production (NPP), and heterotrophic respiration (HR). The diurnal variation in GPP is directly proportional to the relative amount of solar insolation, and HR is assumed to occur at a constant rate and spread evenly throughout the day. There are a number of aspects of plant physiology and ecosystem biogeochemistry that cause actual diurnal variation in NEE to be more muted than the solar-radiation-driven pattern we have modelled. These include reductions in photosynthesis during the middle of the day and afternoon as stomata close due to drought or leaf temperature stress. Similarly, reduced respiration can be expected at night due to cooler temperatures. Synthetic CO₂ concentrations were created for every day at 15:00–16:00 LT by propagating the hourly fluxes through the hourly model. This set of synthetic mole fractions was then assimilated in an inversion analogous to the reference inversion, whereby oceanic prior fluxes, fossil emissions, and the data uncertainties were kept the same as in the reference case. The weekly averages of the hourly fluxes then represent the “truth” against which the posterior fluxes can be compared. The weekly averages of the true fluxes are also used as a prior in the inversion, albeit with an uncertainty twice as large as in the reference inversion, to limit its influence on the results.

Unrepresented diurnal variability led to biases in the annual mean flux estimates for some regions in our inversion, but these errors were much smaller than our uncertainty estimates for most regions on an annual scale (Fig. 14). Im-

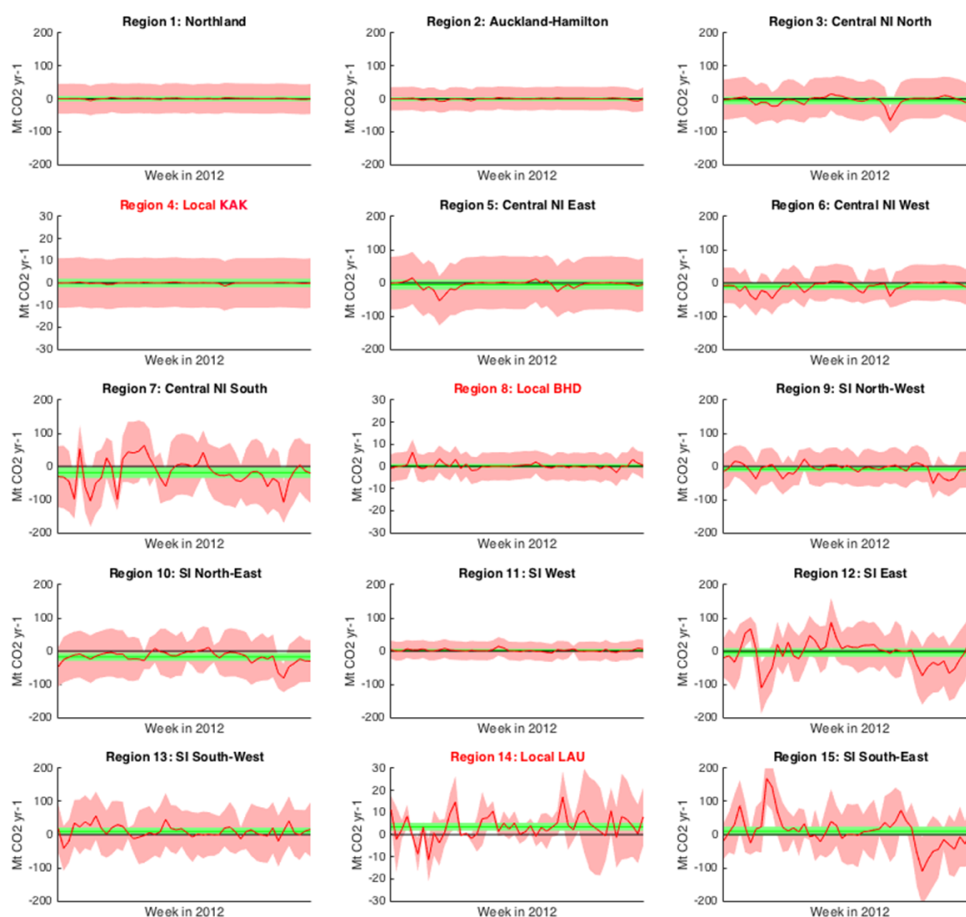


Figure 14. Deviation of the posterior flux from the “true” flux (red lines) in the synthetic data experiment with a diurnal flux cycle for the 15 land regions in 2012. Patched areas give the posterior uncertainty (excluding sensitivity cases). Green bands represent the annual mean deviation, with their thickness corresponding to the annual uncertainty. Note there is a change in scale of the flux axis for the three local regions (red font).

portantly, this experiment does not reveal a substantial systematic bias in the annual mean across all regions. Two of the best constrained regions (upwind of LAU during average wind conditions), regions 13 and 14 (south-western South Island and local LAU region), show a suppressed, rather than overestimated, annual mean sink and this bias emerges without a distinctive seasonal pattern. This is especially surprising in the local region around the LAU station where the effect should be strongest. This suggests that a diurnal bias is unlikely to be the cause of the observed 15:00–16:00 data mismatch, as the latter is characterized by a clear seasonal cycle. Adding to this, the other local region (around BHD, region 8) does not show any distinctive bias for both weekly and annual sinks. This is less surprising as the BHD site is characterized by very high wind speeds for much of the year (Stephens et al., 2013) and therefore less sensitive to local effects. Overall, diurnal variability leads to an overestimation of the annual mean land sink in the South Island of

3.5 Tg CO₂ yr^{−1}, which is well within the uncertainty envelope of the reference inversion (Table 1).

On a weekly timescale, estimates generally agree to within their uncertainties for most regions, with the exception of the eastern South Island (Regions 12 and 15), the southern central North Island (region 7), and the local area region around Lauder (region 14). Regions 12 and 15 (eastern and south-eastern South Island) show an increased sink late in the year (as part of the 2012/2013 summer) but a smaller sink early in the year (as part of the 2011/2012 summer), suggesting that there may be a seasonal bias in our inverse methodology for the eastern South Island. Likewise, diurnal cycle bias leads to significant weekly errors in the central North Island (region 7), although with less seasonal coherence. The Lauder local region (region 14) was created to capture local signals that are not well represented in our inverse model, including diurnal variability, and prevent them from biasing the inverse estimates on larger spatial scales. Thus the larger errors for this region are expected.

7.4.3 Sensitivity scenarios

Results from the sensitivity scenarios (i)–(iii) are incorporated in the figures as an additional uncertainty band on top of the Bayesian posterior uncertainties from the default run (Fig. 10). That band represents the maximum (minimum) value of the flux plus (minus) its uncertainty at every point in time and across all runs, i.e. including the default and sensitivity runs.

While the uncertainty range associated with the suite of sensitivity scenarios is symmetrical around the reference case for most regions, the sensitivity range is characterized by more positive flux estimates than the reference case in the western South Island (Fig. 10), i.e. a slightly smaller annual carbon sink. This can be attributed to sensitivity case (iii), in which the inversion is allowed to adjust air–sea fluxes to any value without penalty. Some of the terrestrial CO₂ uptake is relocated to upwind ocean regions, as this yields a lower Bayesian cost, because fluxes from the western South Island are shifted towards the Biome-BGC prior estimates. However, in order to offset a relatively small flux change on land, the change in ocean flux has to be large due to the distance to the stations and the dilution of CO₂ concentrations on the way. This leads to an ocean sink of 6 Pg CO₂ in 2012 in our regional domain for this sensitivity test, which is more than 10 times larger than estimates for the whole Southern Ocean from global inversions, ocean carbon data, and ocean biogeochemistry models (Gruber et al., 2009). Despite this unrealistic result for the oceans in the sensitivity test, the conclusions about the seasonal pattern in CO₂ uptake and release and its spatial distribution in the New Zealand land regions remain robust.

The inversion assumes an unbiased baseline CO₂ record. Any positive (negative) bias would be interpreted by the inversion as an additional sink (source) of CO₂. From the sensitivity runs we find that a constant bias in the baseline of 0.1 ppm would cause the total CO₂ flux of New Zealand for each year in 2011–2013 to be offset by approximately 20 Tg CO₂ yr^{−1}. This corresponds to about 50 % of the flux uncertainty from the default run (Table 1), thus underscoring the importance of an accurate baseline in a regional inversion. An inversion such as ours can always benefit from advances in air–sea flux datasets, such as *p*CO₂ measurements from a regional cruise network, as well as well-characterized background air concentrations. However, similar to the sensitivity case without ocean prior, the biased baseline has only a minor influence on seasonal flux patterns over land.

Current plans to aid the baseline representation in future top-down studies of the New Zealand region include the addition of in situ measurement sites that observe inflow from the north-westerly and westerly sectors on the upper North Island or west coast of the South Island, which will improve the characterization of synoptic-scale variability in inflow that does not originate from the Southern Ocean sector.

8 Conclusions

We present the first regional inversion estimates of air–land and air–sea CO₂ fluxes for the New Zealand region, which were estimated from two in situ observing stations in New Zealand, ship-based measurements, and Lagrangian model simulations using the NAME dispersion model driven by NZLAM meteorology. The results imply a strong seasonal cycle, especially for fluxes in the western South Island. Regions covered predominantly by indigenous forest appear to have more pronounced photosynthetic and respiratory activity than suggested by the land model's EBF category. This is most apparent in Fiordland, which is a key contributor to the seasonal cycle, as well as the annual mean sink, in the South Island. The timing, magnitude, and regional distribution of seasonal flux patterns are well constrained and robust across sensitivity cases, while uncertainties in annual totals are more significant. Enhanced CO₂ release from the terrestrial biosphere in New Zealand is apparent in response to the 2012/2013 drought period. This response appears most prominent in the North Island and western parts of the South Island, consistent with reports about these regions being most severely affected.

The annual total CO₂ sink in New Zealand is estimated to have decreased over the 3-year period, at 132 ± 36 , 97 ± 36 , and 64 ± 40 Tg CO₂ yr^{−1} in 2011, 2012, and 2013, respectively. The New Zealand national inventory reports a much smaller sink of 28, 27, and 27 Tg CO₂ yr^{−1} for the same years (with uncertainty around 50 %). About 7 Tg CO₂ yr^{−1} of the discrepancy can be attributed to emissions associated with forest harvesting, which are included in the inventory but missed by the inversion due to forestry exports. Another 13 Tg CO₂ yr^{−1} arise from different accounting of fossil emissions between the inventory and the inversion. Additional factors relating to the difference between NEP and NECB in pastures can account for another 9 Tg CO₂ yr^{−1}. Other terms such as erosion, burial, and soil carbon recovery may account for another 4–20 Tg CO₂ yr^{−1}. These differences largely reconcile both results for 2013 but not 2011–2012. Carbon sequestration by grassland and soil carbon could also play an important role in causing differences between the two methods, as these processes are not included or fully resolved in inventory reporting but would be seen by the inversion. Collectively, these factors are likely to reconcile both results only partially, with some differences remaining.

Detailed sensitivity studies suggest that the most important causes of uncertainty in the inverse estimates are uncertainties in the estimate of baseline air entering the domain and air–sea fluxes from the ocean surrounding New Zealand. These uncertainties could be reduced through higher density of *p*CO₂ measurements in the oceans around New Zealand, and extending the ship-based atmospheric CO₂ measurements presently used to estimate the baseline air farther to the south and west. Another possibility is to establish addi-

tional surface stations in strategic locations, i.e. with footprints in areas where Lauder and Baring Head have low sensitivity. While it was out of scope for this study, more detailed analyses of the uncertainties associated with the atmospheric transport model are needed, particularly the role of model resolution given the complex topography of New Zealand.

The inversion methodology developed here is a powerful tool to validate net regional CO₂ sinks in the New Zealand national inventory report. It offers an independent, top-down view on the national carbon budget. The limited sensitivity to the northern half of the North Island, as well as baseline errors, can lead to large uncertainties for annual mean flux estimates in some regions. Improving on these factors in future studies can further increase the usefulness of the top-down approach.

9 Data availability

The model code has been made publicly available in a GitHub/Zenodo (doi:10.5281/zenodo.215972) repository (Steinkamp et al., 2016).

Appendix A: Lauder site description

A1 The Lauder station

The Lauder atmospheric research station (45.038° S, 169.684° E, 370 m a.m.s.l.) is located in the broad Manuherikia river valley on the South Island of New Zealand. A semi-arid continental climate predominates with an annual rainfall of 450 mm and mean annual temperature of 9.7°C. The prevailing wind is from the westerly quarter (a mean daily wind run of approximately 300 km). Periodic southerly frontal systems bring air masses from the Southern Ocean and Tasman Sea. The station is surrounded by pastoral land dominated by non-intensive sheep and cattle farming practices along with seasonal cropping. The valley is sparsely populated. The land westward (on average upwind) of the valley consists of numerous valley systems and mountainous terrain. The vast majority of this land is undeveloped and is part of New Zealand's national park system. There is no major industry present in the region.

Due to the relatively clear unclouded skies, low light pollution, and low levels of local and regional anthropogenic emissions, “clean air” ground-based remote sensing, balloon sonde, and in situ measurements are routinely conducted at the station as part of NDACC (formerly known as NDSC) (Kurylo, 1991), GAW (WMO-GAW, 2007), TCCON (Wunch et al., 2011), and GRUAN (Seidel et al., 2009) activities.

A2 Lauder in situ trace gas measurements

Long-term routine in situ measurements began at Lauder in 2003 with the installation of a TEI-49C Ozone monitor (Zellweger et al., 2010). Previous to this only sporadic short-term campaigns focusing on tropospheric nitrogen dioxide had been undertaken (Johnston and McKenzie, 1984). Continuous in situ measurements of carbon dioxide (CO₂), methane (CH₄), nitrous oxide (N₂O), and carbon monoxide (CO) began in March 2007 when a prototype FTIR trace gas analyser was installed (Griffith et al., 2012; Sepúlveda et al., 2014). In June 2008 a well-calibrated continuous CO₂ NDIR (differential, non-dispersive, infrared) analyser (LI-7000, manufactured by LI-COR, Inc., USA, <http://www.licor.com>) was installed at Lauder. This was followed by regular fortnightly flask samples analysed for CO₂, CH₄, N₂O, CO, and $\delta^{13}\text{C}$ -CO₂ concentrations, starting in May 2009. An added advantage of employing the NDIR analysers at both sites (Lauder and Baring Head) is that they share common data processing code and calibration routines.

A3 Air inlet system

The air inlet system consists of a permanent 10 m high metal NIWA meteorological mast erected at a distance of 33 m to the north, of the nearest building, housing the in situ instrumentation. Two sets of 60 m long stainless steel (SS) tubing

(ID 8.8 mm) are used to sample air from the mast (inlet at 10 m) to two distribution manifolds; prior to 2012 these were baked copper tube (ID 8.8 mm). The inlets are fitted with inverted funnels with coarse mesh (0.7 mm) to provide rain and dust protection. For electrical isolation of the instrumentation and the meteorological mast a 100 mm length of PFA 9.5 mm tubing is inserted between the sampling lines and the manifolds. The manifolds are constructed from 25 mm SS diameter tubing 200 mm in length (volume = 0.086 L). Each port consists of a SS tube (OD 6.3 mm) welded perpendicular to the main body, these extend 15 mm centrally into the main body and terminated with a 45° angle cut facing the direction of flow. Sample air is drawn into the two 4-port manifolds with a roughing pump (KNF Neuberger, N035 AN18) at 10–15 L min⁻¹ giving an effective residence time of approximately 35 s and an associated pressure drop of 40 mbar. The roughing pump allows sample air to be drawn at a higher flow rate and allows multiple instruments to sample air without front end pressure coupling between co-sampling instruments.

The LI-7000 samples air, at 2.6 L min⁻¹, from one of the manifolds connected to the 10 m sampling inlet, with Synflex (Eaton Synflex 1300, OD = 6.3 mm) tubing. An FTIR trace gas analyser samples air from a port on the same manifold through a 6.3 mm PFA tube (300 mm length) at a rate of 3.5 L min⁻¹, and a flask sampling system is connected to the same manifold port with Nylon (Ledalon[®] 1200 series nylon 12 tubing, OD = 3.2 mm) tubing. When flask samples are taken a flow rate of up to 2 L min⁻¹ is used.

A4 LI-7000

A commercially available dual cell LI-COR NDIR analyser the LI-7000 is able to calculate CO₂ mole fractions via measurements in the CO₂ 4.255 μm absorption band. The LI-7000 has been proven to be a low maintenance, robust CO₂ analyser able to meet GAW measurement criteria when operated in the correct manner (WMO, 2001). A gas delivery and data acquisition and diagnostics system designed by NIWA (Gomez, 1997) on a LabView[®] platform is used to automate and manage the delivery sample and reference air along with calibration gas. The Lauder gas handling and data acquisition system is similar to that described by Brailsford et al. (2012). The main difference is that a LI-7000 is employed at Lauder whereas at Baring Head a Siemens Ultramat 3 gas analyser is used.

The LI-7000 draws air from the manifold via a diaphragm pump (KNF Neuberger, KNF 86KNE, 2.6 L min⁻¹). A set of four field standards, with a calibration lineage to the mole fraction scale maintained by the CCL and a target-archive tank, is connected to a valve manifold consisting of five three-way (Parker B16DK1175) valves in a daisy chain configuration, along with the dried air allowing selection of either field standards, target tank, or sample air for the analyser. Gas regulators (Scott Marrin Inc., 1-SS30-590-DAT)

and 1.6 mm SS tubing are used to connect tanks to the gas delivery system. On the outlet of the sample pump an over-pressure is maintained on the inlet to a Nafion drier (Perma Pure LLC, MD-110-144S-4) with the excess flow vented at this point; this removes the bulk of the water content from the sample flow. The air or selected calibration gas then passes through a magnesium perchlorate trap to ensure all the same low water content for gas being introduced to the analyser by a 100 sccm mass flow controller (McMillian, 80SD-5). One of the field standards is also used in the reference cell as a reference gas and is controlled using a similar mass flow controller (McMillian, 80SD-3) at 10 sccm. The exhaust sample and reference gas are joined together and then dried again on molecular sieve trap before acting as the counter flow on the Nafion drier; in this way dew points of -65°C are consistently met.

The data acquisition system selects the calibration gas to measure and monitors each field standard for stability to optimize the gas consumption. When a field standard has a standard deviation of less than 0.015 ppm over a minute it is defined as stable and the next gas is measured. Sample air is continuously measured with 5 min averages collated and reported. Every 4–6 h the suite of four field standards is measured. A target–archive tank is measured every 23 h. Each week the field standards and target tank are measured as a separate aliquot multiple times. This sampling sequence is akin to the calibration protocol employed by Brailsford et al. (2012) and Stephens et al. (2011). Data processing is performed by Lauder LI-7000 specific scripts adapted from those used by Stephens et al. (2011) and written in the free statistical analysis software R.

Allan variance measurements (Allan, 1966) show the precision of the Lauder LI-7000 system as 0.004 ppm (1σ in 5 min). Calibration of the LI-7000 is obtained by fitting a third-order polynomial to the measurements of the four field standards to characterize the concentration-dependent non-linear response of the instrument every 4–6 h. This calibration curve is then used to calibrate sample air measurements, putting the measurements on the WMO X2007 scale. Thirty-minute zero offsets are calculated using the interspersed individual field standard measurements. Instrument-dependent artefacts (e.g. instrument temperature and flushing times) are accounted for in the processing code by calculating a linear fit of known field standard concentrations and the parameter in question.

The field standards and the target tank are filled at BHD and characterized at the NIWA GASLAB, Greta Point, Wellington, New Zealand. In preparing the field standards mixtures are prepared to evenly span the typical air sample concentration range encountered, including elevated nocturnal levels (typical span of 380–450 ppm). The field standard CO₂ concentrations are calibrated to the WMO X2007 scale, along with $\delta^{13}\text{C}$ –CO₂ (VPDB referenced to the WMO CCL scale) (Brailsford et al., 2012). Field standards require changing every 12–18 months. The target tank requires changing

every 6–12 months as in parallel it also functions as a target tank for the FTIR trace gas analyser.

A5 Meteorological sensors

Meteorological sensors were installed onto the sampling mast. Wind speed is measured at three heights (2.8, 5.8, and 10.1 m) using Vector instruments A100LK anemometers. A Vector instruments W200P wind vane mounted at a height of 10.1 m is used to record wind direction. Relative humidity and temperature are measured using Vaisala Humitter 50U/50Y sensors, placed at heights of 2.6 and 9.9 m. In addition, a Vaisala PTB100 analog barometer was installed adjacent to the in situ instruments (inside the building). All these sensors are connected to a Campbell CR10X data logger and SDM-INT8 logger module. Ten-minute averages of all sensor output are recorded independently of in situ gas measurement instrumentation output.

Appendix B: Baseline analysis

A CO₂ baseline is constructed as a weighted average of a southern and northern baseline, which takes into account whether the modelled trajectories originated to the north or south of the inversion domain for a given data point.

B1 Southern baseline

The southern baseline represents a continuous record of steady background CO₂ mole fractions during southerly wind conditions at BHD. A multi-step filter is applied to the BHD record to obtain a CO₂ baseline representative of a large region over the Southern Ocean, as described by Brailsford et al. (2012) and Stephens et al. (2013). In short, the filter selects measurements during extended periods of southerly winds at the site, during which a maximum standard deviation of 0.1 ppm is achieved. Additional meteorological conditions must be fulfilled to preclude the influence of local sources and to ensure the air has not passed over the South Island before arriving at BHD. The result of this filtering process is similar to selecting observations from the southern cluster in Fig. 4. The southern baseline based on the filtering is used in this study.

After filtering the data for baseline conditions, a continuous baseline is constructed using the STL algorithm (Cleveland et al., 1990), which allows estimation of a long-term trend and interannually varying seasonal patterns. The STL algorithm uses two time windows for the seasonal cycle and the trend, which are set by the user and define the respective time periods over which variations in the data are considered. The monthly averaged data fulfilling the baseline conditions are used as input, and the algorithm is run first with a seasonal cycle window of 5 years and a trend window of 121 months to single out the decadal trend. This trend is then removed before a second run with a trend window of 25 months

to capture the interannual and seasonal patterns. Finally the decadal, interannual, and seasonal time series are summed and the resulting baseline subsampled at the 13:00–14:00 and 15:00–16:00 LT windows.

The remainder time series, which is the difference between the monthly record and the sum of the seasonal and trend components from the STL analysis, is used as uncertainty estimate for the southern baseline.

B2 Northern baseline

The northern baseline is based on in situ CO₂ observations using a NDIR analyser on board the TF5, a ship of opportunity that cruised the triangle Japan–Australia–New Zealand about once a month during the period 2011–2013 (Chierici et al., 2006). The cluster analysis showed that during northerly events the air is usually coming from the northern Tasman Sea or the subtropical waters to the north and only occasionally from the South Pacific eastward of New Zealand. The layout of the TF5 cruises with legs crossing the Tasman Sea as well as subtropical legs therefore offers the possibility to characterize the CO₂ concentrations in these regions with monthly resolution. The TF5 dataset provides CO₂ concentrations averaged over 10 min intervals along with the standard deviation of the high-frequency measurements within these intervals.

We defined a regional mask to keep observations from the open ocean and avoid observations taken close to the land, especially near the Australian east coast as it is located upwind during average south-westerly conditions and hosts large urban centres with significant CO₂ emissions. The mask spans the latitudes 39 to 24° S. The mask was then further partitioned into bands spanning 1° of latitude and the data within each band averaged for each month in 2011–2013. The uncertainty of the resulting monthly record is taken as the quadrature sum of the standard deviation of high-frequency data points measured during the ship's transit of the respective latitude band and the standard deviation of the 10 min data about the monthly mean.

The monthly record within each latitude band was analysed using the same STL routine as for the southern baseline. An overall uncertainty estimate was formed by root mean square combination of the monthly uncertainty and the time series of the remainder from the STL analysis. The remainder time series is the difference between the monthly record and the sum of the seasonal and trend components from the STL analysis. Finally, the STL baseline for the two latitude bands for 27–26° S were averaged to produce a baseline representative of the northern edge of the inversion domain at 26° S.

B3 Weighted baseline

A day-to-day baseline was constructed as a weighted superposition of the southern and northern baselines, with weights depending on the proportional latitude of air origin for the twice-daily measurements at BHD and LAU. The daily NAME station footprints for the 13:00–14:00 and 15:00–16:00 LT windows were integrated along the southern and northern edges of the domain to determine the relative fraction of back trajectories leaving the domain to the south and north. These fractions are then used to weigh the two baselines and create a baseline associated with each of the twice-daily data points. Uncertainties are weighed in the same way.

For most days the 13:00–14:00 and 15:00–16:00 LT footprints are similar with regard to the origin of the air, so the weighted baselines for both time windows are almost identical. For both stations on a typical day the region where the air originates is either clearly in the north or the south, so that the weights for the southern and northern baselines are close to zero and one, or vice versa. However, middle cases can occur when the wind conditions at a site rapidly changed during the 1-hour period over which measurements are collected, which often results in two main branches of trajectories originating in the north and south, respectively. In this case both baselines are weighted proportionally to reflect the mixed origin of air during the 1-hour averaging period for the measurements. However, other days, or periods of days, are characterized by slow wind speeds, sometimes slow enough that most back trajectories end before reaching either the northern or southern edge of the domain. In this case, the midpoint of the footprint is determined and its latitude used to proportionally weigh the baselines. The same procedure applies for days with trajectories leaving the domain predominantly to the west or the east.

Acknowledgements. The author(s) wish to acknowledge the contribution of New Zealand eScience Infrastructure (NeSI) to the results of this research. New Zealand's national compute and analytics services and team are supported by the NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation and Employment (<http://www.nesi.org.nz>). In addition, Kay Steinkamp, Gordon Brailsford, Dan Smale, and Sara E. Mikaloff Fletcher would like to acknowledge NIWA core funding through the Greenhouse Gases, Emissions and Carbon Cycle Science Programme. We thank our colleagues from New Zealand's Ministry for the Environment (MfE), especially the LUCAS team, for very fruitful meetings. None of this work could have been accomplished without the station operation teams at Baring Head and Lauder. We thank Hinrich Schaefer for valuable discussions about this paper. We are grateful for access to radiosonde data from Lauder, and would like to thank Ben Liley for his PBL calculations. We would also like to thank Paul Wennberg and his TCCON team for the generous loan of their Li-7000 that is currently at Lauder. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Edited by: C. Gerbig

Reviewed by: two anonymous referees

References

- Allan, D. W.: Statistics of atomic frequency standards, *IEEE Proc.*, 54, 221–230, 1966.
- Baisden, W. T. and Manning, M. R.: Editorial: The New Zealand carbon cycle: from regional budget to global cycle, *Biogeochemistry*, 104, 1–4, doi:10.1007/s10533-011-9579-x, 2011.
- Baker, D., Law, R., Gurney, K., Rayner, P., Peylin, P., Denning, A., Bousquet, P., Bruhwiler, L., Chen, Y. H., and Ciais, P.: TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO₂ fluxes, 1988–2003, *Global Biogeochem. Cy.*, 20, GB1002, doi:10.1029/2004GB002439, 2006.
- Bergamaschi, P., Krol, M., Meirink, J. F., Dentener, F., Segers, A., van Aardenne, J., Monni, S., Vermeulen, A. T., Schmidt, M., Ramonet, M., Yver, C., Meinhardt, F., Nisbet, E. G., Fisher, R. E., O'Doherty, S., and Dlugokencky, E. J.: Inverse modeling of European CH₄ emissions 2001–2006, *J. Geophys. Res.*, 115, D22309, doi:10.1029/2010jd014180, 2010.
- Blunden, J. and Arndt, D. S.: State of the Climate in 2013, *B. Am. Meteorol. Soc.*, 95, 1–257, 2014.
- Brailsford, G. W., Stephens, B. B., Gomez, A. J., Riedel, K., Mikaloff Fletcher, S. E., Nichol, S. E., and Manning, M. R.: Long-term continuous atmospheric CO₂ measurements at Baring Head, New Zealand, *Atmos. Meas. Tech.*, 5, 3109–3117, doi:10.5194/amt-5-3109-2012, 2012.
- Campbell, D. I., Smith, J., Goodrich, J. P., Wall, A. M., and Schipper, L. A.: Year-round growing conditions explains large CO₂ sink strength in a New Zealand raised peat bog, *Agr. Forest Meteorol.*, 192–193, 59–68, doi:10.1016/j.agrformet.2014.03.003, 2014.
- Chierici, M., Fransson, A., and Nojiri, Y.: Biogeochemical processes as drivers of surface *f*CO₂ in contrasting provinces in the subarctic North Pacific Ocean, *Global Biogeochem. Cy.*, 20, GB1009, doi:10.1029/2004GB002356, 2006.
- Ciais, P., Canadell, J. G., Luysaert, S., Chevallier, F., Shvidenko, A., Poussi, Z., Jonas, M., Peylin, P., King, A. W., and Schulze, E.-D.: Can we reconcile atmospheric estimates of the Northern terrestrial carbon sink with land-based accounting?, *Current Opinion in Environmental Sustainability*, 2, 225–230, 2010.
- Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I.: STL: A seasonal-trend decomposition procedure based on loess, *Journal of Official Statistics*, 6, 3–73, 1990.
- Davies, T., Cullen, M. J. P., Malcolm, A. J., Mawson, M. H., Staniforth, A., White, A. A., and Wood, N.: A new dynamical core for the Met Office's global and regional modelling of the atmosphere, *Q. J. Roy. Meteor. Soc.*, 131, 1759–1782, 2005.
- Dymond, J. R.: Soil erosion in New Zealand is a net sink of CO₂, *Earth Surf. Proc. Land.*, 35, 1763–1772, doi:10.1002/esp.2014, 2010.
- Dymond, J. R., Shepherd, J. D., Newsome, P. F., Gapare, N., Burgess, D. W., and Watt, P.: Remote sensing of land-use change for Kyoto Protocol reporting: the New Zealand case, *Environ. Sci. Pol.*, 16, 1–8, 2012.
- EC-JRC/PBL: European Commission, Joint Research Center/Netherlands Environmental Assessment Agency, Emission Database for Global Atmospheric Research (EDGAR), release version 4.2, available at: <http://edgar.jrc.ec.europa.eu/index.php> (last access: 21 December 2016), 2011.
- Enting, I. G.: Inverse problems in atmospheric constituent transport, Cambridge University Press, 2002.
- Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., and Knutti, R.: Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks, *J. Climate*, 27, 511–526, 2014.
- Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO₂ with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework, *J. Geophys. Res.-Atmos.*, 108, 4757, doi:10.1029/2003JD003770, 2003.
- Gomez, A. J.: The CO₂ remote control system at Baring Head, Aspendale, Australia, 1997.
- Griffith, D. W. T., Deutscher, N. M., Caldow, C., Kettlewell, G., Riggensbach, M., and Hammer, S.: A Fourier transform infrared trace gas and isotope analyser for atmospheric applications, *Atmos. Meas. Tech.*, 5, 2481–2498, doi:10.5194/amt-5-2481-2012, 2012.
- Gruber, N., Gloor, M., Mikaloff Fletcher, S. E., Doney, S. C., Dutkiewicz, S., Follows, M. J., Gerber, M., Jacobson, A. R., Joos, F., and Lindsay, K.: Oceanic sources, sinks, and transport of atmospheric CO₂, *Global Biogeochem. Cy.*, 23, GB1005, doi:10.1029/2008GB003349, 2009.
- Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Pak, B. C., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y. H., and Ciais, P.: Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks, *Global Biogeochem. Cy.*, 18, GB1010, doi:10.1029/2003GB002111, 2004.
- Heffter, J.: Transport layer depth calculations, paper presented at: Second Joint Conference on Applications of Air Pollution Meteorology, Am. Meteorol. Soc., New Orleans, LA, 24–28, 1980.

- Hendy, J., Kerr, S., and Baisden, T.: The Land Use in Rural New Zealand Model Version 1 (LURNZv1): Model Description, Motu Working Paper 07-07, Motu Economic and Public Policy Research, Wellington, NZ, 2007.
- Ho, D. T., Law, C. S., Smith, M. J., Schlosser, P., Harvey, M., and Hill, P.: Measurements of air-sea gas exchange at high wind speeds in the Southern Ocean: Implications for global parameterizations, *Geophys. Res. Lett.*, 33, L16611, doi:10.1029/2006GL026817, 2006.
- Hunt, J. E., Laubach, J., Barthel, M., Fraser, A., and Phillips, R. L.: Carbon budgets for an irrigated intensively grazed dairy pasture and an unirrigated winter-grazed pasture, *Biogeosciences*, 13, 2927–2944, doi:10.5194/bg-13-2927-2016, 2016.
- IPCC: Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change, Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press, 2013.
- Johnston, P. and McKenzie, R.: Long-path absorption measurements of tropospheric NO₂ in rural New Zealand, *Geophys. Res. Lett.*, 11, 69–72, 1984.
- Jones, A., Thomson, D., Hort, M., and Devenish, B.: The UK Met Office's next-generation atmospheric dispersion model, NAME III, Air Pollution Modeling and its Application XVII, 580–589, 2007.
- Keller, E. D., Baisden, W. T., Timar, L., Mullan, B., and Clark, A.: Grassland production under global change scenarios for New Zealand pastoral agriculture, *Geosci. Model Dev.*, 7, 2359–2391, doi:10.5194/gmd-7-2359-2014, 2014.
- Kerr, S., Anastasiadis, S., Olssen, A., Power, W., Timar, L., and Zhang, W.: Spatial and temporal responses to an emissions trading scheme covering agriculture and forestry: simulation results from new zealand, *Forests*, 3, 1133–1156, 2012.
- Kidson, J. W.: An automated procedure for the identification of synoptic types applied to the New Zealand region, *Int. J. Climatol.*, 14, 711–721, 1994.
- Kirschbaum, M. U. F., Saggar, S., Tate, K. R., Thakur, K. P., and Giltrap, D. L.: Quantifying the climate-change consequences of shifting land use between forest and agriculture, *Sci. Total Environ.*, 465, 314–324, doi:10.1016/j.scitotenv.2013.01.026, 2013.
- Kurylo, M. J.: Network for the detection of stratospheric change, Orlando'91, Orlando, FL, 168–174, 1991.
- Landcare Research: Fundamental Soil Layers, available at: <https://soils.landcareresearch.co.nz/soil-data/fundamental-soil-layers> (last access: 21 December 2016), 2015.
- Lauvaux, T., Uliasz, M., Sarrat, C., Chevallier, F., Bousquet, P., Lac, C., Davis, K. J., Ciais, P., Denning, A. S., and Rayner, P. J.: Mesoscale inversion: first results from the CERES campaign with synthetic data, *Atmos. Chem. Phys.*, 8, 3459–3471, doi:10.5194/acp-8-3459-2008, 2008.
- Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G. R., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., Jourdain, C., Kato, E., Keeling, R. F., Klein Goldewijk, K., Levis, S., Levy, P., Lomas, M., Poulter, B., Raupach, M. R., Schwinger, J., Sitch, S., Stocker, B. D., Viogy, N., Zaehle, S., and Zeng, N.: The global carbon budget 1959–2011, *Earth Syst. Sci. Data*, 5, 165–185, doi:10.5194/essd-5-165-2013, 2013.
- Lin, J., Gerbig, C., Wofsy, S., Andrews, A., Daube, B., Davis, K., and Grainger, C.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, *J. Geophys. Res.-Atmos.*, 108, 4493, doi:10.1029/2002JD003161, 2003.
- Lowe, D., Guenther, P., and Keeling, C.: The concentration of atmospheric carbon dioxide at Baring Head, New Zealand, *Tellus*, 31, 58–67, 1979.
- Manning, A. J.: The challenge of estimating regional trace gas emissions from atmospheric observations, *Philos. T. R. Soc. A*, 369, 1943–1954, doi:10.1098/rsta.2010.0321, 2011.
- Manning, A. J., O'Doherty, S., Jones, A. R., Simmonds, P. G., and Derwent, R. G.: Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modeling approach, *J. Geophys. Res.*, 116, D02305, doi:10.1029/2010jd014763, 2011.
- Matross, D. M., Andrews, A., and Pathmathevan, M.: Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data, *Tellus B*, 58, 344–358, 2006.
- MfE: New Zealand's Greenhouse Gas Inventory 1990–2013, Ministry for the Environment (MfE) – Manatu Mo Te Taiao, Wellington, New Zealand, 2015.
- Mikaloff Fletcher, S. E., Gruber, N., Jacobson, A. R., Gloor, M., Doney, S. C., Dutkiewicz, S., Gerber, M., Follows, M., Joos, F., and Lindsay, K.: Inverse estimates of the oceanic sources and sinks of natural CO₂ and the implied oceanic carbon transport, *Global Biogeochem. Cy.*, 21, GB1010, doi:10.1029/2006GB002751, 2007.
- MPI (Ministry for Primary Industries): National exotic forest description – as at 1 April 2012, Edn. 29, 2012.
- Morrison, N. and Webster, H.: An assessment of turbulence profiles in rural and urban environments using local measurements and numerical weather prediction results, *Bound.-Lay. Meteorol.*, 115, 223–239, 2005.
- Mudge, P. L., Wallace, D. F., Rutledge, S., Campbell, D. I., Schipper, L. A., and Hosking, C. L.: Carbon balance of an intensively grazed temperate pasture in two climatically: Contrasting years, *Agr. Ecosyst. Environ.*, 144, 271–280, doi:10.1016/j.agee.2011.09.003, 2011.
- Page, M. J., Trustrum, N. A., Brackley, H., and Baisden, W. T.: Erosion-related soil carbon fluxes in a pastoral steepland catchment, New Zealand, *Agr. Ecosyst. Environ.*, 103, 561–579, 2004.
- Parfitt, R. L., Schipper, L. A., Baisden, W. T., and Elliott, A. H.: Nitrogen inputs and outputs for New Zealand in 2001 at national and regional scales, *Biogeochemistry*, 80, 71–88, 2006.
- Parfitt, R. L., Baisden, W. T., Ross, C. W., Rosser, B. J., Schipper, L. A., and Barry, B.: Influence of erosion and deposition on carbon and nitrogen accumulation in resampled steepland soils under pasture in New Zealand, *Geoderma*, 192, 154–159, doi:10.1016/j.geoderma.2012.08.006, 2013.
- Pike, C.: Statistics New Zealand – Tatauranga Aotearoa, Logs to China drive our forestry export growth: http://www.stats.govt.nz/tools_and_services/media-centre/additional-releases-archive/forestry-exports-13-january-14.aspx, (last access: 21 December 2016), 2014.
- Porteous, A. and Mullan, B.: The 2012–13 drought: an assessment and historical perspective, Ministry for Primary Industries (MPI), available at: <https://www.niwa.co.nz/sites/>

- niwa.co.nz/files/2013-18-The%202012-13%20drought%20an%20assessment%20and%20historical%20perspective.pdf (last access: 21 December 2016), 2013.
- Rutledge, S., Mudge, P. L., Campbell, D. I., Woodward, S. L., Goodrich, J. P., Wall, A. M., Kirschbaum, M. U. F., and Schipper, L. A.: Carbon balance of an intensively grazed temperate dairy pasture over four years, *Agr. Ecosyst. Environ.*, 206, 10–20, doi:10.1016/j.agee.2015.03.011, 2015.
- Schipper, L. A., Parfitt, R. L., Fraser, S., Littler, R. A., Baisden, W. T., and Ross, C.: Soil order and grazing management effects on changes in soil C and N in New Zealand pastures, *Agr. Ecosyst. Environ.*, 184, 67–75, doi:10.1016/j.agee.2013.11.012, 2014.
- Scott, D. T., Baisden, W. T., Davies-Colley, R., Gomez, B., Hicks, D. M., Page, M. J., Preston, N. J., Tate, K. R., Trustrum, N. A., and Woods, R. A.: Localized erosion affects New Zealand's national C budget, *Geophys. Res. Lett.*, 33, L01402, doi:10.1029/2005GL024644, 2006.
- Seidel, D. J., Berger, F. H., Diamond, H. J., Dykema, J., Goodrich, D., Immler, F., Murray, W., Peterson, T., Sisterson, D., Sommer, M., and Thorne, P.: Reference upper-air observations for climate: Rationale, progress, and plans, *B. Am. Meteorol. Soc.*, 90, p. 361, 2009.
- Sepúlveda, E., Schneider, M., Hase, F., Barthlott, S., Dubravica, D., García, O. E., Gomez-Pelaez, A., González, Y., Guerra, J. C., Gisi, M., Kohlhepp, R., Dohe, S., Blumenstock, T., Strong, K., Weaver, D., Palm, M., Sadeghi, A., Deutscher, N. M., Warneke, T., Notholt, J., Jones, N., Griffith, D. W. T., Smale, D., Brailsford, G. W., Robinson, J., Meinhardt, F., Steinbacher, M., Aalto, T., and Worthy, D.: Tropospheric CH₄ signals as observed by NDACC FTIR at globally distributed sites and comparison to GAW surface in situ measurements, *Atmos. Meas. Tech.*, 7, 2337–2360, doi:10.5194/amt-7-2337-2014, 2014.
- Shepherd, J. and Newsome, P.: Establishing new zealand's Kyoto land use and land use-change and forestry 2008 map, Contract Report LC0809/133 prepared for the Ministry for the Environment, Lincoln, New Zealand, Landcare Research, 2009.
- Smith, R. W., Bianchi, T. S., Allison, M., Savage, C., and Galy, V.: High rates of organic carbon burial in fjord sediments globally, *Nat. Geosci.*, 8, 450–453, doi:10.1038/ngeo2421, 2015.
- Steinkamp, K. and Gruber, N.: A joint atmosphere-ocean inversion for the estimation of seasonal carbon sources and sinks, *Global Biogeochem. Cy.*, 27, 732–745, 2013.
- Steinkamp, K., Mikaloff Fletcher, S. E., Brailsford, G., Smale, D., Moore, S., Keller, E. D., Baisden, W. T., Mukai, H., and Stephens, B. B.: New-Zealand-CO₂-inversion: First release of the New Zealand CO₂ inversion system, doi:10.5281/zenodo.215972, 2016.
- Stephens, B. B., Miles, N. L., Richardson, S. J., Watt, A. S., and Davis, K. J.: Atmospheric CO₂ monitoring with single-cell NDIR-based analyzers, *Atmos. Meas. Tech.*, 4, 2737–2748, doi:10.5194/amt-4-2737-2011, 2011.
- Stephens, B. B., Brailsford, G. W., Gomez, A. J., Riedel, K., Mikaloff Fletcher, S. E., Nichol, S., and Manning, M.: Analysis of a 39-year continuous atmospheric CO₂ record from Baring Head, New Zealand, *Biogeosciences*, 10, 2683–2697, doi:10.5194/bg-10-2683-2013, 2013.
- Stohl, A., Seibert, P., Arduini, J., Eckhardt, S., Fraser, P., Grealley, B. R., Lunder, C., Maione, M., Mühle, J., O'Doherty, S., Prinn, R. G., Reimann, S., Saito, T., Schmidbauer, N., Simmonds, P. G., Vollmer, M. K., Weiss, R. F., and Yokouchi, Y.: An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons, *Atmos. Chem. Phys.*, 9, 1597–1620, doi:10.5194/acp-9-1597-2009, 2009.
- Sweeney, C., Gloor, E., Jacobson, A. R., Key, R. M., McKinley, G., Sarmiento, J. L., and Wanninkhof, R.: Constraining global air-sea gas exchange for CO₂ with recent bomb ¹⁴C measurements, *Global Biogeochem. Cy.*, 21, GB2015, doi:10.1029/2006GB002784, 2007.
- Tait, A., Henderson, R., Turner, R., and Zheng, X.: Thin plate smoothing spline interpolation of daily rainfall for New Zealand using a climatological rainfall surface, *Int. J. Climatol.*, 26, 2097–2115, 2006.
- Takahashi, T., Sutherland, S. C., and Kozyr, A.: Global ocean surface water partial pressure of CO₂ database: measurements performed during 1968–2008 (Version 2008), ORNL/CDIAC-152, NDP-088r, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Dept. of Energy, Oak Ridge, Tenn., 2009a.
- Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., and Sabine, C.: Climatological mean and decadal change in surface ocean pCO₂, and net sea-air CO₂ flux over the global oceans (Vol. 56, p. 554, 2009), *Deep-Sea Res. Pt. I*, 56, 2075–2076, 2009b.
- Tarantola, A.: Inverse problem theory and methods for model parameter estimation, Society for Industrial and Applied Mathematics (SIAM), 2005.
- Tate, K. R., Giltrap, D. J., Claydon, J. J., Newsome, P. F., Atkinson, I. A. E., Taylor, M. D., and Lee, R.: Organic carbon stocks in New Zealand's terrestrial ecosystems, *J. Roy. Soc. New Zeal.*, 27, 315–335, 1997.
- Tate, K. R., Scott, N. A., Parshotam, A., Brown, L., Wilde, R. H., Giltrap, D. J., Trustrum, N. A., Gomez, B., and Ross, D. J.: A multi-scale analysis of a terrestrial carbon budget – Is New Zealand a source or sink of carbon?, *Agr. Ecosyst. Environ.*, 82, 229–246, 2000.
- Thornton, P., Law, B., Gholz, H. L., Clark, K. L., Falge, E., Ellsworth, D., Goldstein, A., Monson, R., Hollinger, D., and Falk, M.: Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests, *Agr. Forest Meteorol.*, 113, 185–222, 2002.
- Thornton, P. E., Running, S. W., and Hunt, E.: Biome-BGC: terrestrial ecosystem process model, Version 4.1.1, Model product, available at: <https://daac.ornl.gov> (last access: 21 December 2016), Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, 2005.
- Timar, L.: Rural Land Use and Land Tenure in New Zealand, SSRN 1970100, Motu Economic and Public Policy Research Trust, GNS Science, 2011.
- Trotter, C. M., Tate, K. R., Saggart, S., Scott, N. A., and Sutherland, M. A.: A multi-scale analysis of a national terrestrial carbon budget and the effects of land-use change, in: *Global environmental change in the ocean and on land*, edited by: Shiyomi, M., Kawahata, H., Koizumi, H., Tsuda, A., and Awaya, Y., TERRAPUB, Tokyo, 311–341, 2004.
- Trotter, C. M., Tate, K. R., Scott, N. A., Townsend, J. A., Wilde, R. H., Lambie, S. M., Marden, M., and Pinkney, E. J.: Afforesta-

- tion/reforestation of New Zealand marginal pasture lands by indigenous shrublands: the potential for Kyoto forest sinks, *Ann. For. Sci.*, 62, 865–871, 2005.
- Turner, R.: Seasonal Climate Summary, NIWA National Climate Centre, available at: <https://www.niwa.co.nz/climate/summaries/seasonal/summer-2012-13> (last access: 21 December 2016), 2013.
- Uglietti, C., Leuenberger, M., and Brunner, D.: European source and sink areas of CO₂ retrieved from Lagrangian transport model interpretation of combined O₂ and CO₂ measurements at the high alpine research station Jungfraujoch, *Atmos. Chem. Phys.*, 11, 8017–8036, doi:10.5194/acp-11-8017-2011, 2011.
- Wanninkhof, R.: Relationship between wind speed and gas exchange over the ocean, *J. Geophys. Res.-Oceans*, 97, 7373–7382, 1992.
- WMO: Global Atmosphere Watch Measurements Guide, World Meteorological Organization (WMO), WMO TD No. 1073, 2001.
- WMO-GAW: WMO/GAW Strategic Plan: 2008–2015 – A Contribution to the Implementation of the WMO Strategic Plan: 2008–2011, World Meteorological Organization (WMO)/Global Atmosphere Watch (GAW), 2007.
- Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W., Sherlock, V., and Wennberg, P. O.: The total carbon column observing network, *Philos. T. R. Soc. Lond. A*, 369, 2087–2112, 2011.
- Zellweger, C., Steinbacher, M., Buchmann, B., and Scheel, H.: System and Performance Audit of Surface Ozone, Methane, Carbon Dioxide, Nitrous Oxide and Carbon Monoxide at the Global GAW Station Lauder, New Zealand, March 2010, WCC-Empa Report 10/3, 2010.