University of BRISTOL

QUEST Quantifying and Understanding the Earth System

DEPARTMENT OF EARTH SCIENCES

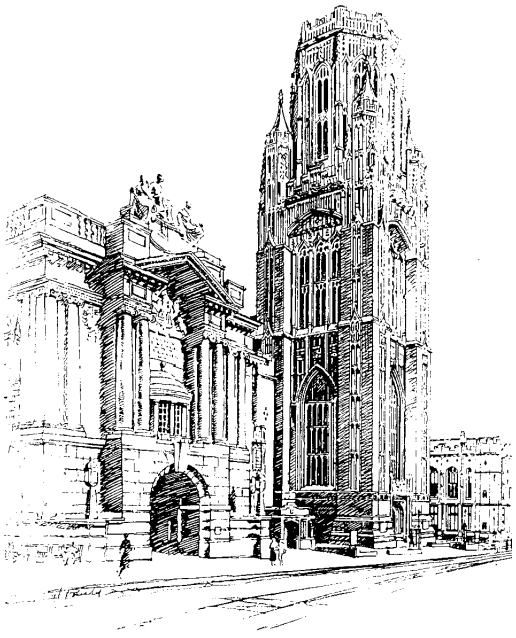
NAME: Alice Hooker-Stroud, MSci

DEGREE: Earth System Science MSc

TITLE: Anthropogenic CO_2 — Seasonal Fossil Fuel Emissions in CCDAS

YEAR OF SUBMISSION: 2008

SUPERVISOR: Dr W. Knorr/Dr M. Scholze



Wills Memorial Building
University of Bristol
Queen's Road, Bristol BS8 1RJ

2008

Anthropogenic CO_2

— Seasonal Fossil Fuel Emissions in CCDAS

Alice Hooker-Stroud, MSci

University of Bristol

18^{th} Sep 2008

Acknowledgements

Many many thanks to both my supervisors - Marko Scholze and Wolfgang Knorr for their patience and kindness regarding all of my queries; and for their guidance though this project. Also, to those involved in the successful organisation of the first year of the Earth Systems Science MSc, my congratulations. Thank you to all who shared their knowledge and expertise with us, without whom this year would not have been as insightful or interesting. Finally, to the EPSRC for awarding me the scholarship which enabled me to study on this course.

Contents

Abstract

A seasonal cycle of fossil fuel CO_2 emissions replaces an intra-annually invariant background flux in the global carbon cycle model CCDAS (Carbon Cycle Data Assimilation System). CO_2 emissions are divided into five regions. The fossil fuel seasonality function is a basic sinusoidal cycle with terms varying latitudinally and monthly. A tuneable parameter representing the amplitude and phase of the cycle in each region is calibrated through data assimilation of atmospheric CO_2 concentrations at 41 observation stations.

There is an overall improvement in fit to observational data — a reduced χ-*squared* of 1.73 (relative to 2.23 with the background fossil fuel emissions). Geographically, most improvement is shown in the Southern Hemisphere, where previously fit to data was most poor.

The realism of the predicted seasonal cycles for the five geographical regions, however, is low. For the worst result (North and Central America), the amplitude of seasonality is near 100 $\times$ that suggested in some studies, and peaks during the summer months as opposed to winter. Prediction that the anthropogenic CO_2 emissions for three of the five regions near zero in winter months is equally unrealistic.

Better estimates for the seasonality parameters are necessary due to the sensitivity of the calibration process to their initial values. Research currently under way will hopefully produce better estimates for some of the highest emitting countries, which may be sufficient to better constrain the calibration. Uncertainty regarding the parameters associated with seasonality in emissions are high. It is recommended that limits be imposed on the tuneable parameters so that the predicted seasonality is not compensating for large changes in other seasonal fluxes that occur during calibration.

Furthermore, the seasonality function is deemed over-simplistic. Suggestions are made relating to a superposition of a number of sinusoidal terms which should introduce more flexibility into the calibration — allowing for more variation in cycle within a region, and increasing the realism of the emissions in accordance with other studies.

1 Introduction

In 1896, the Nobel Prize-winning chemist Svanté Arrhenious proposed that the burning of fossil fuels, and the resulting release of carbon dioxide into the atmosphere, was causing a change in the transparency in the atmosphere to radiation, relating to J. B. Fourier's greenhouse theory introduced at the beginning of the 18th century [27]. He further proposed that said change may result in warming *outside previous human experience* [27].

Arrhenious' nonchalant comment took just short of a century to reach our (and politicians') ears. Now his concept has been developed into a theory widely ascribed to, attributing temperature change within the Earth system to the *radiative forcing* of certain greenhouse gases (GHGs) [23].

It is estimated that a net anthropogenic change in radiative forcing of $+1.6\mathrm{Wm}^{-2}$ (+0.6 to +2.4) has occurred over the last century [24, 23], leading to a temperature change of between 0.3 - 0.6°C [27].

The most recent IPCC assessment report (published in 2007), states in its synthesis that *"Most of the observed increase in global average temperatures since the mid 20th Century is* very likely *due to the observed increase in anthropogenic GHG [greenhouse gas] concentrations."* Global GHG emissions have risen by 70% between 1970 and 2004; CO_2, has increased by 80% within that time period [23].

CO_2 is the second most abundant GHG (to water vapour). It currently contributes to over 50% of the net anthropogenic change in positive radiative forcing [23]. It is also a long-lifetime GHG — much of the CO_2 from fossil fuel emissions will remain in the atmosphere for a significant period of time. According to Archer (2005), 17-33% of CO_2 emitted now will dominate climate effects over the next 1,000 years, and up to 7% will still be in the atmosphere 100,000 years from now [3].

Therefore, it is cumulative, not annual emissions which determine the ultimate equilibrium atmospheric concentration. The IPCC and the Kyoto protocol both cite future CO_2 emissions projections with reference to *stabilisation scenarios* which attempt to quantify emissions cuts necessary for atmospheric concentrations to arrive at various levels.

Increasing emissions and hence increasing atmospheric CO_2 concentrations have been earmarked with the potential to cause a change in climate extraneous to normal variations. Concerns regarding these digressions are highly policy relevant. Gurney (2007) states *"our quantitative knowledge about these [CO_2] emissions is insufficient to satisfy current scientific and policy needs"* [17]. It is necessary to have the scientific knowledge to back up policy decisions so that future protocol may be increasingly effective.

Further research is necessary to ratify gaps in knowledge regarding the global carbon cycle — how emissions are effecting it currently, and how they may effect it in the future. Effective mitigation requires

process-based and quantitate data to back up policy decisions. Quantification and comprehension of the oceanic and terrestrial CO_2 sinks is one of the most important and policy-relevant areas of scientific research at this time [45]. Quantification of anthropogenic CO_2 emissions are tantamount to understanding the future evolution of the global carbon cycle, and the determinants of such emissions necessary for verification of policy adherence.

2 Background

2.1 The Global Carbon Cycle

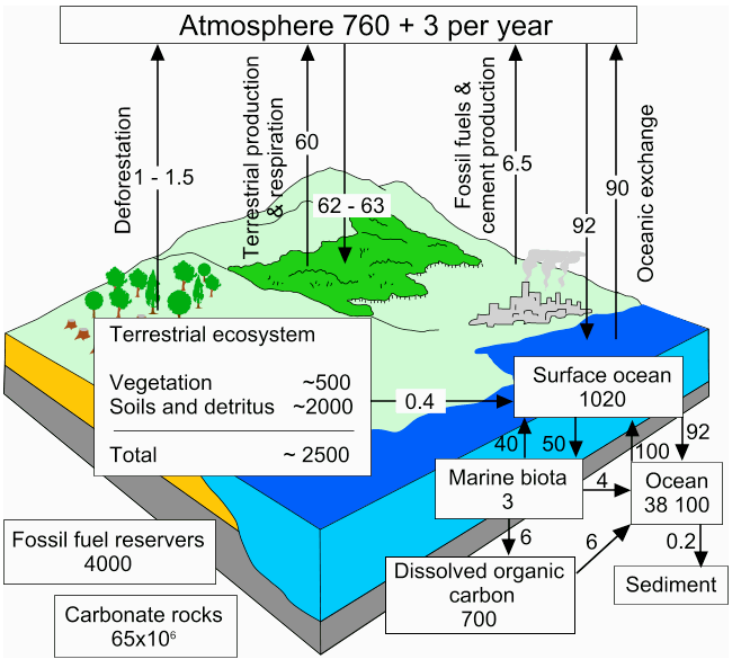


Figure 1: Global Carbon Cycle showing net carbon fluxes (values accompanying arrows in $GtCyr^{-1}$ and carbon reservoirs (boxed values in GtC). Values are in accordance with those in IPCC AR4 and represent current best estimates. [40]

Previous to industrialisation (circa 1750), atmospheric CO_2 concentrations had remained between ~180ppm (glacial) ~300ppm (inter-glacial) for 650kyr (six glacial-interglacial cycles) [22]. In the more recent past (the 10kyr previous to 1750), the global atmospheric concentration fluctuated only within a ~260 and 280ppm range [22]. The stability of the level of CO_2 in the atmosphere is governed by the global carbon cycle — a complex interplay of transference and recycling of carbon between three main components - the atmosphere,

the ocean and the terrestrial biosphere [58]. Fig. 1 succinctly depicts the main processes of transference which work on a variety of time scales.

The status quo in atmospheric CO_2 concentrations was broken by the advent of the industrial era. Fossil fuel combustion, deforestation, cement production, land use-change and biomass burning all contributed to a much larger net source of CO_2, and consequently, atmospheric concentrations rose. Today, the concentration sits at 379ppm — almost 100ppm higher than seen previously in the geological record [23]. The burning of fossil fuels is the largest contributor to elevated CO_2 levels [27, 22, 3].

Even under elevated atmospheric concentrations, however, the ocean and terrestrial biosphere still serve to remove almost 60% of the carbon released into the atmosphere [52]. The mechanisms driving the terrestrial and oceanic sinks, and their evolution into the future are currently poorly understood, with large uncertainties relating to the many of the processes involved [23, 22].

2.2 Carbon Cycle Modelling

Traditionally, process-based modelling (bottom-up) and inverse modelling (top-down) techniques have been applied independently to probe questions relating to Earth systems and the two communities have worked independently [46]. A typical global carbon cycle model may be described by a set of equations including a number of parameters (and uncertainties belonging to these parameters); in the case of inverse modelling — *tuneable* parameters, whose values are known only to within a certain degree of accuracy. Inverse models aim at reducing uncertainty associated with these parameters. Calibration of these parameters is usually guided by intuition rather than mathematical computation [28].

For the bottom-up approach, validity of the model is determined by a post-run comparison of results with observations, or a cross-model comparison [59]. The latter is commonly used to infer information relating to the accuracy of the model, and the calculation of errors associated with it [4].

Until recently, combining the methods has resulted in model construction and run-times far in excess of the immediate necessity (and often policy-relevancy) of the results themselves [28]. Data assimilation without the use of adjoint or tangent linear inverse models is usually highly inefficient and/or inaccurate. More recently, computational and technical advances, in specific the development of FastOpt's AD tool 'Transformation of Algorithms in FORTRAN' (**TAF**), has enabled much cheaper and more efficient data assimilation techniques [28].

2.3 CCDAS

CCDAS (Carbon Cycle Data Assimilation System) combines the inverse approach with a CO_2 transport model *and* a process-based model. The use of TAF allows for a model of this complexity to run efficiently. Within the main mandate of quantifying and understanding underlying processes of the global carbon cycle, one objective is to quantify terrestrial (or *Kyoto*) sinks, and to make suggestions regarding how it is possible to optimise these sinks with a view to reducing emissions.

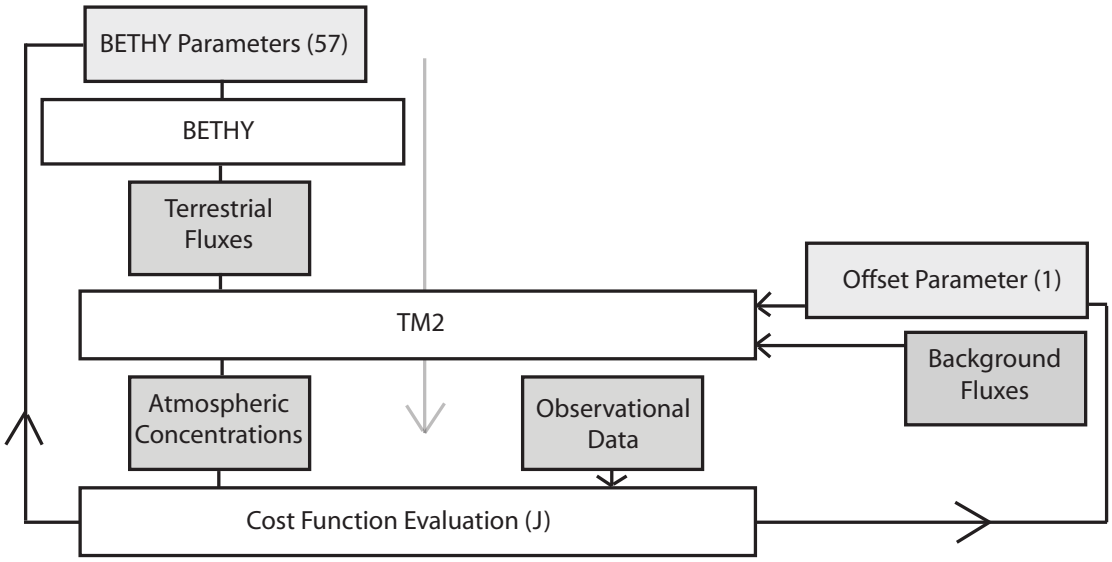


Figure 2: CCDAS set up for calibration. Lightly shaded boxes indicate tuneable parameters (57 associated with the terrestrial biosphere model, BETHY; 1 offset parameter for base atmospheric concentration); dark shaded boxes indicate CO_2 fluxes and atmospheric concentrations. White boxes indicate mapping between quantities, and the downward arrow indicates overall model flow. The final calculation of J directs the next iteration in the calibration (changes the parameters towards a better fit to observations). Once a minimum in$J$ is found, calibration is stopped. The background fluxes consist of fossil fuel emissions, CO_2 flux due to land use change and oceanic carbon transference. All parameters are listed in Table 1

It consists of biosphere model (BETHY) and an atmospheric transport model (TM2), together with prescribed background CO_2 fluxes representing ocean flux, land use change and fossil fuel emissions (as in Rayner (2005) [48]; oceanic flux has magnitude and distribution from Takahashi et al. (1999) and inter-annual variability from Le Qu*é*r*é* (2003); flux from land use change is from Houghton (2003) [56, 34, 21])
.

Within the model are a number of free parameters, which are tuned via data assimilation in the calibration mode of the model, using TAF [30, 28]. Fig. 2 depicts the model flow in calibration mode. Uncertainties for the parameters are calculated combining observational, model and data uncertainties. Initial values of the

parameters are based on current scientific knowledge, and the calibration seeks optimised values, attempting to minimise the margin of the associated uncertainty. The model may also be run in a forward (or prognostic) mode, using calibrated parameters to predict additional qualities of the system [28]. This mode is not used within this study and hence is not detailed here, however forecasting is kept in mind when making decisions relating to modelling preferences. The following subsections describe briefly the two main constituents of CCDAS (the carbon cycle model, and the surrounding data assimilation system), with a final focus on the anthropogenic CO_2 emissions and their current role within the system.

2.3.1 BETHY and TM2

The "Biosphere-Energy-Transportation-Hydrology" model (BETHY), the core of CCDAS, is a process-based model of the terrestrial biosphere [52, 33]. It works on a $2^\circ \times 2^\circ$ grid, and classifies vegetation into 13 plant functional types (PFTs) [28]. Each grid cell may be allocated up to 3 different PFTs; the relative amount of each specified by a fractional weighting [52]. At each grid cell, carbon assimilation via photosynthesis is calculated, dependent simultaneously on a balance of light, heat, soil water and nitrogen [33]. The respiration from both soil and plant are calculated via energy and water balance, and a phenology scheme [52]. It simulates the diurnal cycle of CO_2 assimilation and respiration, resolved at an hourly time step [28]. BETHY is a fully prognostic model, currently driven by climate and radiation data between 1979-2003. Furthermore, it is possible to forecast the future nature of the terrestrial biosphere under prescribed climate situations [52].

In the current version of BETHY, atmospheric CO_2, as input into the model is kept constant. Both direct and indirect effects of increased atmospheric CO_2 levels on the terrestrial biosphere (a *fertilisation effect*) are poorly constrained by current research [24], and thus difficult to include adequately in the model.

The CO_2 fluxes with the atmosphere are computed by BETHY and fed into the atmospheric transport model, TM2 via a mapping routine onto a (roughly) $8^\circ \times 10^\circ$ resolution grid [20]. For a passive tracer such as CO_2, TM2 acts as a linear function. It features vertical transport via convection and turbulent eddy mixing, and latitudinal/longitudinal transporting via analysed winds from the European Centre for Medium-Range Weather Forecasts (ECMWF) [48] (in this case, from 1986 [52]). This single year meteorological driving data means that often inter-annual variations in transport may appear as variations in the sources, and can have an adverse effect on inter-annually dependent parameters.

2.3.2 Data Assimilation

During calibration mode, CCDAS seeks an optimal parameter set, $\mathbf{x}_p$, for which the difference between model simulation, M, and observations are at a minimum. It employs the *cost function* (J in Eq. 1), to evaluate the fit of a model run to observational data (atmospheric CO_2 concentration data from GLOBALVIEW network [16] — $\mathbf{d}$). The derivative provides information regarding the gradient of J with respect to the parameters, guiding optimisation [48]. $\mathbf{C_d}$ and $\mathbf{C_p}$ are covariance matrices expressing the uncertainty in observations, $\mathbf{d}$, and the priors, $\mathbf{p}$ respectively (Eqs. 2 and 3) [28, 48, 30].$M(\mathbf{x})$ is the model output (atmospheric CO_2 concentrations calculated via the model using parameter set $\mathbf{x}$) and ^{T} denotes a transpose. Using an iterative loop, the model tunes the parameter set, starting from the initial input values, $\mathbf{x}_0$, moving through various $\mathbf{x}$, to reach $\mathbf{x}_p$ — the predicted (or *optimal*) parameter set.

$$J(\mathbf{x}) = \frac{1}{2}\left((M(\mathbf{x}) - \mathbf{d})^T {\mathbf{C_d}}^{-1}(M(\mathbf{x}) - \mathbf{d}) + (\mathbf{x} - \mathbf{p})^T {\mathbf{C_p}}^{-1}(\mathbf{x} - \mathbf{p})\right) \tag{1}$$

$$\mathbf{C_d} = \left(\frac{\partial^2 J(\mathbf{x})}{\partial \mathbf{x}^2}\right)^{-1} = \mathbf{H}(\mathbf{x})^{-1} \tag{2}$$

$$\mathbf{C_p} = \left(\frac{\partial M(\mathbf{x})}{\partial \mathbf{x}}\right)\mathbf{C_d}\left(\frac{\partial M(\mathbf{x})}{\partial \mathbf{x}}\right)^T \tag{3}$$

The approach is Baysian. Both observations ($\mathbf{d}$) and priors ($\mathbf{p}$) are assumed to have Gaussian probability distributions [28]. The posterior uncertainties on the tuned parameters ($\mathbf{x}$ from the final iteration — or $\mathbf{x}_p$, with a minimum in J) is approximated by the inverse Hessian of J at this point [28, 48].

Finding a minimum in J is non-trivial due to the non-linear nature of the model. The function may contain many local minima, sharp changes in gradient, troughs, peaks and shear edges. It is therefore important to make the best possible estimates for the initial values of the tuneable parameters ($\mathbf{x}_0$) to guide the optimisation towards a physically meaningful minimum.

2.3.3 Current Fossil Fuel Emissions in CCDAS

Currently, fossil fuel CO_2 emissions are designated a prescribed background flux in combination with ocean flux and contributions from land use change (see Fig. 2) [48, 52].

The annually invariant flux magnitudes are from Marland et al. (2006), resulting in step increases between years from 1979 and 2003 [38]. Two spatial distributions of population density are used to weight

the emissions on a national level. For the years prior to 1991, 1990 density distribution (from Andres et al. (1996) [2]) is used, and for the years after 1994, 1995 distribution (from Brenkert (1998) [9]) is used (i.e. the distribution is static). The years in between are a linear interpolation of the two [52]. Quantification of the effect of this temporally changing distribution is not cited in the literature. It is unclear to what extent this minimal interpolation (5 years out of a usual 21 year calibration) corrects the distribution. Static emissions distribution is cited as the cause of poor fit to observations in a number of models, however attempts to quantify this are few [48, 17, 15].

3 Theory

3.1 Fossil Fuel Emissions

The main shortcoming of the data sets currently used in CCDAS (and many other carbon cycle models) is the lack of seasonality [48]. Inappropriate treatment of a seasonal cycle within emissions may lead to spatial and temporal errors in results [44, 18].

$$C_{obs} = C_{ff} + C_{ot} + \sum_{i=1}^{N} C_{res} \tag{4}$$

Inversion studies employ existing atmospheric CO_2 and fossil fuel flux data to infer residual carbon exchange with the terrestrial biosphere and the oceans [18]. Fossil fuel CO_2 emissions are prescribed as a *background* or *pre-subtracted* flux [18, 4, 26, 8, 11], as in Eq. 4, where C_{obs} is the observational atmospheric CO_2 data, C_{ff} is fossil fuel emission data, C_{ot} represent other background fluxes that may be considered and C_{res} is the contribution to residual (biosphere and oceanic) fluxes from N discrete regions [18]. This latter flux is calculated via a final step of subtracting the background fluxes - C_{ff} and C_{ot} from the observational data.

Typically, fossil fuel emissions are considered well quantified (in comparison to biotic fluxes) and hence are provided with small uncertainty. Despite regionally different intra-annual variability, it is often deemed unnecessary to include since its rate of growth (or depletion in recent years in some areas) is much less than the inter-annual variability of the atmospheric CO_2 [46].

There are two main problems with this approach. First of all, often, the fossil fuel emissions are much greater than the net terrestrial sinks and so when subtracted, the final result is very sensitive to errors in the fossil fuel emissions [26]. As both temporal and spatial resolutions of the inversion models increases, more

and more bias is inherent in the process. Gurney (2005) used a number of atmospheric transport models to compare flux estimates with and without the inclusion of seasonal variation in fossil fuel emissions. In many cases the introduction of seasonality lead to a bias to the order of 50% on the residual fluxes (C_{res}) in the model [18]. Furthermore, introducing a peak-to-peak seasonal variation of roughly 30% to mid-Northern latitudes incurred a difference in annual flux of up to 0.5 GtC. This suggests that the inversion as a means of inferring the residual fluxes, on a smaller than annual temporal scale, seasonal variations of fossil fuel emissions are important.

Although attached to a process-based terrestrial biosphere model, it is likely that calibration in CCDAS will be similarly sensitive to seasonality in fossil fuel emissions. It is the information regarding correct reconstruction of the atmospheric concentrations (including seasonality) that directs calibration and tunes parameters.

Seasonality in fossil fuel emissions are often omitted from inverse (or data assimilation) models because of the large uncertainties attached to the phase, amplitude and progression over time of the cycle. Few studies have managed to quantify these important factors, especially on a global basis. Rotty et al. (1987) assumed the driving factors behind these seasonalities in emissions are likely to lie mainly in energy demand through heating (dependent on seasonal variations in climate and weather) [50]. Studies on areas in Northern European CO_2 emissions display significant winter highs with a mean seasonal peak-to-peak of 1.4ppm (2 $\times$ that of Rotty et al.), contributing to just over 10% of the seasonal cycle in atmospheric concentration in some areas [35].

Blasing (2005) showed that in the United States, there are two main overriding cycles in emissions. Coal, used for heating, produces emission peaking in the winter months, in agreement with Rotty et al, although with a substantially higher mean peak-to-peak of $\sim$0.035GtC. Natural gas however displayed two peaks — one attributed to electricity usage through indoor activities in winter, and another in the summer due to usage the of air conditioning with an average peak-to-peak of $\sim$0.02GtC [6]. Furthermore, over time it is suggested that this seasonality decreases. Summer emissions displayed more rapid increase than winter emissions, suggesting a change in the annual pattern of energy use [6]. Other studies in energy use have showed similar results. In hotter climes such as Spain, evidence shows that energy consumption has changed phase. Whereas previously highs in winter months dominated the seasonal cycle, for the past decade or so, summer peaks in consumption are far in excess of those in winter [44]. Similarly, in the US, demand depends greatly on the geographic location of the State — energy use profiles from states further south show more dependence on high summer temperatures than low winter ones [51].

On other spatial and temporal scales, further demand drivers may need to be considered [44, 18, 29, 26]. Since CCDAS works on a monthly scale, these need not be considered. However, the case for inclusion of seasonality in emissions is clear. Without intra-annual variation in emissions, errors in the quantification of residual fluxes may be large. It is also hoped that through inclusion in CCDAS, some of the unknowns in the seasonal trends of fossil fuel emissions may be constrained.

3.2 Models and Approaches

Research with regards to the carbon cycle has thus far tried to encompass the anthropogenic inputs in ways which vary substantially in depth and realism. The following subsections describe various approaches to modelling anthropogenic CO_2 emissions. Over the last decade, largely due to vast and varied data sets being made accessible, bottom-up CO_2 emissions modelling has been made possible. There are two main approaches - one which identifies the direct sources of anthropogenic CO_2 (largely based around fossil fuel combustion) and one which concentrates on the drivers behind the emissions (a socio-economic treatment).

3.2.1 Energy Decomposition

There are two main approaches to energy decomposition techniques [25]. The *reference approach* concentrates on the *supply* side of energy data, accounting for carbon mainly on the supply of primary fuels and net quantities of secondary fuels brought into the economy (trade) [59]. The *sectorial approach* focuses on the *consumption* of energy - a sectorial breakdown of aggregate CO_2 emissions by the amount of each type of fuel consumed by each sector of the country [59].

Reference Approach. The breakdown of emissions is by fuel type. Raupach (2007) broke down these contributors into seven sources, as in Eq. 5: combustion of solid, liquid and gaseous fuels (F_{solid}, F_{liquid} and F_{gas} respectively); flaring of gas from wells and industrial processes (F_{flare}); cement production (F_{cement}); oxidation of non-fuel hydrocarbons ($F_{non-fuelHC}$) and fuel from 'international bunkers' used for shipping and air transport ($F_{bunkers}$) [47]. These vary on a national basis as well as temporally [39, 7]

$$F = F_{solid} + F_{liquid} + F_{gas} + F_{flare} + F_{cement} + F_{non-fuelHC} + F_{bunkers} \quad (5)$$

$$Q_i = \text{consumption}_i = \text{production}_i + \text{imports}_i - \text{exports}_i - \text{bunkers}_i - \text{changes in stocks}_i \quad (6)$$

$$F_i = Q_i \times FO_i \times C_i \tag{7}$$

For each of these emissions-producers, i, the consumption nationally is calculated as in Eq. 6, and then converted into CO_2 emissions via Eq. 7. Q represents the quantity of fuel i (i.e. the 'consumption' from Eq. 6), FO the fraction oxidised and C the carbon content [1, 7, 39]. On a national level, this allows for a fairly accurate database of emissions to be built up. Andres et al. (1999) quote an average error of 8% in their final database, although this error may range anywhere between -340% and 90% given a particular year and country [1].

The main source of error in this type of analysis comes from omissions in data sets, and changes in boundaries and geographical definitions of countries. Often older trades inventories for fossil fuel data are grouped together, limiting the accuracy of calculation in Eq. 7. Furthermore, any higher resolution of data (spatially or temporally) must be done by weighting. Frequently seasonality is excluded for this reason (data is not available due to legalities in collection) [18]. Positively however, this approach allows for global coverage over relatively long periods of time, due to the available data [37]. It also makes emissions proxies easier to detect (on a national level) due to data being available for countries in a diverse range of economic and social situations (more is discussed on this topic in the socio-economic treatment). BP quote that coal is the fastest growing energy source over the globe, and that coal usage can in fact dictate the % contribution of a country to global CO_2 emissions [10].

Sectorial Approach. The sectorial approach aims to provide highly resolved estimates for CO_2 emissions that come from a process-based (electricity generation, residential heating, vehicle propulsion etc.) model [17]. It has thus far only been applied to individual, and predominantly developed, nations [17, 15, 59]. Globally, freedom and quality of data are the main restrictions to this approach [17].

An example is the **Vulcan** project (and soon **Hestia**, which will build on the findings of **Vulcan**) — a model for the USA's CO_2 emissions resolved by 10km and an hourly time step [19]. CO_2 emissions are grouped by types of source - *point sources* which encompasses stationary sources such as power stations; *area sources* such as office buildings and wildfires; and *mobile sources* relating to vehicular emissions.

Another approach is to deconstruct emissions into economic sectors such as public power; refineries, extraction and distribution of fossil fuels; industry; road traffic; waste and agriculture. These sectors are then further sub-divided, providing a process-based approach to the drivers. As an example, the *public power* sector's contribution to emissions is dictated by energy demand, which in turn may be related to

climate-drivers. Temperature, according to Sailor (2001), accounts for 80% of the variance in electricity consumption [51].

Uncertainty is introduced to the model through simplification on other levels. For example, homogeneity in demand may be applied across whole countries. Other drawbacks again are necessity for large amounts of data (from household energy bills to meteorological data), and the time necessary for models to run [19]. If this technique were applied worldwide then the time for such a model to run would be costly.

3.2.2 Socio-economic Treatment

Studies which attempt to reconstruct CO_2 emissions through economic and social proxies attempt to find correlation and trends in national statistics which reflect those in emissions data [41].

Socio-economic models tend to work on the same scales as inversion processes — annual and national, though they may forecast tens of years into the future [32, 36]. This is the main advantage over all previously detailed approaches. Economic theory and forecasting is well-established [23]. Global economics follow relatively predictable trends, which, if linked to emissions, may be exploited in estimations of how CO_2 emissions trends may alter in the future. Integrated Assessment Models (IAMs) combine the economic an scientific drivers of climate change, with depth in both areas associated to the purpose of the study - usually policy evaluation or optimisation [32]. Often uncertainty with prediction is only assessed crudely [32, 57]. The uncertainty attached to such predictions are only as good as the modeller's ability to predict decisions [32]. The largest uncertainty with regards to a socio-economic approach lies in societies' preferences. Once issues of social choice are restricted (or reduced) through policy and co-operation, the uncertainty in prediction decreases [55].

EKC. The *Environmental Kuznets Curve* (EKC) describes a hypothesised situation where as GDP increases, pollutants increase almost linearly, *until* a certain point, where greater income dictates that emissions hit a maximum and will thereafter decrease with increasing wealth (the *inverted U* - see Fig. 3) [12]. At a certain level of economic development, more efficient infrastructure and pollution controls are afforded. There is little doubt that both CO_2 emissions and energy consumption are correlated with the size of a country's economy, but since 1960 this correlation has been weakening [49, 55]. Research roughly a decade ago into a number of pollutants and their correlation to GDP stated that this 'turning point' tended to occur before countries reached a GDP of US$8,000 [49].

Turning points of per capita emissions of CO_2 in particular generally fall well outside the range of other

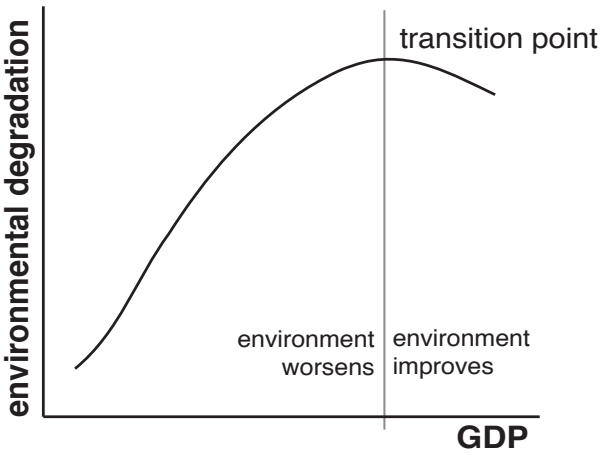


Figure 3: Example of the 'inverted U' relationship between environmental degradation (in this case, increasing CO_2 emissions) and GDP in the *Environmental Kuznets Curve*, featuring the *transition point* from environmental destruction to improvement.

pollutants [12]. The uncertainty related to the turning point is high and results in little confidence in any estimations made [12]. Cases in which the turning point is well defined with relatively little uncertainty tend to relate to pollutants which have a localised effect. If a pollutant has a more global or indirect effect, a downward trend in emissions corresponding with increased GDP is not necessarily instigated [12].

Using the same correlation, monotonic relationships between CO_2 emissions and economic growth have also been found [49], demonstrating that results are highly dependent on the samples and data used in the study (i.e. which country is tested) [54]. When separated into low, middle and high income countries, it is found that a relatively small number of wealthy countries (with the highest emissions) dictate the *global* trend by becoming more efficient, while the rest of the countries worsen [49]. Few countries have moved up development stages with regards to efficiency, despite economic growth.

Furthermore, there is no reason to believe that most countries will reach the transition point [49]. Stern (2003) argues that one possible reason for this is because of the nature of a free trade economy. Economic growth resulting in migration from 'developing' to 'developed' incites a transference from capitol gains through labour and natural resources to human and manufactured capitol - in effect, an 'outsourcing' of emissions intensive industry. Once a significant number of countries have made this transition, there are no longer any countries to outsource to - the trend is finite [54]. In fact, trade is not discussed in the majority of literature on the EKC and often may result in spatially transferred emissions, rather than reductions [54, 55]. This also means that it is not viable to model developing countries' futures on historical evidence

from the development of the wealthier countries.

IPAT and Kaya. The IPAT identity in Eq. 8 cites three constituent *drivers* of the change in emissions - population, affluence (income per capita), and technology (emissions per unit income) [41]. When discussing CO_2 emissions in specific, this relationship is often referred to as the Kaya identity (see Eq. 9), and is frequently summed over *regions* of the globe with similar emissions and economic profiles, or countries, i [47, 55]. Here, P is the population, G is GDP (Gross Domestic Product) and E is energy consumption. The emissions (F), are defined as amount of carbon per unit of fossil energy [47].

$$\text{impact} = \text{population} \times \text{affluence} \times \text{technology} \tag{8}$$

$$\text{CO}_2 \text{ emissions}_i = P_i \left(\frac{G_i}{P_i}\right)\left(\frac{E_i}{G_i}\right)\left(\frac{F_i}{E_i}\right) \tag{9}$$

Using these simple relationships, it is possible to factorise the growth of emissions - to *weight* the dependence of an impact on each. For example, a 1.7% increase in CO_2 emissions per year (on average) since 1850 may be factorised as follows:

$$\begin{aligned} 1.7\%\ \text{CO}_2 \text{ increase} &= 3\% \text{ growth in GDP} \\ &- 1\% \text{ decline in energy intensity of GDP} \\ &- 0.3\% \text{ decline in carbon intensity of the primary energy} \end{aligned} \tag{10}$$

The growth in GDP may be factorised further:

$$3\% \text{ growth in GDP} = 1\% \text{ growth in population} + 2\% \text{ growth in per capita income} \tag{11}$$

Unfortunately, growth in emissions is not linear over time [23]. The drivers in global emissions after 2000 shows a reversal in the long-term trends with global carbon intensity (CO_2 emissions per unit of energy, E) and energy intensity (unit of energy, E per unit of GDP, G) *decreasing* with increasing emissions [47]. Distribution of emissions has higher dependence on different factors for different regions - F_i and E_i must be weighted towards developed regions and P_i must be weighted towards developing regions in order to correct for this [47]. Stern also concluded that population growth and GDP were the most important proxies for

developing countries and energy intensity reductions were important in *transition* (fast developing) economies such as China [55, 59]. The greatest uncertainty in driver quantification and forecasting is in developing regions, whose transition is difficult to predict.

However, no single driver is fundamental or independent [41]. For example, population growth depends on fertility and death rates, which in turn depend upon education, income, social norms and health provisions. All of these aspects can be cited as contributors to the long-term productivity, economic growth and structure and technological change within a country. The main drivers behind emissions growth depend upon shared factors [41]. This results in large uncertainties which makes probabilistic scenarios impossible with current knowledge [42].

Edmonds and Joos (2004) recognise the importance of this technological advancement in stabilisation of CO_2 within the atmosphere. Not all driving forces are adequately encompassed within this process, which is why they do not determine emissions concisely [13, 43].

3.2.3 Evaluation

Unfortunately, socio-economic theory is too qualitative for this study. Poorly constrained factors driving CO_2 emissions may lead to indeterminacy — there may not be sufficient information in the atmospheric concentration data to constrain such uncertain and inter-dependent parameters. Furthermore, socio-economic factors do not exhibit intra-annual variations. Process-based modelling, whilst encompassing seasonality intrinsically, is likely to be too costly computationally, with too many factors requiring tuneable parameters, and too much globally unavailable data.

Information in observational data may not be sufficient to constrain more than a few parameters, and so the following fossil fuel CO_2 emissions model is constructed — concentrating on cost-effective running and minimal parameter addition.

4 Model

4.1 Base Level Emissions

In order to reduce the number of tuneable parameters, land masses are divided into five regions (See Fig. 4). Antarctica and the Arctic are excluded as they have zero emissions.

1. Europe and Northern Asia
2. Mid-North Africa, the Middle-East and Central and Southern Asia

Oceania

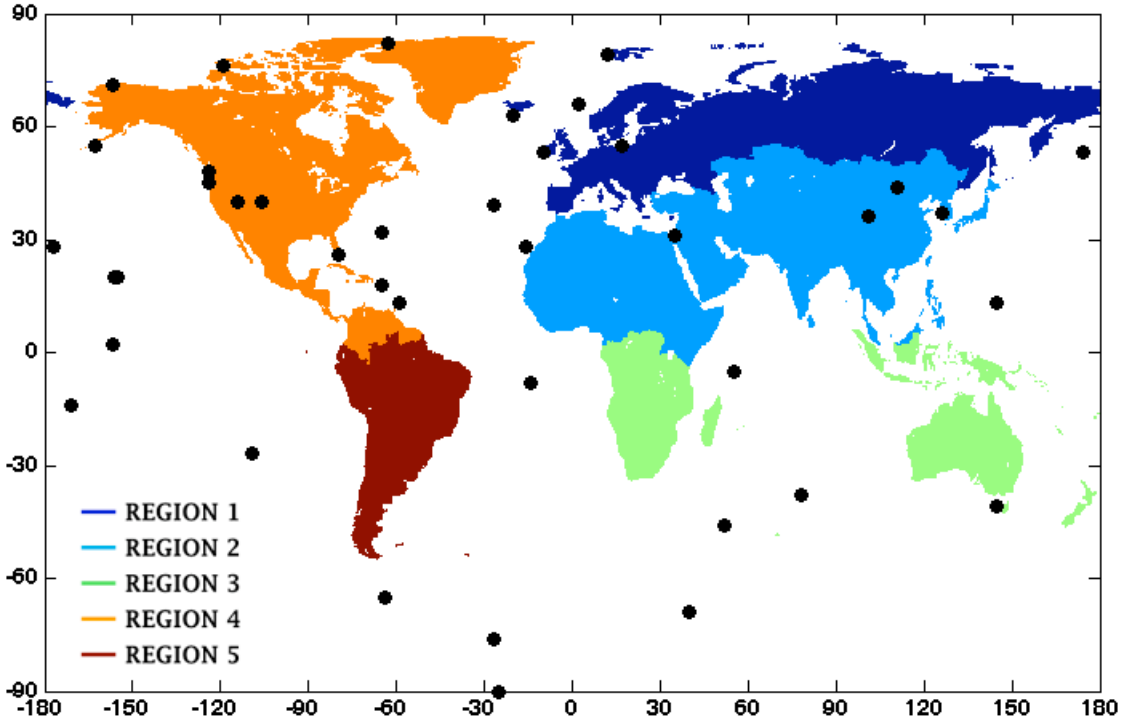


Figure 4: Geographical distribution of regions (0.5°× 0.5° resolution). Black circles indicate positioning of atmospheric CO_2 observation stations used in this study. Lists of observation stations and their locations may be found in Appendix A and the countries in each region in Appendix B.

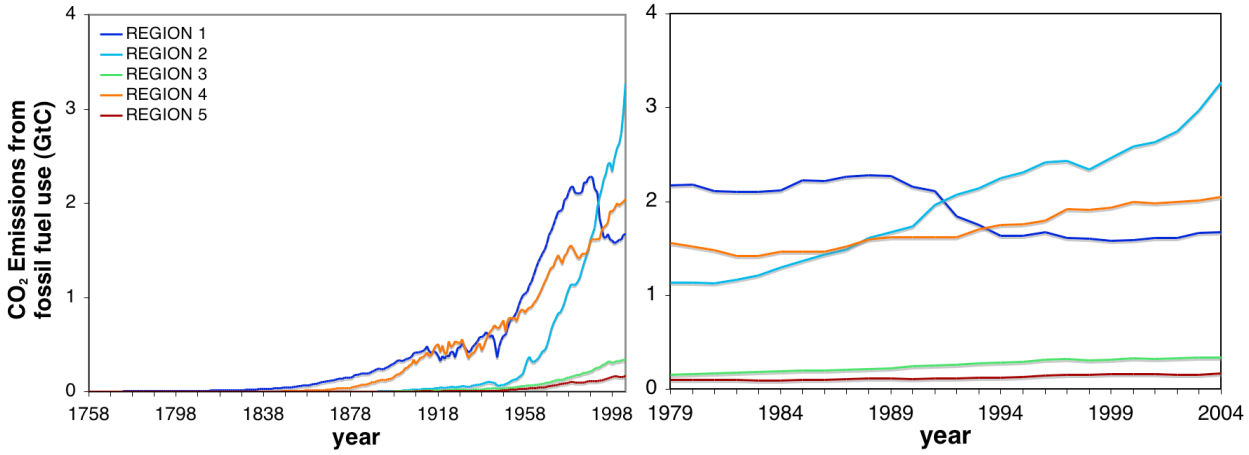


Figure 5: Total annual CO_2 emissions from fossil fuel use per region (F_n^r) as from Marland et al. (2006) [38]. Trends over the past ∼250 years (left); Data required for calibration period (1979 - 1999) (right).

3. Southern Africa, the tip of South-East Asia and Oceania
4. Central and Northern America
5. South America

Since seasonality depends on hemisphere (opposite winter and summer dictating energy consumption) [51, 18], regions do not cross the equator where possible. Regions also attempts to encapsulate countries with similar emissions, or emissions growth in order to reduce misallocation within the region by population weighting. This is not always possible.

Fossil fuel emissions data is taken from the **CDIAC** data set for the years 1978 to 2000 (from Marland et al. (2006) [38]. Annual emissions are totalled for each region. This grouping is a source of uncertainty, essentially taking a countries' emissions and reallocating them across a larger area. Annual emissions per region (F^r_n, where n denotes year, and r the region) are linearly interpolated so that monthly (m) estimates, F^r_m, may be made:

$$F^r_m = \left(\frac{F^r_{n+1} - F^r_n}{12} \right) \times m \tag{12}$$

$$F^0_{i,j,m} = F^r_m \times W^r_{i,j} \tag{13}$$

Data for the year before and the year after the calibration are required for interpolation. The annual average emission is taken to be the value for the centre of the year, and other values adjusted around it. The total annual flux is therefore not necessarily equal to the original value in the data set, and the largest errors are incurred on the first and last years of the calibration period. The greater the number of years the model is calibrated, the lower the error from this source. All emissions values are monthly, but in gC yr^{-1}.

The emissions are then multiplied by the population fraction weighting (as in Eq. 13) to provide the spatial distribution of base level emissions ($F^0_{i,j,m}$). Population data is from **SEDAC** (the Socio-economic Data and Applications Center), for the year 1995 on a $0.5^\circ \times 0.5^\circ$ resolution [5]. A grid cell is allocated to a region and population summed on a regional basis. Each grid cell is then given a *population fraction* ($W^r_{i,j}$, where i and j are the co-ordinates of the grid cell, and r is the region), corresponding to the fraction of the total population of the region in that grid cell. This data is then aggregated to TM2's low resolution grid — $\sim 8^\circ \times 10^\circ$. On these scales therefore, one grid cell may contain population relating to more than one region. Gridded population fraction representation may be found in Appendix C.

This is likely to be one of the largest sources of error temporally. The distribution is static throughout the calibration run. Spatially, as the resolution of the model decreases, the error in allocation of emissions

increases. Any single populated grid cell on the $0.5^\circ \times 0.5^\circ$ resolution will mean that the corresponding grid cell on the lower resolution will too be allocated emissions - they will in effect be distributed over much larger areas. In cases where there is very little actual land mass in the low resolution grid cell, this likely to be unimportant - the relative fraction of population is likely to be so low the emissions from this area will not significantly affect atmospheric concentration. In very densely populated coastal areas close to CO_2 observation stations, this is likely to be more of a problem.

4.2 Seasonality

A seasonal cycle taken from Gurney (2005), is imposed upon the base level emissions ($F^0_{i,j,m}$) [18]. The total emissions per grid cell per month, belonging to region r — $F^r_{i,j,m}$ — is then as in Eq. 14.

$$F^r_{i,j,m} = F^0_{i,j,m} + S_r F^0_{i,j,m} sin(\theta_j) cos\left(\frac{2\pi(m-1)}{12}\right) \tag{14}$$

S_r is the tuneable parameter to be optimised through CCDAS calibration, and determines the amplitude of the seasonal cycle. As a factor multiplied by the total base level emissions, S_r represents what fraction of those emissions are governed by seasonality. For example, if $S = 0.5$, then the seasonal cycle has a peak-to-peak equal to half of the total emissions (or an amplitude equal to quarter of the emissions). The $sin(\theta_j)$ term encompasses the latitudinal dependence (greater seasonality at higher latitudes, with opposite seasonal cycles for Northern and Southern hemispheres) and the $cos(...)$ term depicts variation dependent on the month, m.

Monthly flux in $GtCyr^{-1}$ are output from the model on a $\sim 8^\circ \times 10^\circ$ grid, where the total emission per grid cell is equal to the sum of emissions from each region contained within that grid cell. (A high resolution — $2^\circ \times 2^\circ$ — optional version is also provided, should TM2 be upgraded to the BETHY resolution).

4.3 Model Flow

The overall model flow within CCDAS is now as in Fig. 6. The fossil fuel emissions now run in parallel to BETHY figuratively. The fluxes from CO_2 emissions are added to those from terrestrial processes in BETHY and the remaining background fluxes on the resolution of the transport model (TM2). The total flux is then passed through TM2 and atmospheric CO_2 concentrations (in ppm) are compared to observations from GLOBALVIEW [16].

The cost function, J and its gradient with respect to the parameter set $\mathbf{x}$ (now totalling 62) is calculated

(as in Eq. 1) in an iterative loop, until a minimum in J is found. This is the calibration (or *optimisation*) process.

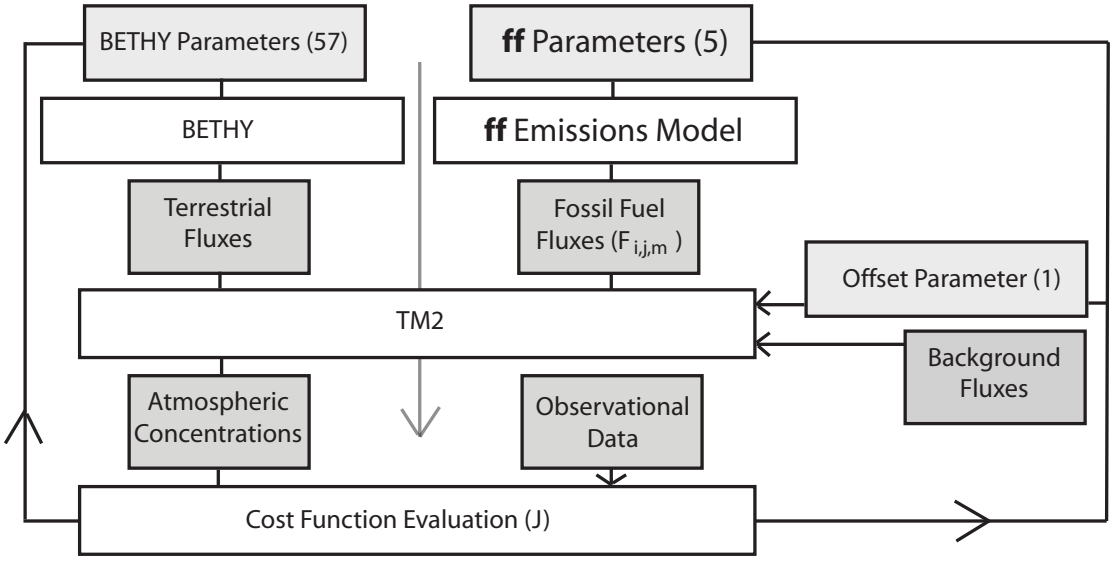


Figure 6: **ff** model flow diagram in calibration. BETHY and the fossil fuel emissions model now run in parallel. The total number of parameters (in lightly shaded boxes) to be optimised through calibration is now 62 — 1 representing base atmospheric concentration, 57 associated with processes in the terrestrial biosphere model, BETHY, and the additional 5 representative of the amplitude factor of the seasonal cycle of fossil fuel emissions (one for each region). Dark shaded boxes indicate CO_2 fluxes and atmospheric concentrations. White boxes indicate mapping between quantities, and the downward arrow indicates overall model flow. The final calculation of J directs the next iteration in the calibration (changes the parameters towards a better fit to observations). Once a minimum in$J$ is found, calibration is stopped. The background fluxes now consist only of CO_2 flux due to land use change and oceanic carbon transference. All parameters are listed in Table 1

5 Methods

Two versions of CCDAS are used in this study — one with fossil fuel emissions prescribed as a background flux — the *head version* (referred to as **hv** from this point onward), and the version described in Section 4 with an emissions model running parallel to BETHY (referred to as **ff**). The difference between the models' output is analysed and compared to gain insight into the effect of the introduction of seasonality in emissions.

5.1 Model Run — Calibration

S_r is given an initial value ($\mathbf{x}_0$) of 0.3 (30% seasonal variation of total emissions) for all regions. This follows suggested *realistic* seasonal variation according to Gurney (2005) [18]. The uncertainty on this (σ_i) is estimated at ± 0.5. The distribution therefore encompasses the higher level of seasonality cited by Gurney,

whilst the lower range encompasses the possibility of a switched seasonal cycle. S_r is unconstrained, meaning that there are no limits on what value it may take within the model. All other initial parameter values and associated uncertainties are kept as in previous studies of CCDAS for both the **hv** and **ff** model versions [48, 52].

Both models are calibrated on high resolution ($2° \times 2°$), for 21 years (1979 - 1999), with a 5 year spin-up phase [48, 52]. After calibration, the predicted parameters ($\mathbf{x}_p$) for both models are given, and their uncertainties (σ_p) provided by the square root of the diagonal elements of the inverse *de-noised* Hessian in Eq. 2.

5.2 Analysis

5.2.1 Fit to Data

An overall fit to observational data is dictated by the value of the cost function (J) — if the fossil fuel emissions are improved, the cost function (as in Eq. 1) is expected to be smaller. The quantification of the fit is made by calculating the χ-*squared*. This is equal to twice J divided by the number of observations.

$$C_{GLOB} = 0.25C_{SPO} + 0.75C_{MLO} \tag{15}$$

An estimate of fit to a *global* concentrations is made by calculation of C_{GLOB} as in Eq. 15, where C_{SPO} is the atmospheric CO_2 concentration at the South Pole and C_{MLO} is that at Mauna Loa. An emphasis is put on investigating the fit of the seasonality of the concentrations by removing the *moving average* of both data and model output from the overall totals ($\mathbf{d}$ and $M(\mathbf{x}_p)$ respectively). This process is detailed in Appendix D.1. An average seasonal cycle is constructed, and an R^2 fit of model output to observational data is calculated (R^2 fitting is detailed in Appendix D.2).

The method for concentration deconstruction for C_{GLOB} (removal of the running mean) is repeated individually for the MLO and SPO stations to assess the fossil fuel model's ability to reconstruct atmospheric concentrations with both high and low seasonality. The BAL (Baltic Sea) and TAP (Tae-Ahn Peninsula) observation stations' data and fit are also investigated — both are cited as stations whose concentrations are poorly reconstructed by CCDAS, thought to be due to a lack of seasonality in the fossil fuel emissions [48].

5.2.2 Parameters and Seasonality

The optimised (or *predicted*) parameters ($\mathbf{x}_p$), and their associated uncertainties (σ_p) are analysed to provide insight into error reduction within the model (provided by the formula 1 - σ_p/σ_i, where σ_i and σ_p are the uncertainties relating to the initial parameter set ($\mathbf{x}_0$) and predicted parameter set ($\mathbf{x}_p$), respectively), and any differing importance of various processes within the models. Since each optimised parameter is related to a process within CCDAS, information regarding any changing importance in a process may be found from a change in parameter value.

Time series of total CO_2 flux from each region are taken from model output (summing over all grid cells for a given month — $\sum_{i,j} F^r_{i,j,m}$). Seasonality at the upper and lower bounds given by the predicted uncertainties on the parameters are also investigated, by running the model with these values (i.e. $S_r = x_p \pm \sigma_p$). Mean monthly terrestrial and fossil fuel fluxes are calculated by summing the total flux over all grid cells for each month in turn, and then averaging over the number of years in the calibration period. This provides an overall impression of their seasonal cycles, and highlights and large changes between **hv** and **ff** models.

5.2.3 Fluxes

Global Basic comparison of the two models' mean annual, and mean monthly terrestrial and fossil fuel CO_2 flux is made to assess any large changes through the addition of the fossil fuel model. This is done by summing the flux over all grid cells over all months within a given year (or month). These annual totals are then averaged over the model calibration period. The standard deviation provides information relating to the inter-annual changes in each flux. Mean Northern and Southern Hemispheric fluxes are also studied.

For the fossil fuel emissions in **ff**, the mean annual flux from each region is also calculated and checked against original totals from the original CDIAC data set.

Spatial Distribution Mean monthly and annual distributions of both terrestrial flux and fossil fuel emissions are constructed, and the difference between them studied. The areas, (and times of year) with the largest difference is the focus, with a view to highlighting distribution changes in seasonality and mean annual flux between **hv** and **ff**.

Fit to observations of atmospheric CO_2 concentrations at stations located near large perturbations in mean seasonal or annual flux are calculated using the R^2 technique. An estimation of improvement in fit between **hv** and **ff** models is given via Γ; a percentage expression of the change in R^2 towards a prefect fit to observations (calculation details in Appendix D.3).

A range of stations distributed globally are also chosen for this analysis, to obtain a more general representation of the effects of introducing a seasonal cycle in fossil fuel emissions. Patterns latitudinally or regionally are investigated as an assessment of whether there are trends in the improvement of model fit to observational data.

6 Results

6.1 Fit to Data

6.1.1 Global

After calibration, the values of J are 7749 and 6024 for the **hv** and **ff** model versions respectively. These translate into χ-*squared* values of 2.23 and 1.73, meaning that the introduction of seasonality in the fossil fuel emissions has improved the fit to observational data. Relative to a base of spatially and intra-annually invariant background fossil fuel emissions, the introduction of seasonality improves the fit by 37.3%, whereas the introduction of a (limited) temporally changing emissions source improves the fit by only 22.4%.

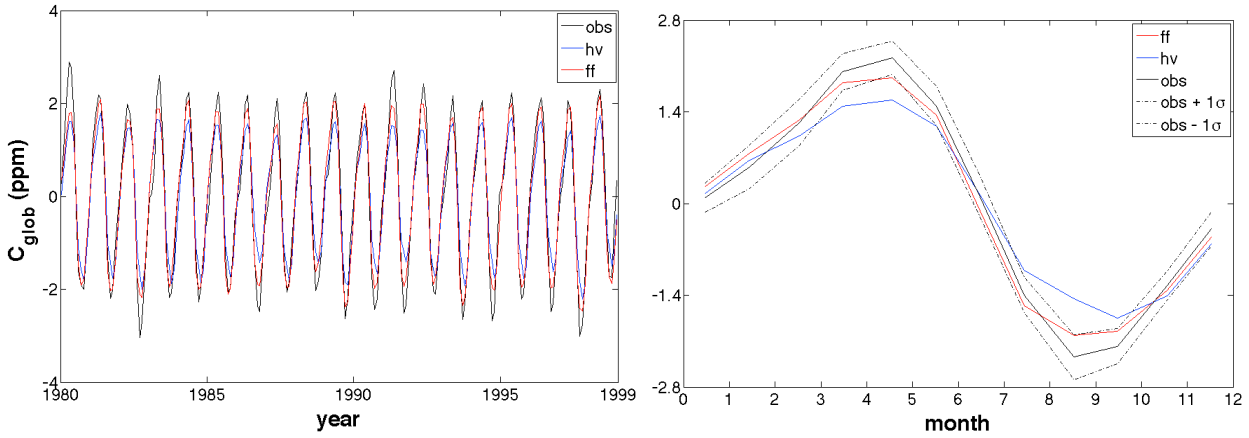


Figure 7: **hv** (red), **ff** (blue), observations (black). Left: C_{GLOB} seasonal cycle over calibration period (inter-annual changes removed via moving average). Mean seasonal cycle (right).

Seasonality in C_{GLOB} has been reproduced more accurately in the **ff** model. Fig. 7 shows that the mean seasonal cycle in atmospheric CO_2 concentration of the model output now falls within 1σ of the mean of the observations. The R^2 of the overall fit is now 0.91 in comparison to the **hv** fit of 0.75. The seasonal fit has improved from 0.91 to 0.98. The slight lag in the seasonal cycle has also been corrected somewhat, with both high and low amplitude peaks corresponding temporally with those in the observations (in the **hv**

model output, the low amplitude peak was delayed by ∼2 months).

6.1.2 Individual Stations

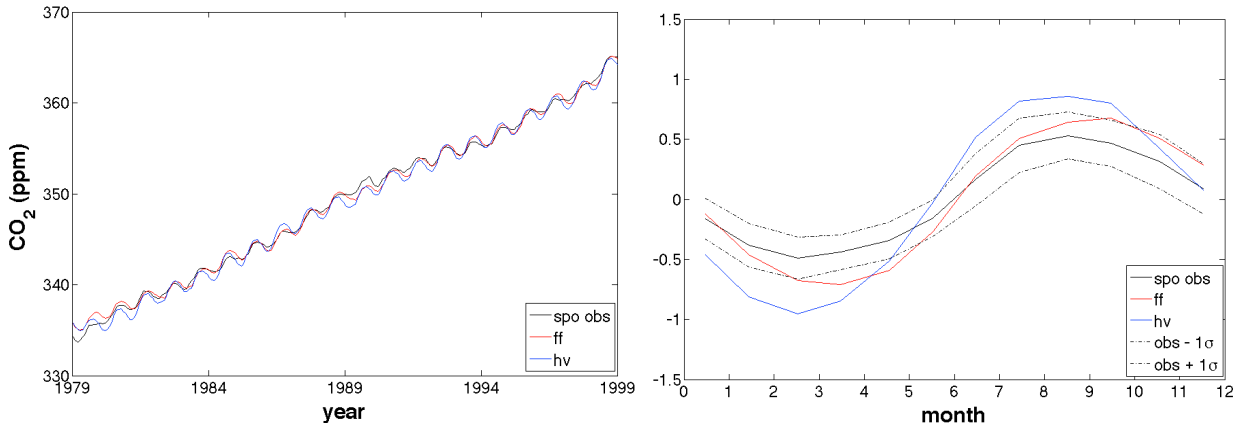


Figure 8: Station = SPO (South Pole). Black = observational data, blue = **hv** model output, red = **ff** model output. Left: monthly atmospheric CO_2 concentration, over model calibration period. Right: running mean (inter-annual trend) removed, and mean seasonal cycle constructed over the central 19 years of calibration. Dashed lines signify 1σ of distribution either side of the mean.

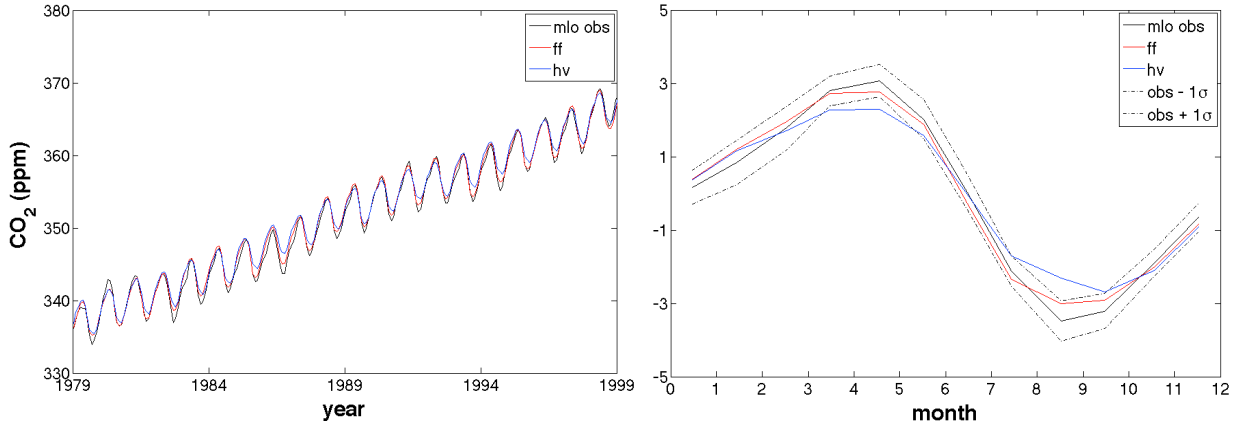


Figure 9: Station = MLO (Mauna Loa). Black = observational data, blue = **hv** model output, red = **ff** model output. Left: monthly atmospheric CO_2 concentration, over model calibration period. Right: running mean (inter-annual trend) removed, and mean seasonal cycle constructed over the central 19 years of calibration. Dashed lines signify 1σ of distribution either side of the mean.

The overall fit is improved for atmospheric concentrations at both the South Pole and Mauna Loa stations (from an R^2 of -0.67 to 0.38 for SPO, and from 0.79 to 0.92 for MLO — **hv** and **ff** respectively). The model reconstructs both areas with high and low seasonal cycles well (See Figs. 8 and 9). Most importantly from this result, the phase lag of the seasonal cycle at SPO has been corrected. With regards to the seasonal

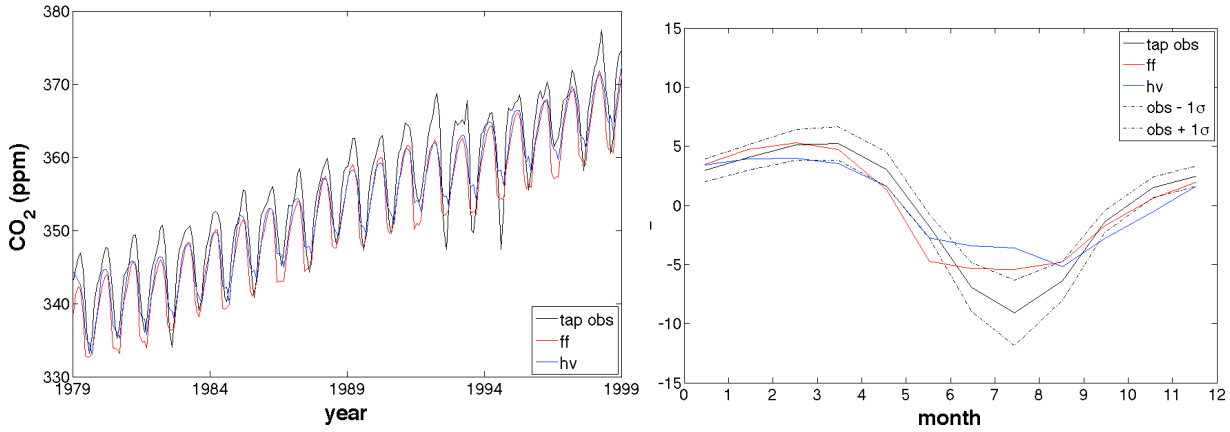


Figure 10: Station = TAP (Tae-Ahn Peninsula, Korea). Black = observational data, blue = **hv** model output, red = **ff** model output. Left: monthly atmospheric CO_2 concentration, over model calibration period. Right: running mean (inter-annual trend) removed, and mean seasonal cycle constructed over the central 19 years of calibration. Dashed lines signify 1σ of distribution either side of the mean.

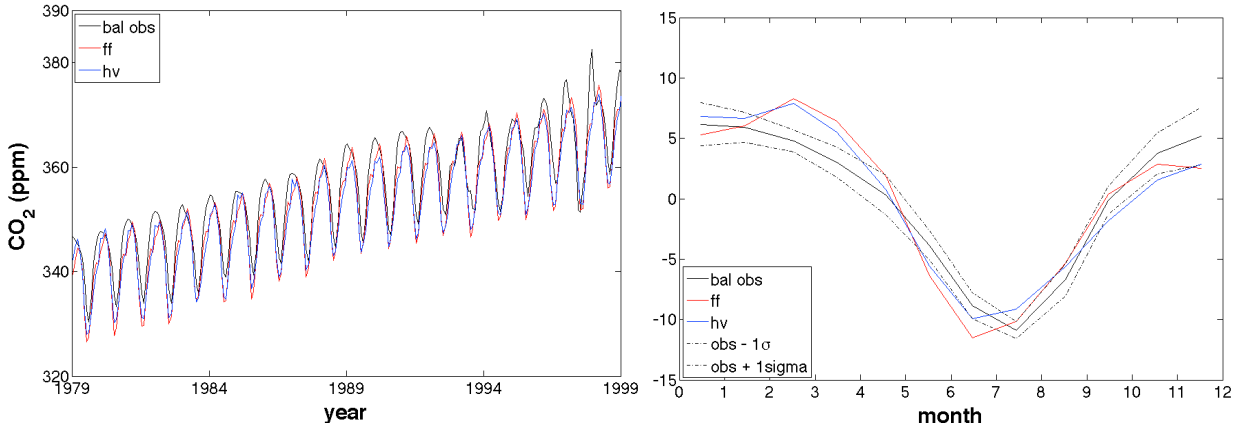


Figure 11: Station = BAL (Baltic Sea). Black = observational data, blue = **hv** model output, red = **ff** model output. Left: monthly atmospheric CO_2 concentration, over model calibration period. Right: running mean (inter-annual trend) removed, and mean seasonal cycle constructed over the central 19 years of calibration. Dashed lines signify 1σ of distribution either side of the mean.

variation alone the R^2 fit is improved from 0.17 to 0.78 for SPO, and from 0.94 to 0.98 for MLO. Fig. 10 shows that for the Tae-Ahn Peninsula (TAP), the seasonal cycle (in itself is an adequate fit — R^2 is 0.80 for **hv**, and 0.89 for **ff**), is offset from the observational inter-annual trend (both curves fall below that of the observational data). This deconstruction technique suggests that it is the background term which is in fact the source of most of the error, not the seasonality as suggested in previous studies [48].

Similar results are shown at the Baltic Sea station (BAL). Despite being a little too high in March/April, the seasonal fit is good (R^2 fit = 0.90 for **hv**, and 0.87 for **ff**). However, the inter-annual variation is displaced. Peaks fail to be high enough, while troughs are too low.

6.2 Parameters

Under the unbound calibration, parameters and associated uncertainties as in Table 1 are predicted by **hv** and **ff**. Between the two model calibrations, most parameters have changed only within a small range. The parameters whose values have changed most significantly relate to GPP in BETHY.

S_r parameters are listed in Table. 2. The parameters for regions 1 and 4 (two of the largest emitters, and the most northerly regions) are negative, predicting a reversed seasonal cycle. The parameters for regions 3, 4, and 5 are very high, with values around 2. This suggests that the amplitude of their seasonal cycles are roughly equal to the total emissions for the regions.

Fig. 12 displays the reduction in error with respect to all parameters through calibration of the **ff** model. There is little or no error reduction in the fossil fuel seasonality parameters (S_r — numbers 58-62); all predicted uncertainties being $< 2\%$ below their original values.

Fig. 13 shows regionally divided output of the fossil fuel model over the calibration period ($\sum_{i,j} F^r_{i,j,m}$). Since the uncertainty on the parameters (σ_p) is still large, the range of seasonal variation for each region is extensive. Fig. 14 shows that for region 2, the lower range on S_r constitutes a 6 month change in phase of the seasonal cycle. For region 4, the troughs of the cycles constitute negative emissions - a physical impossibility.

6.3 Fluxes

6.3.1 Global

Mean Annual Fluxes Globally, the mean annual CO_2 flux from fossil fuel emissions has increased by 0.14GtCyr^{-1} in the **hv** model. The terrestrial biosphere overcompensates for this, and decreases by 0.15GtCyr^{-1}. This makes the mean annual global flux slightly less in the **hv** model (see Fig. 15).

Table 1: Parameter numbers, symbols (where bracketed acronyms denote various PFTs), values and associated uncertainties: Initial ($\mathbf{x}_0$) and predicted as per **hv** and **ff** model calibration ($\mathbf{x}_p$). Units for each parameter are detailed below.

Parameters		**Initial**		**Predicted**			
No.	Symbol	x_0	σ_i	x_p (**hv**)	σ_p	x_p (**ff**)	σ_p
1	V^{25}_{max} (TrEv)	60.0	20.00	24.19	49.60	25.57	36.19
2	V^{25}_{max} (TrDec)	90.0	20.00	38.89	52.47	58.30	7.96
3	V^{25}_{max} (TmpEv)	41.0	20.00	89.05	8.17	-45.99	1782.85
4	V^{25}_{max} (TmpDec)	35.0	20.00	31.11	18.43	91.84	7.38
5	V^{25}_{max} (EvCn)	29.0	20.00	2.67	0.00	47.29	11.86
6	V^{25}_{max} (DecCn)	53.0	20.00	93.31	2.82	71.41	14.78
7	V^{25}_{max} (EvShr)	52.0	20.00	92.39	17.92	89.18	10.96
8	V^{25}_{max} (DecShr)	160.0	20.00	51.02	57.69	52.09	29.85
9	V^{25}_{max} (C3Gr)	42.0	20.00	23.71	7.36	2.88	26.40
10	V^{25}_{max} (C4Gr)	8.0	20.00	1.96	46.48	2.68	31.66
11	V^{25}_{max} (Tund)	20.0	20.00	37.68	0.00	7.23	26.20
12	V^{25}_{max} (Wetl)	20.0	20.00	34.47	11.62	21.90	18.17
13	V^{25}_{max} (Crop)	117.0	20.00	67.15	3.99	42.41	13.23
14	$a_{J,V}$ (TrEv)	1.96	5.00	1.46	6.32	0.76	12.02
15	$a_{J,V}$(TrDec)	1.99	5.00	1.72	5.78	1.97	4.99
16	$a_{J,V}$(TmpEv)	2.00	5.00	1.42	0.00	2.00	5.00
17	$a_{J,V}$(TmpDec)	2.00	5.00	2.52	0.00	1.55	6.43
18	$a_{J,V}$(EvCn)	1.79	5.00	1.28	0.00	0.63	8.94
19	$a_{J,V}$(DecCn)	1.79	5.00	2.41	0.00	2.31	3.88
20	$a_{J,V}$(EvShr)	1.96	5.00	1.98	0.00	1.82	5.36
21	$a_{J,V}$(DecShr)	1.66	5.00	1.88	0.00	1.62	5.09
22	$a_{J,V}$(C3Gr)	1.90	5.00	2.48	0.00	2.00	4.75
23	$a_{J,V}$(C4Gr)	140.0	20.00	178.74	0.00	24.36	112.95
24	$a_{J,V}$(Tund)	1.85	5.00	2.27	0.00	1.98	4.66
25	$a_{J,V}$(Wetl)	1.85	5.00	2.02	0.00	2.44	3.79
26	$a_{J,V}$(Crop)	1.88	5.00	0.86	0.00	2.20	4.27
27	$f_{R,leaf}$	0.40	25.00	0.14	0.00	0.11	6.35
28	$f_{R,growth}$	1.25	5.00	01.82	0.00	1.00	5.52
29	$Q_{10,f}$	1.5	100.00	1.75	0.00	1.05	4.01
30	$Q_{10,s}$	1.5	100.00	0.35	0.00	1.04	4.31
31	τ_f	1.5	270.95	0.85	0.00	28.29	2.96
32	κ	1.0	230.00	3.72	0.00	2.51	17.13
33	f_S	0.2	43.07	0.75	0.00	16.27	14.72
34	E_{R_d}	45000.0	5.00	40114.24	0.00	24049.30	8.84
35	$E_{V_{max}}$	58520.0	5.00	72953.87	0.00	53516.15	5.45
36	E_{K_O}	35948.0	5.00	34522.58	0.00	35892.81	5.01
37	E_{K_C}	59356.0	5.00	53429.59	0.00	57681.20	5.14
38	E_k	50967.0	5.00	47429.39	0.00	49409.11	5.16
39	α_q	0.28	5.00	0.39	0.00	42.25	3.24
40	α_i	0.04	5.00	0.0002	0.00	0.000009	21.96
41	K^{25}_C	460.0	5.00	02.70	0.00	0.018	12.78
42	K^{25}_O	330.0	5.00	320.00	0.00	346.60	4.76
43	$\alpha_{\Gamma,T}$	1.7	5.00	0.68	0.00	0.93	8.83
44	β(TrEv)	1.0	25.00	1.07	0.00	0.54	31.71
45	β(TrDec)	1.0	25.00	0.99	0.00	0.92	24.27
46	β(TmpEv)	1.0	25.00	0.71	0.00	1.02	24.57
47	β(TmpDec)	1.0	25.00	2.32	0.00	1.42	17.49
48	β(EvCn)	1.0	25.00	0.88	0.00	0.72	12.48
49	β(DecCn)	1.0	25.00	0.71	0.00	0.66	11.97
50	β(EvShr)	1.0	25.00	2.17	0.00	2.60	9.54
51	β(DecShr)	1.0	25.00	0.25	0.00	0.31	52.19
52	β(C3Gr)	1.0	25.00	1.64	0.00	0.52	35.04
53	β(C4Gr)	1.0	25.00	0.75	0.00	1.36	18.34
54	β(Tund)	1.0	25.00	0.86	0.00	0.84	26.28
55	β(Wetl)	1.0	25.00	0.36	0.00	.35	56.94
56	β(Crop)	1.0	25.00	0.51	0.00	2.71	9.17
57	offset	338.0	0.30	335.98	0.30	336.89	0.05

Units are as follows: V_{max}, μmol(CO_2)m^{-2} s^{-1}; $a_{G,T}$, μmol(CO_2)mol(air)$^{-1}$(°C)$^{-1}$; activation energies E, J/mol; t_f, years; offset, ppmv; all others unitless.
Uncertainties are in percentage (if equal to zero, the error is too small to be expressed).
Uncertainties represent one standard deviation.

Table 2: Additional parameters in **ff**: numbers, symbols, description, values and associated uncertainties: Initial ($\mathbf{x}_0$) and predicted as per **ff** model calibration ($\mathbf{x}_p$). Parameters are unitless. Uncertainties are expressed as percentages of parameter values and represent one standard deviation.

Parameters			**Initial**		**Predicted**	
No.	Symbol	Description	x_0	σ_i	x_p (**ff**)	σ_p
58	S_1	Seasonality amplitude factor for region 1	0.3	166.67	-0.7022	70.91
59	S_2	Seasonality amplitude factor for region 2	0.3	166.67	0.1009	492.84
60	S_3	Seasonality amplitude factor for region 3	0.3	166.67	2.0474	24.42
61	S_4	Seasonality amplitude factor for region 4	0.3	166.67	-2.1009	23.39
62	S_5	Seasonality amplitude factor for region 5	0.3	166.67	1.9246	25.98

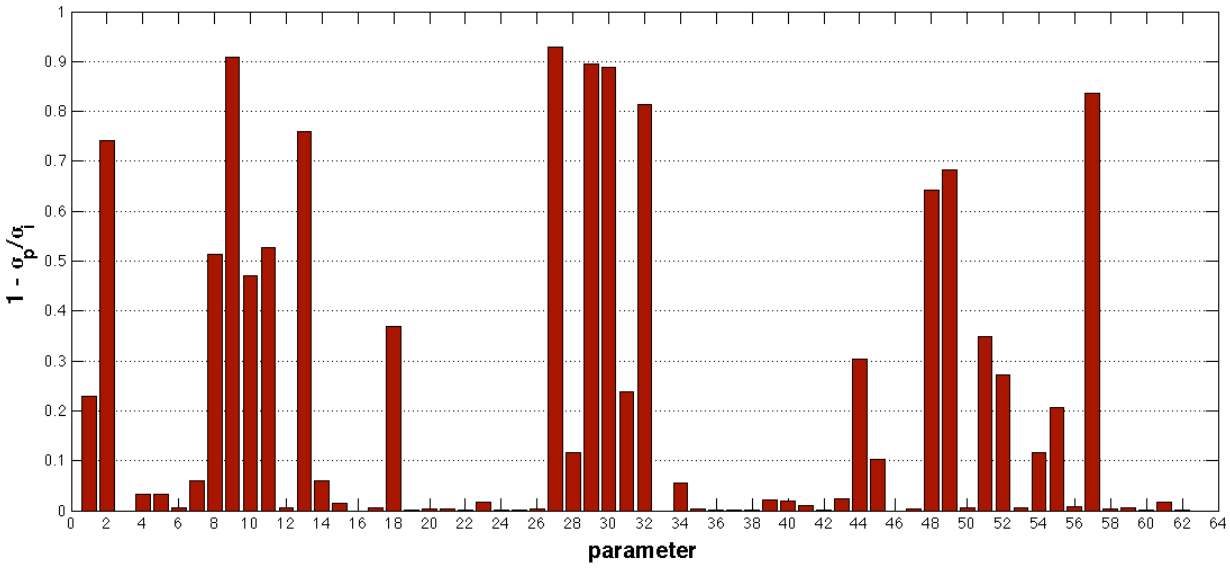


Figure 12: Uncertainty reduction in the **ff** model version. σ_i and σ_p denote the initial (pre-calibration) and predicted (post-calibration) uncertainties respectively, associated with each parameter. Note that there is little/no uncertainty reduction in S_r parameters (numbers 58-62).

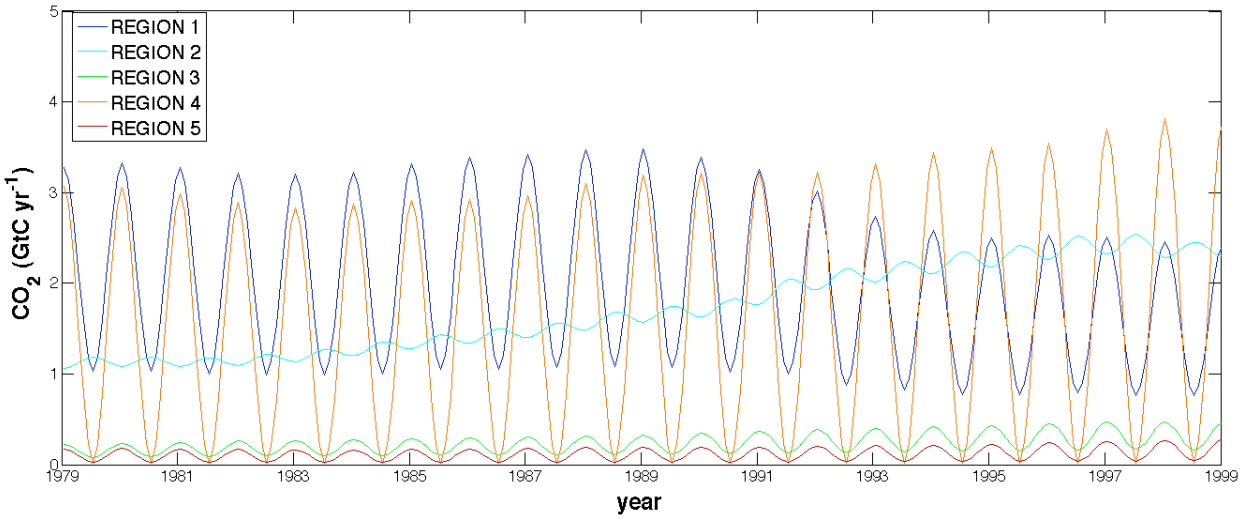


Figure 13: Predicted total monthly CO_2 fossil fuel emissions for each region ($\sum_{i,j} F^r_{i,j,m}$) after calibration of S_r. Summation over all regions is the fossil fuel flux output to TM2.

Mean Monthly Fluxes Fig. 16 displays the effect of introducing a seasonally varying fossil fuel source on the *global* mean monthly terrestrial and fossil fuel seasonal cycles. The main effect on the terrestrial cycle appears to be in increase in the amplitude of the seasonal cycle. The largest increase in emissions (positive flux) occurs around June ($+\sim 3\text{GtCyr}^{-1}$), but this is outweighed by an increase in the sequestration around the end of the year (lower minimum in September $\sim$-6.25GtCyr^{-1}). The phase of the cycle has also shifted - with both maximum and minimum flux almost a month earlier in the year.

The global mean seasonal fossil fuel cycle is a full 6 months out of phase, and has an amplitude of roughly 15% of that of the terrestrial cycle. Fig. 17 shows how the mean seasonal cycle of both fluxes differ between the hemispheres. The opposite seasonal cycle of terrestrial flux between hemispheres is evident, as is the difference in amplitude. In the Southern Hemisphere, seasonal cycles of both terrestrial fluxes and fossil fuel emissions are much lower (for example, the terrestrial flux for the **ff** model has a peak-to-peak of 42.72GtCyr^{-1} in the Northern Hemisphere, but only 10.062GtCyr^{-1} in the Southern Hemisphere). The addition of seasonally varying fossil fuel emissions has increased the amplitude of the seasonality in terrestrial fluxes in both hemispheres.

Regional Fluxes Regionally, Europe and Russia are the combined highest emitters of CO_2 per annum on average, contributing just under 35%. Regions 3 and 5 only contribute a small amount to total fossil fuel emissions ($< 6.5\%$). Region 2 (North-Mid Africa, the Middle-East and Central and Southern Asia) has the

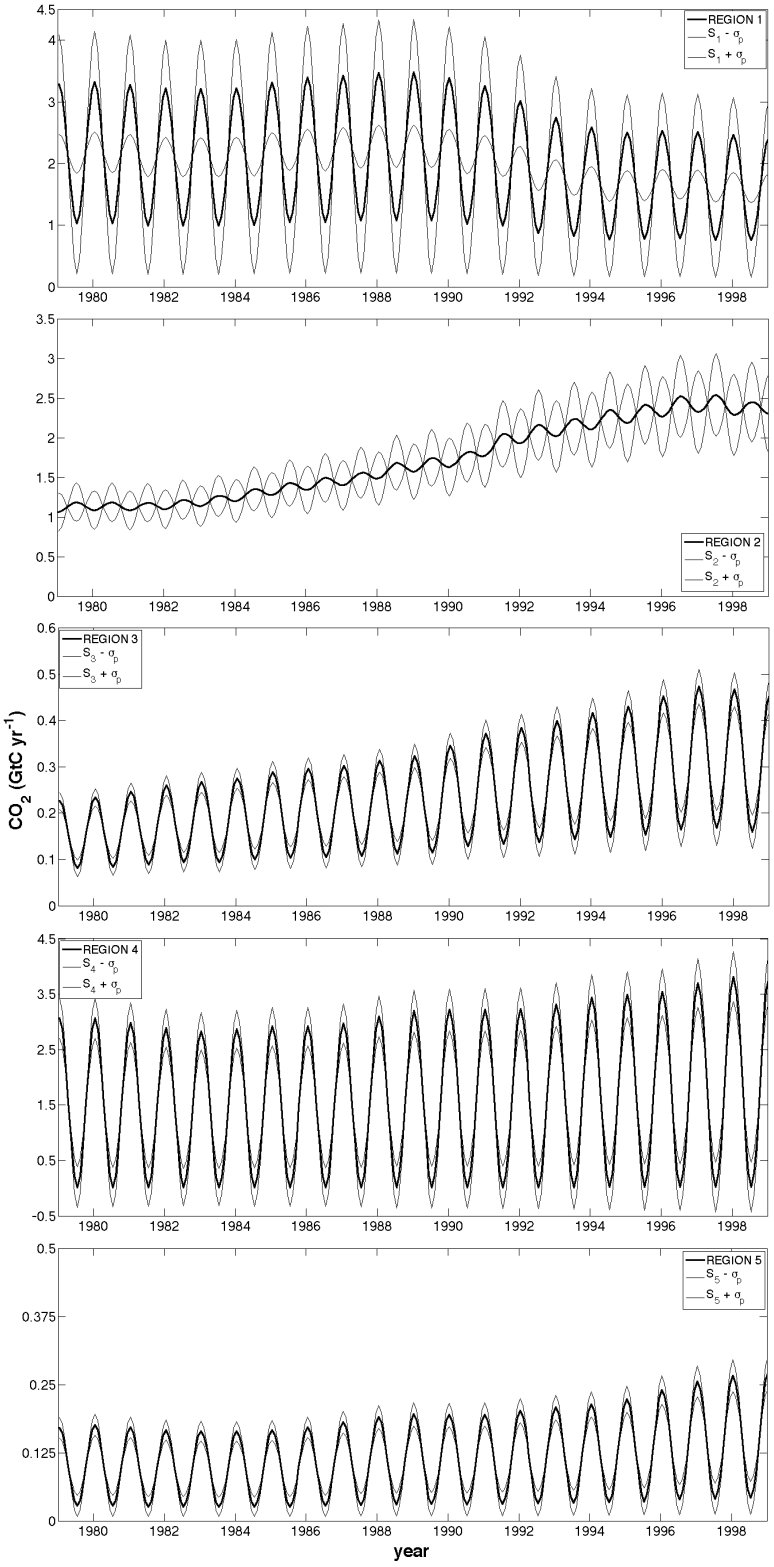


Figure 14: Predicted total CO_2 fossil fuel emissions for each region (full line) after calibration ($\sum_{i,j} F^r_{i,j,m}$). Dashed lines represent simulated total emissions with $S_r = \mathbf{x}_p \pm \sigma$, representing one standard deviation in the spread in seasonality dictated by the uncertainty in S_r (σ_p) of the model output. Regions 1 (top) to 5 (bottom) are displayed.

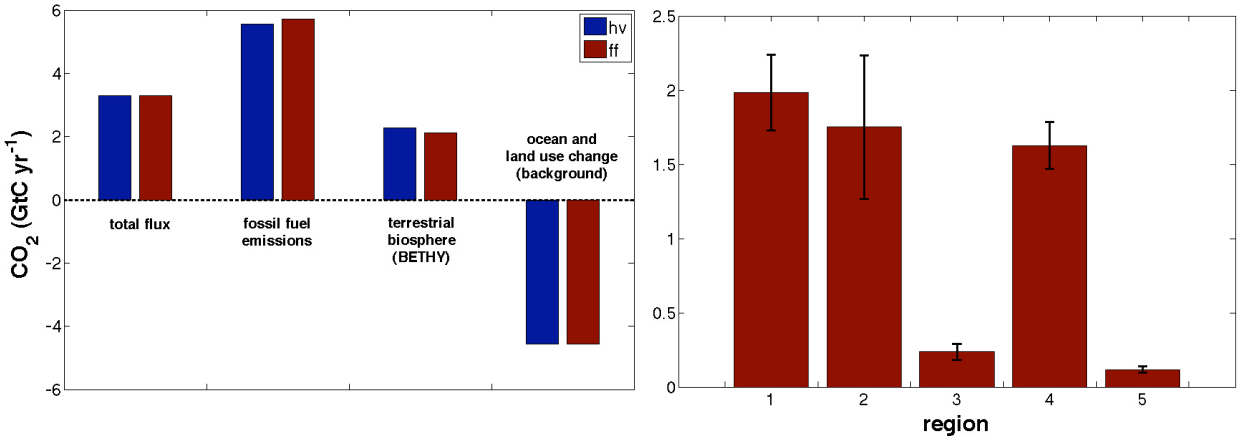


Figure 15: Left: Mean annual CO_2 flux from terrestrial biosphere, fossil fuel emissions and background fluxes for **hv** (blue) and **ff** (red) model calibrations (averaged over calibration period). Right: Mean annual CO_2 flux from each region (averaged over calibration period). Error bars indicate 1σ standard deviation of distribution.

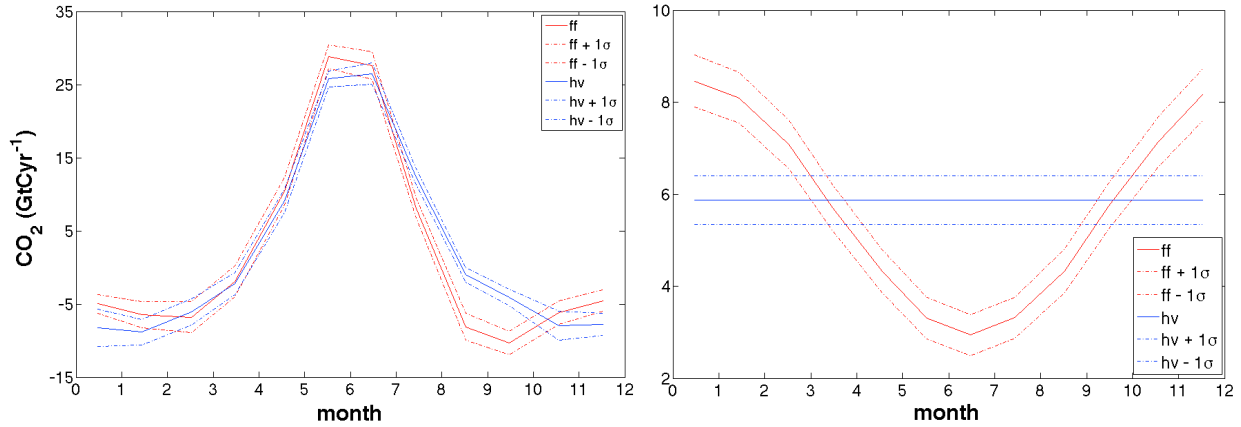


Figure 16: Mean monthly terrestrial flux (left) and fossil fuel emissions (right) in **hv** (blue) and **ff** (red) model calibrations, representing the average seasonal cycle of each. No running mean has been removed. Dotted lines represent 1σ distribution around the mean.

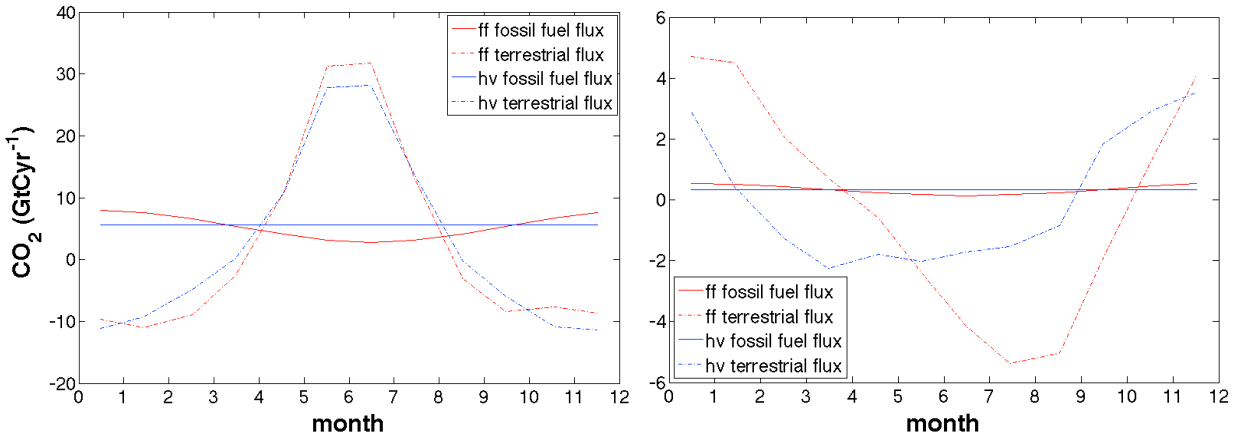


Figure 17: Mean monthly terrestrial and fossil fuel flux from **hv** (blue lines) and **ff** (red lines), totalled over the Northern Hemisphere (left) and the Southern Hemisphere (right). Full lines indicate fossil fuel fluxes; dotted lines denoting terrestrial fluxes. No distribution around the mean is shown, however is estimated to be similar in magnitude to those shown in Fig. 16.

largest standard deviation indicating the largest change in emitted CO_2 over the model calibration period. Regional mean annual fossil fuel fluxes, and their standard deviations and detailed in Appendix E.1, in which all values are shown to be comparable to the original emissions data set.

6.3.2 Spatial Distribution

Mean Seasonal Differences Spatially distributed, the largest increase in mean monthly CO_2 emissions occurs on the Western coast of the United States, with an increase of $0.23 GtCyr^{-1}$ in February, followed closely by Central Europe. The largest decrease occurs in July on the Eastern coast of the United States ($-0.39 GtCyr^{-1}$) with significant decreases also occurring in Japan. Interestingly, whereas in Europe and the United States, emissions changes are likely due to the strongly reversed seasonality in these regions (they are marked as increases in the Northern Hemisphere winter months, and decreases in the summer months), the decrease in emissions in Japan is year-long (appears both in summer and winter months). This means that the emissions for region 2 have been reallocated, possibly to Northern India, where a slight increase is seen at both seasonal extremes (Fig. 18).

There is a high spatial correlation between densely populated areas and those areas with the largest changes in emissions. The allocation of emissions per region (in the **ff** version) as opposed to by country (**hv**), has been altered. For example, although Northern India is highly populated, previous representation of emissions has not allocated so much CO_2 per head as in Japan. This is an example of an introduced uncertainty due to the population weighting scheme used in this study, which must be kept in mind.

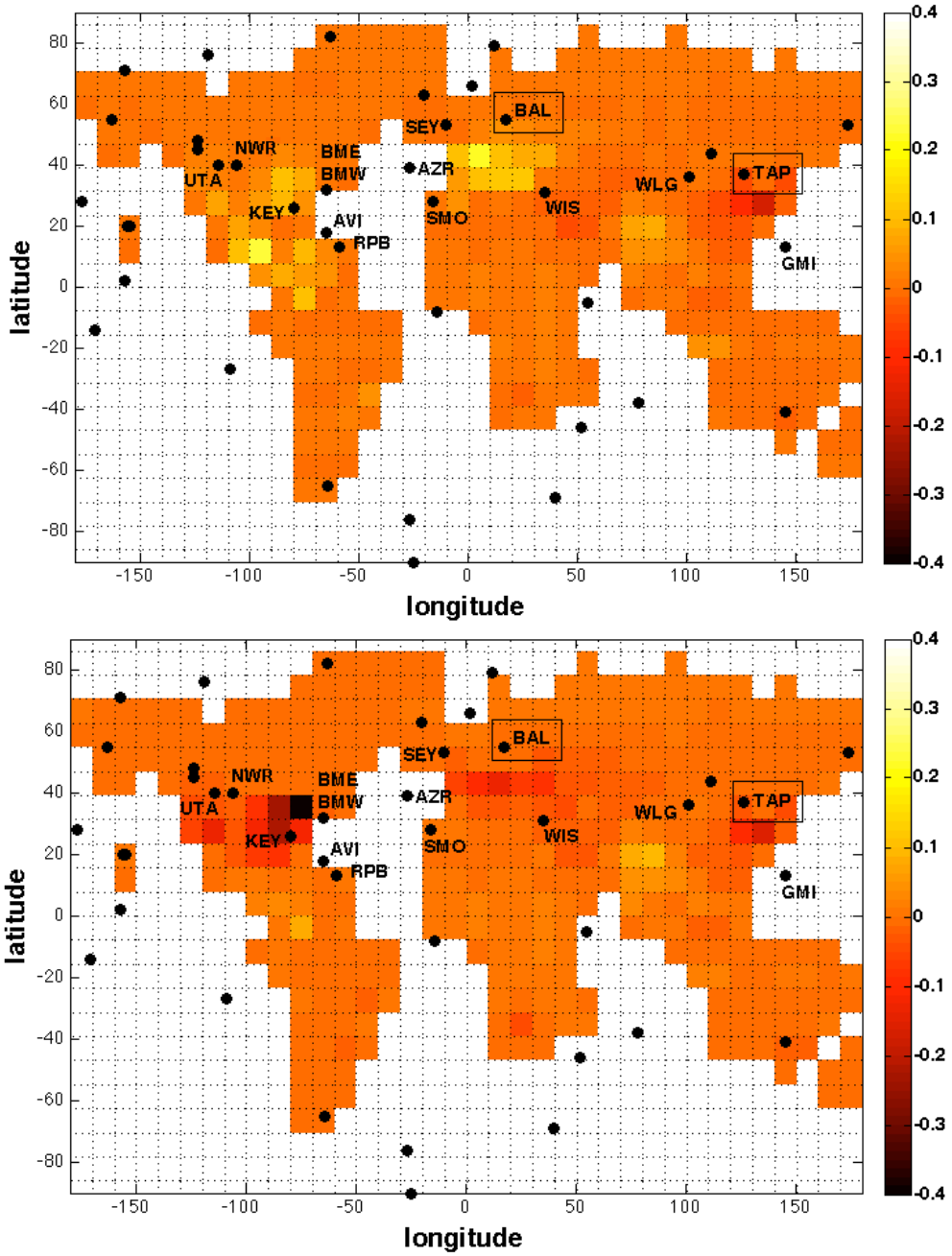


Figure 18: Spatial distribution of the difference between the mean monthly fossil fuel CO_2 flux in the two model versions (**ff** - **hv**), for two months with the largest difference in flux. Scale is in $GtCyr^{-1}$. Top: 14.6 days into the year (0.04 of a year) — Jan. Bottom: 197.1 days into the year (0.54 of a year) —Jul. Black circles represent positions of observation stations. Those closest to the changes in flux are labelled.

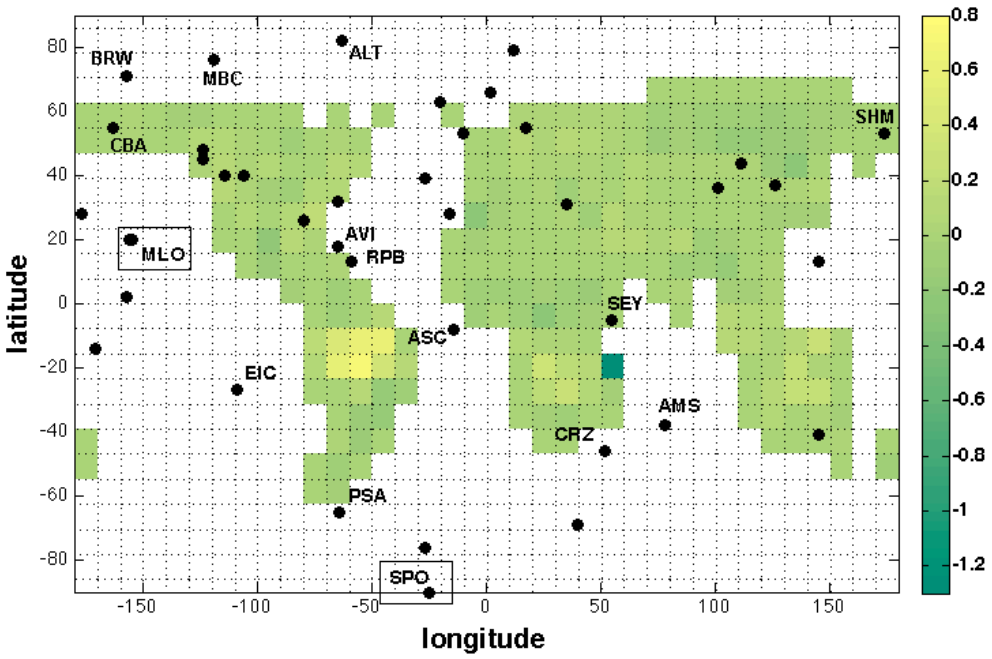


Figure 19: Spatial distribution of the difference between the mean monthly terrestrial CO_2 flux in the two model versions (**ff** - **hv**), for the month with the largest difference in flux. Scale is in $GtCyr^{-1}$. Distribution represents 43.8 days into the year (0.12 of a year) — Feb. Black circles represent positions of observation stations. Those closest to the changes in flux, and others of interest are labelled.

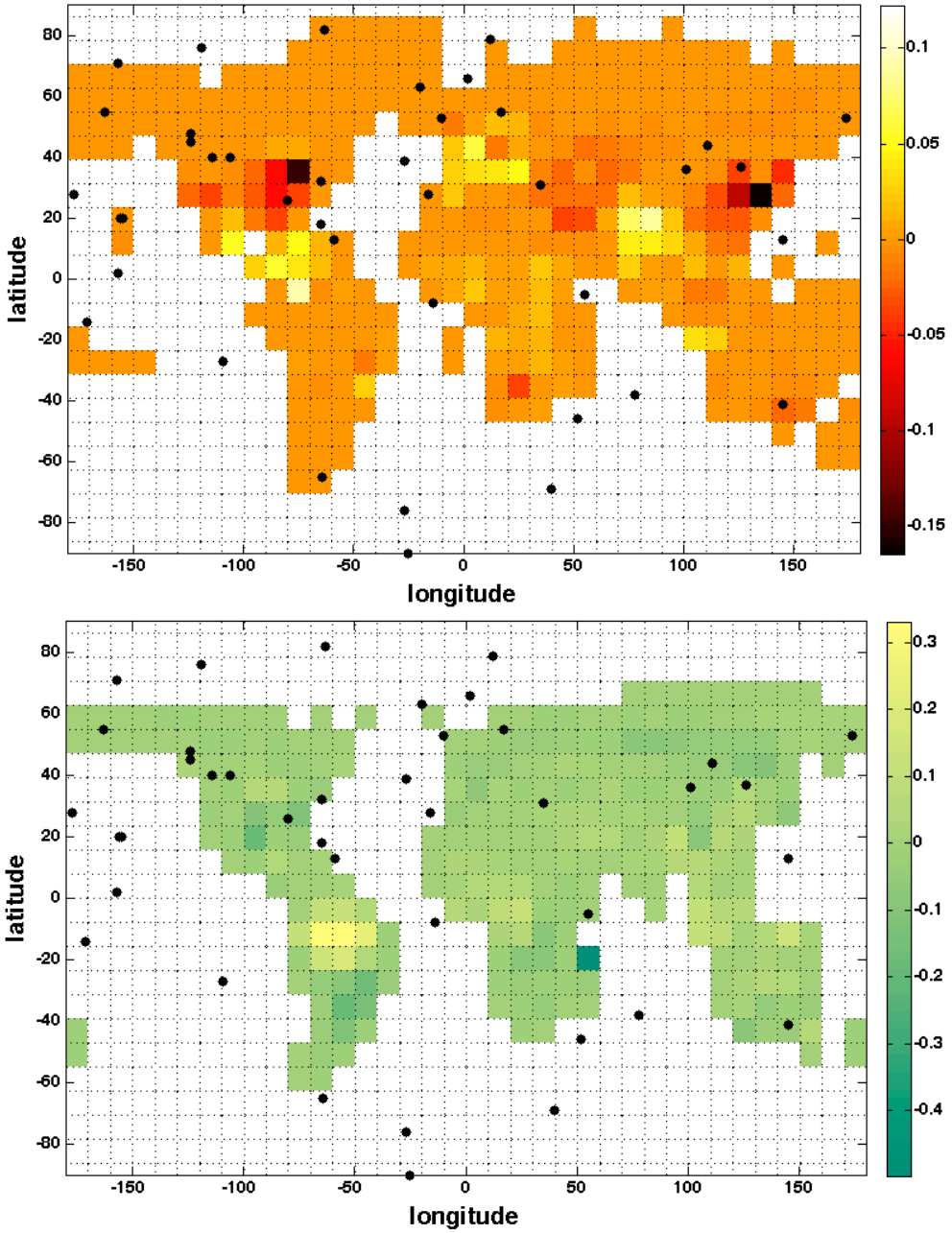


Figure 20: Spatial distribution of the difference between the mean annual fossil fuel (top) and terrestrial biosphere (bottom) CO_2 flux in the two model versions (**ff** - **hv**). Scale is in $GtCyr^{-1}$. Black circles represent positions of observation stations.

The terrestrial biosphere flux in two areas in the Southern Hemisphere are most affected by the introduction of seasonal fossil fuel flux. Fig. 19 shows the largest increase in flux — 0.73GtCyr^{-1}, is located in central South-America. The most northerly tip of Madagascar appears to have the largest decrease in flux at roughly -1.28GtCyr^{-1}. Both of these largest changes occur for the month of February (Southern Hemisphere summer).

Mean Annual Differences Fig. 20 shows the difference between the mean *annual* flux for both fossil fuel emissions and terrestrial biosphere. Once seasonality is removed, it is found that the differences in flux are not as significant as Figs. 18 and 19 suggest. The largest changes in annual fossil fuel flux are less than $\pm$0.17GtC^{-1}, and those in flux from the terrestrial biosphere well below $\pm$0.5GtCyr^{-1}.

6.4 Spatial Distribution of Fit to Data

Fig. 21 shows that for most stations, the fit to observations was improved due to the introduction of a seasonal cycle into fossil fuel emissions. Both bar charts in Fig. 21 divide the stations into region (or geographically closest region in some cases) — denoted by the colour of the bars. Then within each regional classification, the stations are arranged so that those with lowest latitude are to the left, and those at higher latitudes are to the right.

Firstly, it is clear from the top chart, that originally in the **hv** model, region 3 (Southern Africa, the tip of South-East Asia and Oceania) was most poorly reconstructed, with negative R^2 values indicative of a cycle out of phase significantly. Even though the phase is not rectified in all cases in this region, it at least approaches a better fit. This is indicated by large Γ values in the chart below.

The largest improvement in fit to observations seasonally are at Ascension Island (ASC), Easter Island (EIC) and Amsterdam Island (AMS). ASC and EIC now have seasonal cycles close in phase and magnitude to observations. These two stations are located to the East and West of South America respectively (ASC is allocated to region 5, although strictly speaking is it roughly equidistant between the African West Coast and the South American East Coast).

Most stations located in regions 1,2 and 3 already have good fit to data in the **hv** (apart from SMO (located at Tutuila, American Samoa), whose seasonal cycle is significantly out of phase). Any marginal increases or deceases in fit such as at the majority of stations in these regions are disregarded. The size of the uncertainty on atmospheric concentrations and the distribution around the mean when averaging outweigh the fitting accuracy.

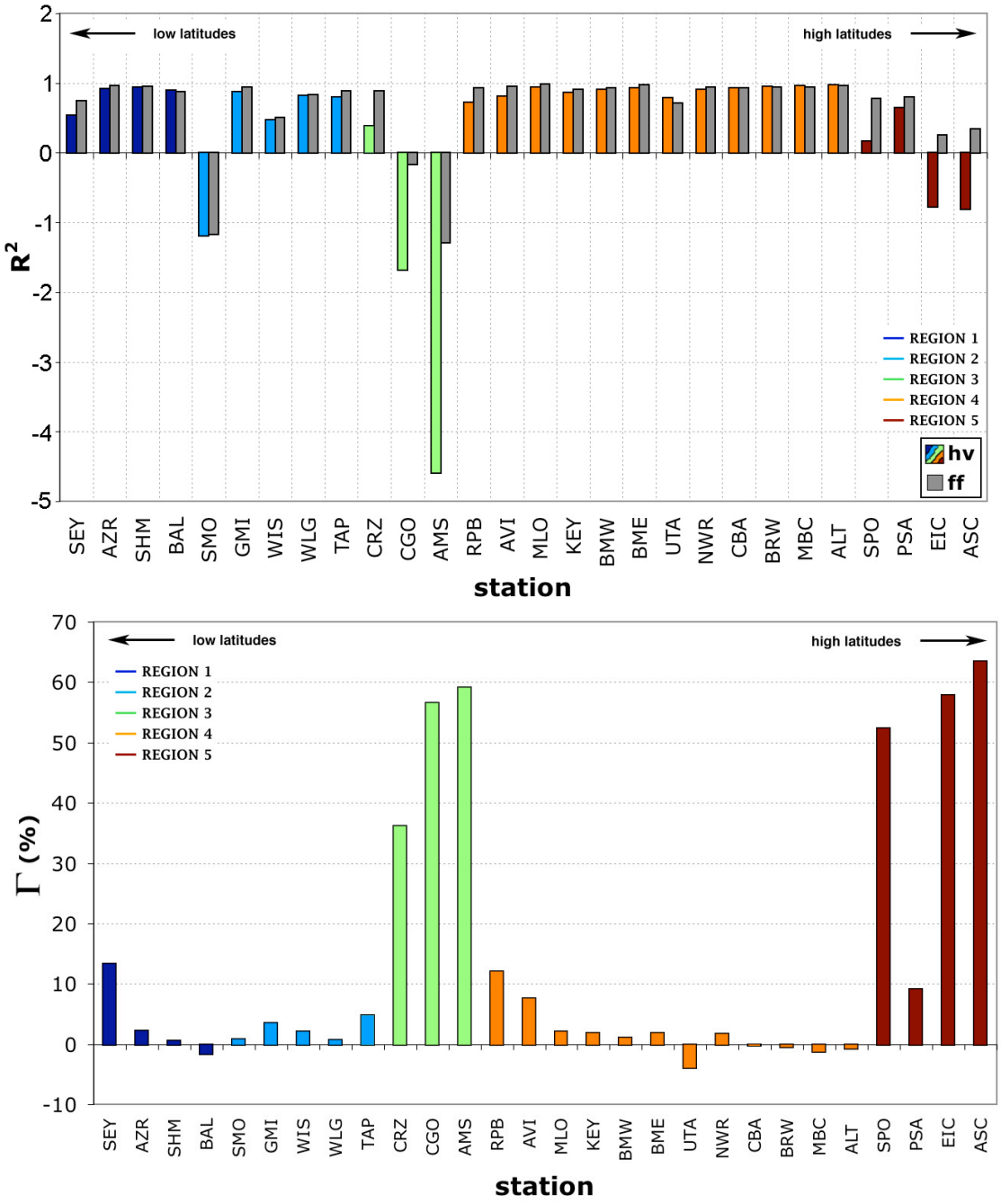


Figure 21: Top: R^2 fit of **hv** (colour) and **ff** (grey) to observations of seasonal cycle (running mean removed from atmospheric concentrations at observation stations). Bottom: Fit improvement, as calculated in Eq. 17 (see Appendix D.3), of atmospheric CO_2 concentration at various observation stations (expressed as a percentage of convergence to perfect fit from fit in **hv**). In both cases, stations are arranged by region, and by latitude (in each regional grouping left to right signifies low latitudes to high latitudes). Colours denote regions (see Fig. 28).

Overall, most fit improvement is made at lower latitudes, with almost all stations which show a vast improvement in fit located South of the equator (regions 3 and 5). Palmer Station, Antarctica (PSA) shows relatively little improvement, although its fit was already vastly better than all other stations within the region. It is important to note here that station distribution is not homogeneous. Therefore a rigourous assessment of the spatial distribution of fit improvement is not possible. Graphical representation of the atmospheric CO_2 concentrations relative to observations at some of these stations (and others of interest) may be found in Appendix E.2.

7 Discussion

The fit of the **ff** model output to observational data is encouraging. Overall, the tendency for CCDAS to reconstruct intra-annual CO_2 variability poorly (relative to inter-annual trends) has been improved upon through the addition of seasonality in fossil fuel emissions, in agreement with suggestions from Levin (2003) [35]. Inter-annual trends are also, in general, improved upon, although fit to observations in this case was already very good in most areas of the globe [52, 48, 53]. Most improvement has been made in the Southern Hemisphere where previously (in the **hv**) reconstruction of atmospheric CO_2 concentrations was poorest. Furthermore, stations cited as having very poor fit to observations by Rayner (2005) — Easter Island (EIC) and Ascension Island (ASC) — display some of the highest improvements in fit due to the addition of a seasonal fossil fuel emission cycle (58% and 64% respectively).

Results also suggest that at the Baltic Sea (BAL) and the Tae-Ahn Peninsula (TAP) stations, poor fitting to observations is more likely to be due to the magnitude of the inter-annual flux, not seasonality as previously suggested [48]. The mean annual fossil fuel emissions around TAP have decreased in the **ff** model, and those near BAL have increased. This explains the slight increase in the overall fit relative to the running mean at BAL (R^2 has increased from 0.53 to 0.60), and a slight decrease in fit at TAP ($R^2 = 0.60$ in **hv**, compared to 0.43 in **ff**).

Despite these initially promising results, analysis of the seasonality parameters for the fossil fuel emissions reveals highly unlikely and uncertain predicted seasonal cycles for emissions. Whilst the mean annual emissions are still comparable to those originally provided in the data set (see Table E.1 in Appendix B for comparison), the magnitude of the seasonality in most regions is much greater than suggested by previous studies and literature regarding seasonality in emissions [6, 35, 44, 51]. Although estimates on seasonal amplitude of emissions are not available for all regions (values from the study by Rotty et al. (1987) have

been shown to largely underestimate the seasonality of emissions [35]), it is possible to make a comparison in some cases. For example, whereas Blasing (2005) cited a peak-to-peak of 0.035GtC in areas of North America with a high in winter months, the estimate after calibration of the **ff** model is a peak-to-peak of almost 3GtC for this region of opposite seasonal cycle (high in summer). Even including uncertainties on this value, it is highly improbable that the model's seasonal cycle will diminish almost 100-fold and reverse the phase of the cycle. This region compares most poorly to previous attempts to quantify the amplitude of the seasonal cycle.

Other regions do not exhibit much more realistic behaviour. The seasonal cycle over Europe and Russia also displays a high in summer months (see Fig. 14), despite having its' population distributed at higher latitudes (mean annual temperatures are lower than in region 4) making it more likely to have a traditional cycle following energy demand through heating [35]. Another important point to note is that emissions within the calibrated **ff** model drop near zero for three of the regions (3, 4, and 5) at the middle of the year. For the Southern regions (3 and 5), this is not so anomalous with our knowledge regarding the regions. They are low emitting in general, accounting for just over 6% of global emissions together, and during summer months (beginning/end of the year), their consumption is likely to be reduced. However, for North America this is an unimaginable result. In some areas of North America (due to the latitudinally dependent nature of the seasonality term in emissions), the model predicts a *negative* fossil fuel emission. This is entirely unphysical.

Region 2 is the most likely candidate for realistic seasonality, with a peak-to-peak ranging between $\sim$0.1 and 0.25GtC relative to an average annual emission of 1.75 ± 0.49GtCyr^{-1}. Even this however is most probably too strong a seasonal cycle, although no studies have been found with details relating to the seasonality in emissions in this region an so no quantitative comparison may be made. Qualitatively, since this is a low-latitude region (with population density at its highest between 20 and 30°N — see Appendix E.3), it is expected that it would exhibit low seasonality in emissions correlated to low seasonality in climate. Furthermore, within this region, some of the least predictable emitters are found. Relating to the socio-economic analysis of emissions, both India and China are *transition* countries, meaning that their emissions have been increasing rapidly for much less time that other established high-emitting countries [41, 13, 14]. Therefore any trends in emissions now are unlikely to have been evident throughout the calibration period.

This highlights one of the main problems within the fossil fuel emissions model. One seasonal trend must represent the whole calibration period - there is no possibility for phase transformation or seasonal cycle reversal within a region in the given time period. Since some studies suggest cycles are changing,

then this possibility should be encompassed in the seasonality term in emissions [6]. One other important result of the manner in which seasonality is constructed by Eq. 14 is the latitudinal dependance. Since the model presumes a high in winter months (for both hemispheres), it increases the seasonality towards higher latitudes. However, if the seasonality is reversed (is likely driven by the use of air conditioning) in hotter climes, then this latitudinal dependence should be reversed somewhat — hotter and longer summers towards the equator are likely to constitute more energy consumption [44]. This, of course, must be balanced with less seasonal climates closer to the equator, but may be true of tropical/sub-tropical latitudes. This is simply not available to the calibration of the model.

The most significant improvement in reproduction of seasonal cycles is seen in the Southern hemisphere. This is likely attributed to changes in the reaction of the terrestrial biosphere to the calibration process. It is evident from Figs. 19, 20 and 17 that areas in the Southern Hemisphere have altered the magnitude of the terrestrial flux fairly significantly. The introduction of seasonality in the low fossil fuel emissions in region 5 has prompted reduction in terrestrial flux throughout the year over the same area. It is possible that because the majority of background flux due to land use change is located in this area then this is indicative of an overestimation in the magnitude of this contribution [52]. Also possible is that previously (in the **hv** model), higher mean annual terrestrial biosphere flux may have been substituting for a lacking in intra-annual variability in order to attempt a better likeness to observations. This suggestion is supported by Fig. 17 where seasonality in terrestrial biosphere flux is increased. Since the seasonality in the Southern Hemisphere is dominated by oceanic uptake of CO_2 (more ocean-covered area) [31], then perhaps it is the intra-annual variability this background flux which is underestimated. This may also be the case for region 3, where the seasonal cycle is stronger then expected.

In the Northern Hemisphere, as has already been mentioned, the seasonality in emissions is overestimated significantly by the calibration. This is most likely in part due to the increase in the mean seasonal cycle of terrestrial biosphere flux as shown in Fig. 17 since the Northern Hemisphere's seasonal cycle in atmospheric CO_2 concentrations is dominated by the terrestrial biosphere (large land masses are located here) [31]. The seasonal trend in emissions is on average, 6 months out of phase with those from the terrestrial biosphere. These two changes almost entirely cancel each other out, resulting in little or no improvement in fit to observations in these areas (see Fig. 21). Whereas the resulting change in terrestrial flux is minimal seasonally (because the amplitude is already high), this translates to a huge seasonal cycle in emissions. Rather than 'filling the gap' between seasonality introduced in atmospheric CO_2 concentrations by terrestrial flux and that in observations, the calibration has made the seasonality in fossil fuel emissions work counteractively

to an increased seasonality in terrestrial flux.

The most likely reason for this lies in the calibration and its dependence on initial parameters. First of all, since the fossil fuel flux is fed directly into the transport model (TM2), atmospheric CO_2 concentrations are very sensitive to this flux. In turn therefore, the model's data assimilation is very sensitive to the parameters associated with the fossil fuel emissions, and predicted values show high dependence on initial values. Better estimates are possibly needed before the model will predict realistic emissions scenarios. Inability to reduce the uncertainty on the S_r parameters is also of concern. The uncertainties associated with this seasonal cycle will in turn have a large effect on the uncertainties attached to overall seasonality of atmospheric CO_2 concentrations at points around the globe. This is likely to be especially so in the case where the uncertainty on S_r encompasses a reversed seasonal cycle. Since S_r is a multiplicative factor on the seasonal cycle, the larger the emissions of a region or area, the greater the effect of the uncertainty.

8 Conclusion

On replacement of a seasonally invariant background flux of fossil fuel emissions with a basic sinusoidal seasonal cycle, there is an overall improvement in fit to observational data — a reduced χ-*squared* of 1.73 (relative to 2.23 with the background fossil fuel emissions), suggesting that the inclusion of a seasonal cycle in emissions is important.

On calibration of the 62 tuneable parameters in the system, the largest change between the two model versions (with and without seasonal variation on CO_2 emissions) is shown in parameters associated with GPP in the terrestrial biosphere model. This transpires as an increased seasonal cycle of the terrestrial biosphere fluxes in both Northern and Southern Hemispheres.

The predicted values of the five regional parameter representing intra-annual variations in CO_2 emissions result in highly unrealistic seasonal cycles for most regions — very large amplitudes (often near 100% of the base level emissions), and with only one region in the Northern Hemisphere (Mid-North Africa, the Middle-East and Central and Southern Asia) exhibiting the expected phase (high emissions in winter months, low in summer).

In the worst case (North and Central America), the amplitude of seasonality is predicted to be near 100 $\times$ that suggested in other studies, and peaks during the summer months as opposed to winter [6]. Prediction that the anthropogenic CO_2 emissions for three of the five regions near zero in winter months is equally unrealistic. In Northern Hemispheric regions, this is largely attributed to counteracting an opposite

and vastly increased seasonal cycle in terrestrial flux. In the Southern Hemisphere, an under-estimated seasonality, or a phase lag in the oceanic background flux is thought to be the main influence.

The most useful knowledge gained though this study is that the seasonal function is over-simplistic. It does not take into account any temporal changes in phase of the seasonality, or encompass any differing elements of seasonal energy use (for example emissions due to energy use through heating versus through air conditioning). Furthermore, the amplitude of the seasonality is relative to the magnitude of the emissions which, according to some studies is not representative of inter-annual trends. For countries whose emissions are increasing, often the seasonal cycle is decreasing. Another important omission within the seasonality function is that if the seasonal cycle is reversed, the amplitude still increases with increasing latitude. This is highly unrealistic.

It is proposed that a more complex algorithm is used to describe seasonality. At least two sinusoidal terms are deemed necessary to make predictions more realistic — one representing emissions produced from energy use through heating with a maximum in winter months, and a latitudinal dependence similar to that used in this study; and one representing a cycle 6 months out of phase (with a maximum in the summer months), encompassing emissions driven by what is thought to be air conditioning usage. This term should have different or no latitudinal dependence. It is possible that a third term is necessary to encompass electricity usage in winter through a rise in indoor activities during colder periods with fewer hours of sunlight. These terms will obviously be much more applicable in certain regions, creating a semi-annual cycle of fossil fuel emissions. The superposition of the two cycles will also provide more freedom with respect to the phase of the seasonal cycle of emissions, hopefully improving the phase lag in atmospheric concentrations seen in some areas. It also removes the need for further division of regions since within this formulation, the differing latitudinally varying terms should allow for variation in seasonal cycle within a region. This should aid keeping the number of additional parameters to a minimum.

For this deconstruction, the annual fossil fuel emissions data set will need to be separated into two portions — one containing fuel types commonly used for electricity production, and the other for heat production. If possible a further third group should be pre-subtracted to represent emissions that are unlikely to have a seasonal dependence, for example, transport-related emissions. This more complex deconstruction may require the re-assessment of regional divisions, dependant on available data, and fuel type usage for heating and electricity production.

Finally, more research is required to make assumptions for the initial parameters that are entered into CCDAS. Since the calibration is very sensitive to these values, where possible more realistic and accurate

estimates should be made. Uncertainty relating to the estimates is currently high, meaning predictions are poorly constrained. Research detailing seasonality in emissions for some regions is beginning to emerge from process-based models. Although these generally relate to more developed regions of the globe (studies relating to the United States and Europe have been detailed here), it is possible that better estimates in these areas alone may be sufficient to better constrain the seasonality elsewhere due to their percentage contribution to global emissions. Finally, It is recommended that limits be imposed on the tuneable parameters so that the predicted seasonality is not compensating for other seasonal fluxes.

References

[1] Andres, R., Fielding, D., Marland, G., Boden, T., Kumar, N., and Kearney, A. Carbon dioxide emissions from fossil-fuel use, 1751-1950. *Tellus Series B-Chemical and Physical Meterology 51*, 4 (—1999—), 759–765.

[2] Andres, R. J., Marland, G., Fung, I., and Matthews, E. A $1^\circ \times 1^\circ$ degrees distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950-1990. *Global Biogeochemical Cycles 10*, 3 (—1996—), 419–429.

[3] Archer, D. Fate of fossil fuel co_2 in geologic time. *Journal of Geophysical Research-Oceans 110*, C9 (—2005—).

[4] Baker, D., Law, R., Gurney, K., Rayner, P., Peylin, P., Denning, A., Bousquet, P., Bruhwiler, L., Chen, Y., Ciais, P., Fung, I., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., and Zhu, Z. Transcom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988-2003. *Global Biogechemical Cycles 20*, 1 (—2006—), –.

[5] Balk, D., and Yetman (CIESIN), G. *Gridded Population of the World (GPWv3)*. SEDAC (Socioeconomic Data and Applications Centre (Available at http://sedac.ciesin.columbia.edu/gpw/global.jsp), —2005—.

[6] Blasing, T. J., Broniak, C. T., and Marland, G. The annual cycle of fossil-fuel carbon dioxide emissions in the united states. *Tellus Series B-Chemical and Physical Meteorology 57*, 2 (—2005—), 107–115.

[7] Boden, T., Marland, G., and Andres, R. Estimates of global, regional, and national annual CO_2 emissions from fossil-fuel burning, hydraulic cement production, and gas flaring: 1950-1992. *ORNL/CDIAC-90, NDP-30/R6* (—1995—).

[8] Bousquet, P., Peylin, P., Ciais, P., Le Quere, C., Friedlingstein, P., and Tans, P. P. Regional changes in carbon dioxide fluxes of land and oceans since 1980. *Science, Amer Assoc Advancement Science 290*, 5495 (—2000—), 1342–1346.

[9] Brenkert, A. L. *Carbon dioxide emission estimates from fossil-fuel burning, hydraulic cement production, and gas flaring for 1995 on a one degree grid cell basis.* Carbon Dioxide Inf. Anal. Cent.,

Oak Ridge Natl. Lab., Oak Ridge, Tenn. (Available at http://cdiac.esd.ornl.gov/ndps/ndp058a.html), —1998—.

[10] British Petroleum, p. l. c. *BP Statistical Review of World Energy*. BP, —2008—.

[11] Ciais, P., Peylin, P., and Bousquet, P. Regional biospheric carbon fluxes as inferred from atmospheric CO_2 measurements. *Ecological Applications, Ecological Soc Amer 10*, 6 (—2000—), 1574–1589.

[12] Cole, M. A., Rayner, A. J., and Bates, J. M. The environmental kuznets curve: an empirical analysis. *Environment and Development Economics 2* (—1997—), 401 – 416.

[13] Edmonds, J., Joos, F., Nakicenovic, N., Richels, R., Sarmiento, Jorge L — Field, C., and Raupach, M. *(eds.), Scenarios, targets, gaps, and costs*. Island Press, —2004—.

[14] Energy Information Administration, E. I. A. International energy outlook, —2008—.

[15] Friedrich, R., Freibauer, A., Gallmann, E., Giannouli, M., Koch, D., Peylin, P., Pye, S., Riviee, E., Sn Hose, R., Winiwartr, W., and Blank, P. Temporal and spatial resolution of greenhouse gas emissions in europe, —2003—.

[16] GLOBALVIEW-CO_2. *Cooperative Atmospheric Data Integration Project: Carbon dioxide*. Global Monit. Div., Earth Syst. Res. Lab., NOAA, Boulder, Colo. (Available at ftp://ftp.cmdl.noaa.gov/ccg/co2/GLOBALVIEW), —2004—.

[17] Gurney, K., Ansley, W., Mendoza, D., Petron, G., Frost, G., Gregg, J., Fischer, M., Pataki, D., and Ackerman, K. Research needs for finely resolved fossil carbon emissions. *Eos, Transactions American Geophysical Union 88*, 49 (—2007—), 52–543.

[18] Gurney, K., Chen, Y., Maki, T., Kawa, S., Andrews, A., and Zhu, Z. Sensitivity of atmospheric CO_2 inversions to seasonal and interannual variations in fossil fuel emissions. *Journal of Geophysical Research - Atmospheres 110*, D10 (—2005—), –.

[19] Gurney, K., Denning, S., Fischer, M., Ojima, D., Marland, G., and Kasibhatla, P. *High-Resolution Fossil Fuel Emissions Estimates in Support of NACP and OCO-Based CO_2 Measurements and Assimilation System*. NCAP, —2005-2008—.

[20] Heimann, M. *The Global Atmospheric Tracer Model TM2: Technical Report No. 10*. Max-Planck-Institut fur Meteorologie Bundesstrasse 55, D-20146 Hamburg, Germany, —1995—.

[21] Houghton, R. A. Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850 2000. *Tellus, Series B 55*, 2 (—2003—), 378 390.

[22] IPCC, Denman, K., Brasseur, G., — Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H. *(eds.), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Chapter 7: Changes in Atmospheric Constituents and in Radiative Forcing.* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, —2007—.

[23] IPCC, Pachauri, R., and Reisinger, A. *(eds.), Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* IPCC, Geneva, Switzerland, —2007—.

[24] IPCC, Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H. *(eds.), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, —2007—.

[25] IPCC/OECD/IEA, Penman, J., Gytarsky, M., Hiraishi, T., Irving, W., and Krug, T. *IPCC Guidelines for National Greenhouse Gas Inventories: Overview.* Integovernmental Panel on Climate Change, —2006—.

[26] Janssens, I. A., Freibauer, A., Ciais, P., Smith, P., Nabuurs, G. J., Folberth, G., Schlamadinger, B., Hutjes, R. W. A., Ceulemans, R., Schulze, E. D., Valentini, R., and Dolman, A. J. Europe's terrestrial biosphere absorbs 7 to 12% of european anthropogenic CO_2 emissions. *Science, Amer Assoc Advancement Science 300*, 5625 (—2003—), 1538–1542.

[27] Joos, F. The atmospheric carbon dioxide perturbation. *Europhysics News 27* (—1996—), 213–218.

[28] Kaminski, T., Giering, R., Scholze, M., Rayner, P., and Knorr, W. An example of an automatic differentiation-based modelling system. *Computational Science and its Applications - Proceedings 2668* (—2003—), 95–104.

[29] Kaminski, T., and Heimann, M. Inverse modeling of atmospheric carbon dioxide fluxes. *Science, Amer Assoc Advancement Science 294*, 5541 (—2001—), U1–U1.

[30] Kaminski, T., and Rayner, P. J. *Assimilation and network design.* Springer-Verlag, —2007—.

[31] Keeling, C. D., Piper, S. C., Bacastow, R. B., Wahlen, M., Whorf, T. P., Heimann, M., and Meijer, H. A. *Exchanges of Atmospheric CO_2 and $^{13}CO_2$ with the Terrestrial Biosphere and Oceans from 1978 to 2000. I. Global Aspects.*, vol. 01-06. Scripps Institution of Oceanography, —2001—.

[32] Kelly, D. L., and Kolstad, C. D. *Integrated Assessment Models For Climate Change Control.* US Department of Energy, —1998—.

[33] Knorr, W. *Doctoral Thesis: Satellite Remote Sensing and Modelling of the Global CO Exchange of Land Vegetation: A Synthesis Study.* Faculty of Earth Sciences of Hamburg University, Max-Planck-Institut für Meteorologie, Examensarbeit Nr. 49., —1997—.

[34] Le Quéré, C. Two decades of ocean CO_2 sink and variability. *Tellus, Series B 55* (—2003—), 649 656.

[35] Levin, I., Kromer, B., Schmidt, M., and Sartorius, H. A novel approach for independent budgeting of fossil fuel CO_2 over europe by (CO_2)-C^{14} observations. *Geophysical Research Letters 30*, 23 (—2003—).

[36] Manne, A. S., and Richels, R. G. *MERGE: An Integrated Assessment Model for Global Climate Change.* Stanford University, EPRI, —2004—.

[37] Marland, G., Andres, R. J., and Boden, T. A. Global, regional, and national CO_2 emissions. *ESD Publication, No. 4195; Trends '93: A compendium of data on global change* (—1994—), 505–584. Boden, T. A.

[38] Marland, G., Boden, T. A., and Andres, R. J. *Global, regional, and national CO_2 emissions, in Trends: A Compendium of Data on Global Change.* Carbon Dioxide Information Analysis Center, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6335, U.S.A., —2006—.

[39] Marland, G., and Rotty, R. Carbon dioxide emissions from fossil fuels: A procedure for estimation and results for 1950-82. *TELLUS 36(B)* (—1984—), 232–61.

[40] Myneni, R. B. *Global Climate Change and Environmental Impacts: Chapter 2. Radiative Forcing of Climate Change.* Climate and Vegetation Research Group, Boston University, —2002—.

[41] Nakicenovic, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grubler, A., Jung, T., and Kram, T. *IPCC Special Report on Emissions Scenarios. Chapter 3: Scenario Driving Forces*. Cambridge Univ Press, Cambridge, UK, —2000—.

[42] Nakicenovic, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grubler, A., Jung, T., and Kram, T. *IPCC Special Report on Emissions Scenarios. Chapter 6.2: Scenario Driving Forces*. Cambridge Univ Press, Cambridge, UK, —2000—.

[43] Nakicenovic, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grubler, A., Jung, T., and Kram, T. *IPCC Special Report on Emissions Scenarios. Chapter 9.1: Carbon Dioxide Emissions*. Cambridge Univ Press, Cambridge, UK, —2000—.

[44] Pardo, A., Meneu, V., and Valor, E. Temperature and seasonality influences on spanish electricity load. *Energy Economics 24*, 1 (—2002—), 55–70.

[45] Prentice, I., Farquhar, G., Fasham, M., Goulden, M., Heimann, M., Jaramillo, V., Kheshgi, H., Le Quere, C., Scholes, R., and Wallace, D. *The carbon cycle and atmospheric carbon dioxide*. Climate Change 2001: The Scientific Basis. Cambridge University Press, Cambridge, —2001—.

[46] Prentice, I., Heimann, M., and Sitch, S. The carbon balance of the terrestrial biosphere: Ecosystem models and atmospheric observations. *Ecological Applications, Ecological Soc Amer 10*, 6 (—2000—), 1553–1573.

[47] Raupach, M., Marland, G., Ciais, P., Le Quere, C., Canadell, J., Klepper, G., and Field, C. Global and regional drivers of accelerating CO_2 emissions. *Proceedings of the National Academy of Sciences of the United States of America 104*, 24 (—2007—), 10288–10293.

[48] Rayner, P., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H. Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (ccdas). *Global Biogechemical Cycles 19*, 2 (—2005—), –.

[49] Roberts, J., and Grimes, P. Carbon intensity and economic development 1962-91: A brief exploration of the environmental kuznets curve. *World Development 25*, 2 (—1997—), 191–198.

[50] Rotty, R. Estimates of seasonal variation in fossil fuel co_2 emissions. *Tellus Series B Chemical and Physical Meteorology B 39* (—1987—), 184–+.

[51] Sailor, D. Relating residential and commercial sector electricity loads to climate - evaluating state level sensitivities and vulnerabilities. *Energy 26*, 7 (—2001—), 645–657.

[52] Scholze, M., Kaminski, T., Rayner, P., Knorr, W., and Giering, R. Propagating uncertainty through prognostic carbon cycle data assimilation system simulations. *Journal of Geophysical Research - Atmospheres 112*, D17 (—2007—), –.

[53] Scholze, M., Rayner, P., Knorr, W., Kaminski, T., Giering, R., and Widmann, H. Non-linear parameter optimisation of a terrestrial biosphere model using atm. co2 observations. *Geophysical Research Abstracts 6*, 06281 (—2004—).

[54] Stern, D. I. *The Environmental Kuznets Curve*. International Society for Ecological Economics, —2003—.

[55] Stern, N. *The Stern Review: The Economics of Climate Change*. Cambridge University Press, —2006—.

[56] Takahashi, T., R., Wanninkhof, H., Feely, R. A., Weiss, R. F., Chipman, D. W., Bates, N., Olafsson, J., Sabine, C., and Sutherland, S. C. Net sea-air CO_2 flux over the global oceans: An improved estimate based on the sea-air pCO_2 difference. *paper presented at 2nd International CO_2 in the Oceans Symposium, Cent. for Global and Environ. Res., Natl. Inst. for Environ. Stud., Tsukuba, Japan, 18 - 22 Jan.* (—1999—).

[57] Weyant, and — Bruce, J. P. *(eds), Climate Change 1995: Economic and Social Dimensions of Climate Change. Integrated Assessment of Climate Change: An Overview and Comparison of Approaches and Results*. Cambridge University Press, Cambridge, 1996.

[58] Wigley, T., and Schimel, D. S. *The Carbon Cycle*. Cambridge University Press, —2000—.

[59] Wu, L., Kaneko, S., and Matsuoka, S. Driving forces behind the stagnancy of china's energy-related CO_2 emissions from 1996 to 1999: the relative importance of structural change, intensity change and scale change. *Energy Policy 33*, 3 (—2005—), 319–335.

APPENDIX

A Observation Stations

Table 3: 41 stations and their geographical co-ordinates, from the GLOBALVIEW CO_2 network used in this study [16]. Station codes are used for ease of reference. Region is classified as the geographically closest region if th station is not positioned directly within a regional boundary.

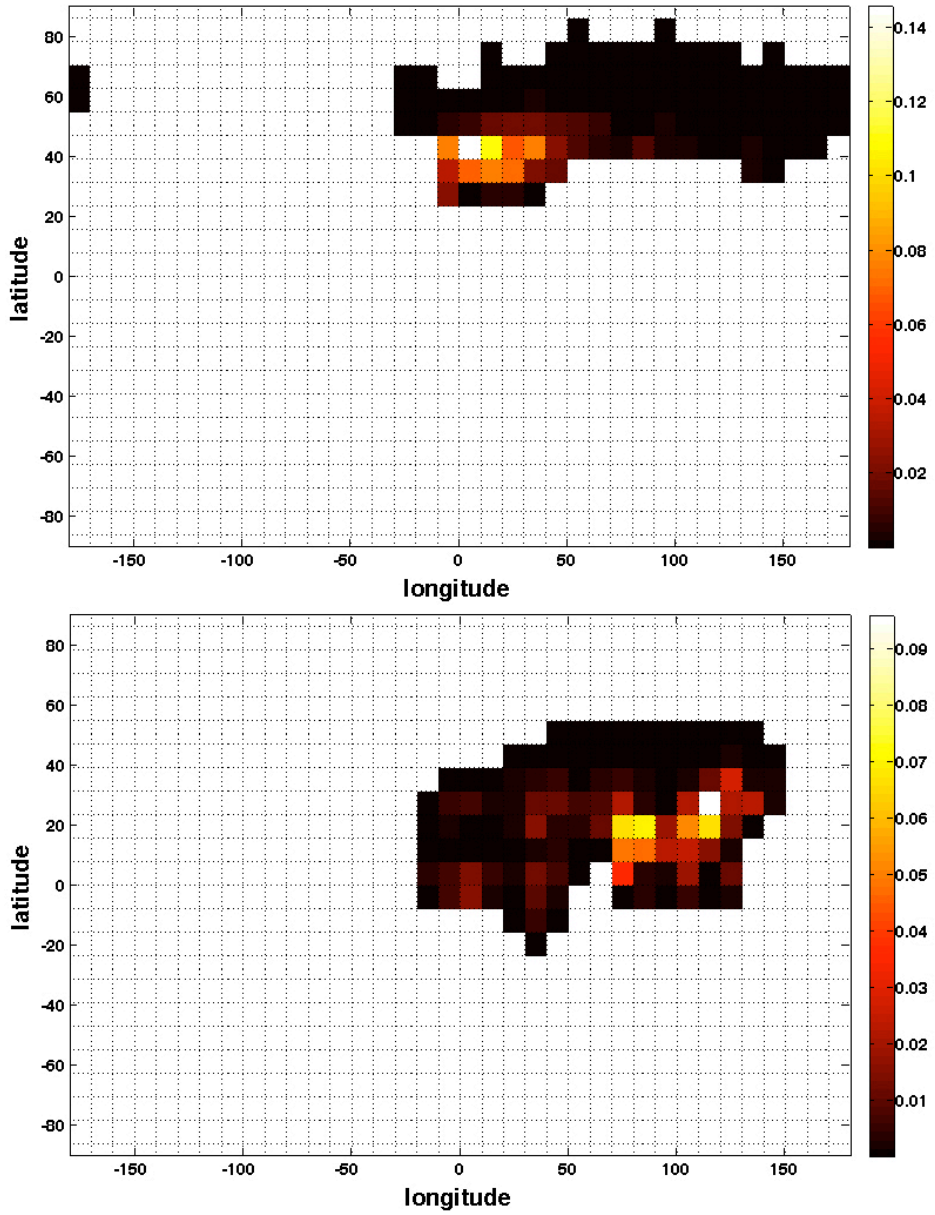
Station Code	Station	Latitude	Longitude
ALT	Alert, Nunavut, Canada	82.45	-62.52
AMS	Amsterdam Island, France	-37.95	77.53
ASC	Ascension Island, U.K.	-7.92	-14.42
AVI	St. Croix, Virgin Islands, U.S.A.	17.75	-64.75
AZR	Terceira Island, Azores, Portugal	38.77	-27.38
BAL	Baltic Sea, Poland	55.35	17.22
BME	St. David's Head, Bermuda, U.K.	32.37	-64.65
BMW	Southhampton, Bermuda, U.K	32.27	-64.88
BRW	Barrow, Alaska, U.S.A.	71.32	-156.6
CBA	Cold Bay, Alaska, U.S.A.	55.2	-162.72
CGO	Cape Grim, Tasmania, Australia	-40.68	144.68
CHR	Christmas Island, Kiribati	1.7	-157.17
CMO	Cape Meares, Oregon, U.S.A.	45.48	-123.97
CRZ	Crozet, Indian Ocean, France	-46.45	51.85
EIC	Easter Island, Chile	-27.15	-109.45
GMI	Guam, Mariana Islands, U.S.A.	13.43	144.78
HBA	Halley Bay, Antarctica, U.K.	-75.58	-26.5
ICE	Storhofdi, Heimaey, Vestmannaeyjar, Iceland	63.34	-20.29
IZO	Tenerife, Canary Islands, Spain	28.3	-16.48
KEY	Key Biscayne, Florida, U.S.A.	25.67	-80.2
KUM	Cape Kumukahi, Hawaii, U.S.A.	19.52	-154.82
MBC	Mould Bay, Nunavut, Canada	76.25	-119.35
MHD	Mace Head, County Galway, Ireland	53.33	-9.9
MID	Sand Island, Midway, U.S.A.	28.21	-177.38
MLO	Mauna Loa, Hawaii, U.S.A.	19.53	-155.58
NWR	Niwot Ridge, Colorado, U.S.A.	40.05	-105.58
OPW	Olympic Peninsula, Washington, U.S.A.	48.25	-124.42
PSA	Palmer Station, Antarctica, U.S.A.	-64.92	-64
RPB	Ragged Point, St. Phillip's Parish, Barbados	13.17	-59.43
SEY	Mahe Island, Seychelles	-4.67	55.17
SHM	Shemya Island, Alaska, U.S.A.	52.72	174.1
SMO	Tutuila, American Samoa, U.S.A.	-14.24	-170.57
SPO	South Pole, Antarctica, U.S.A.	-89.98	-24.8
STM	Atlantic Ocean (Polarfront), Norway	66	2
SYO	Syowa Station, Antarctica, Japan	-69	39.58
TAP	Tae-Ahn Peninsula, Korea	36.73	126.13
UTA	Wendover, Utah, U.S.A.	39.9	-113.72
UUM	Ulaan Uul, Mongolia	44.45	111.1
WIS	Sede Boker (Negev Desert), Israel	31.13	34.88
WLG	Mt. Waliguan Baseline Observatory, Peoples Republic of China	36.29	100.9
ZEP	Zeppelin Station, Ny-Alesund, Svalbard, (Spitsbergen), Norway	78.9	11.88

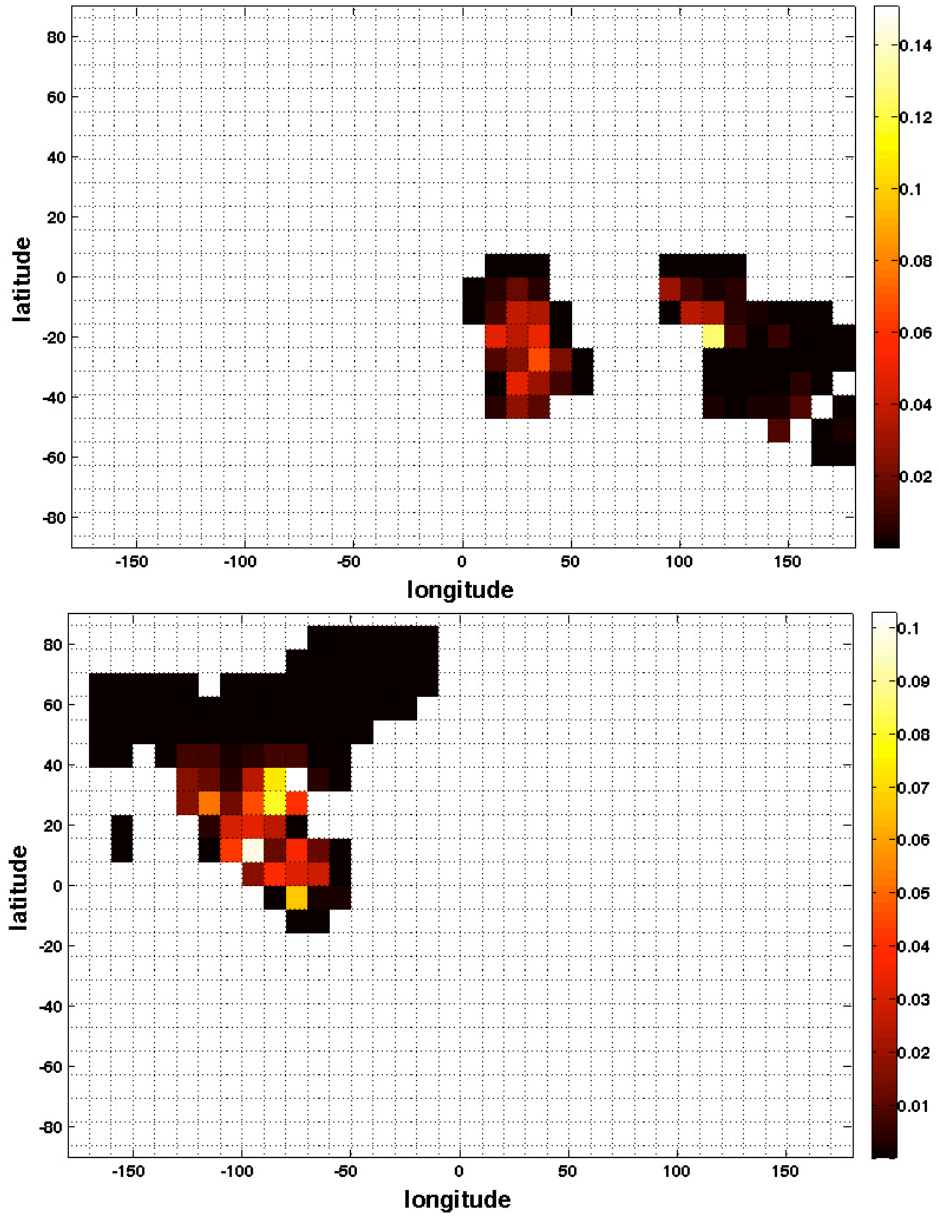
B Regional Division of Countries

Table 4: List of countries from **CDIAC** emissions data set [38], divided into regions. Some countries, whilst geographically covering the same area have different names historically. All variations are included as they appear in the data set. The top 20 emitters are denoted by *****.

Region 1	Region 2	Region 3	Region 4	Region 5
Albania	Afghanistan	American Samoa	Anguilla	Argentina
Andorra	Algeria	Angola	Antigua and Barbuda	Bolivia
Austria	Armenia	Australia *	Aruba	Brazil *
Belarus	Azerbaijan	Botswana	Bahamas	Chile
Belgium	Bahrain	Brunei Darussalam	Barbados	Ecuador
Bosnia Herzegovina	Bangladesh	Burundi	Belize	Falkland Islands (Malvinas)
Bulgaria	Benin	Comoros	Bermuda	French Polynesia
Croatia	Bhutan	Congo	British Virgin Islands	Paraguay
Cyprus	Burkina Faso	Congo, Democratic Republic	Canada *	Peru
Czech Republic	Cote d'Ivoire	Cook Islands	Cayman Islands	Pitcairn
Denmark	Cambodia	East Timor	Colombia	Uruguay
Estonia	Cameroon	Fiji	Costa Rica	
Faeroe Islands	Cape Verde	Gabon	Cuba	
Finland	Central African Republic	Indonesia *	Dominica	
France *	Chad	Kiribati	Dominican Republic	
Germany *	China *	Lesotho	El Salvador	
Gibraltar	Djibouti	Madagascar	French Guiana	
Greece	Egypt	Malawi	Greenland	
Holy See	Equatorial Guinea	Mauritius	Grenada	
Hungary	Eritrea	Mozambique	Guadeloupe	
Iceland	Ethiopia	Namibia	Guatemala	
Ireland	Federated States of Micronesia	New Caledonia	Guyana	
Italy *	Gambia	New Zealand	Haiti	
Latvia	Georgia	Niue	Honduras	
Liechtenstein	Ghana	Papua New Guinea	Jamaica	
Lithuania	Guam	Reunion	Martinique	
Luxembourg	Guinea	Rwanda	Mexico *	
Malta	Guinea-Bissau	Samoa	Montserrat	
Monaco	Hong Kong SAR (China)	Seychelles	Nicaragua	
Netherlands	India *	Solomon Islands	Northern Mariana Islands	
Netherlands Antilles	Iran (Islamic Republic of) *	South Africa *	Panama	
Norway	Iraq	St. Helena	Puerto Rico	
Poland	Israel	Swaziland	Saint Kitts and Nevis	
Portugal	Japan *	Tonga	Saint Lucia	
Republic of Moldova	Jordan	Tuvalu	St. Pierre and Miquelon	
Romania	Kazakstan	United Rep. of Tanzania	St. Vincent and the Grenadines	
Russian Federation *	Kenya	Vanuatu	Suriname	
San Marino	Korea, Dem. People's Rep. of	Wallis and Futuna Islands	Trinidad and Tobago	
Slovakia	Korea, Republic of *	Zambia	Turks and Caicos Islands	
Slovenia	Kuwait	Zimbabwe	United States of America *	
Spain *	Kyrgyzstan		United States Virgin Islands	
Svalbard (Norway)	Lao People's Dem. Rep.		Venezuela	
Sweden	Lebanon			
Switzerland	Liberia			
The former Yugoslav Rep. of Macedonia	Libyan Arab Jamahiriya			
Ukraine *	Macau			
United Kingdom *	Malaysia			
Yugoslavia	Maldives			
	Mali			
	Marshall Islands			
	Mauritania			
	Mongolia			
	Morocco			
	Myanmar			
	Nauru			
	Nepal			
	Neutral zone (IRQ and SAU)			
	Niger			
	Nigeria			
	Oman			
	Pakistan			
	Palau			
	Palestinian National Authority			
	Philippines			
	Qatar			
	Sao Tome and Principe			
	Saudi Arabia *			
	Senegal			
	Sierra Leone			
	Singapore			
	Somalia			
	Sri Lanka			
	Sudan			
	Syrian Arab Republic			
	Tajikistan			
	Thailand			
	Togo			
	Tunisia			
	Turkey			
	Turkmenistan			
	Uganda			
	United Arab Emirates			
	Uzbekistan			
	Vietnam			
	Western Sahara			
	Yemen			

C Population Fraction Distribution




latitude
longitude
80
60
40
20
0
-20
-40
-60
-80
-150
-100
-50
0
50
100
150
0.14
0.12
0.1
0.08
0.06
0.04
0.02
latitude
longitude
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01

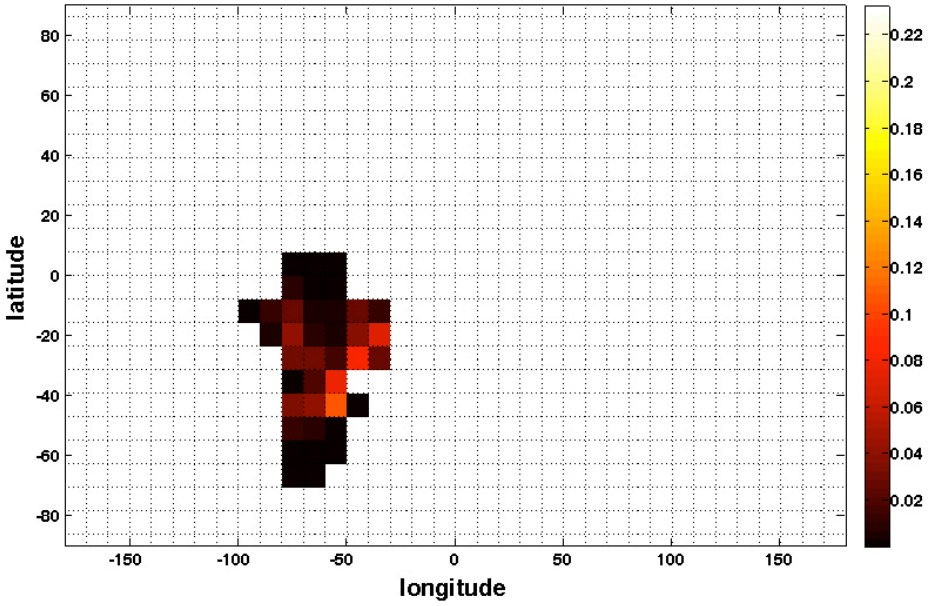


Figure 22: Population fraction ($W^r_{i,j}$) spatial distribution for all regions 1 to 5 (top - bottom)) on $\sim 8^\circ \times 10^\circ$ (TM2) resolution. Darker areas indicate low population fraction of the region, whereas those lighter in tone are indicative of highly populated areas.

D Fit Statistics

D.1 Deconstruction of Atmospheric CO_2 Concentrations

Within atmospheric CO_2 concentration time series (either model output, or from observational data — $M(\mathbf{x}_p)$ and $\mathbf{d}$ respectively), there are two main components — an inter-annual trend and a seasonal cycle. In order to estimate the changes in both of these due to the addition of seasonal fossil fuel emissions, the two are separated. A *running mean* from the concentration data is calculated over 24 months (or occasionally 48 if the seasonality is very strong) and removed. The mean seasonal cycle is then calculated from what remains. Comparison may then be made between both models' relative fit to the observations. Fig. 23 shows an example of this deconstruction technique.

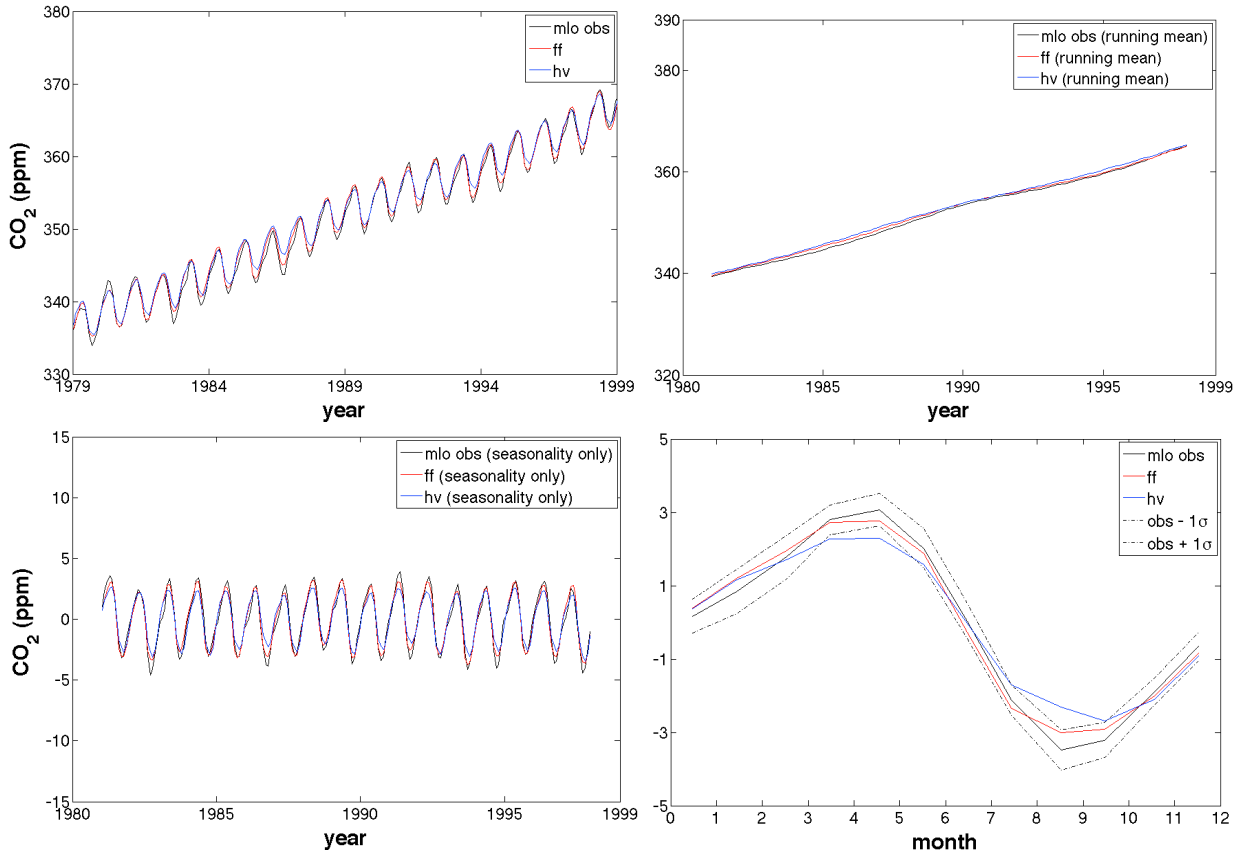


Figure 23: From top-left to bottom-right: Atmospheric CO_2 concentration data from observations (black), and model output (**hv** in blue, **ff**in red); Running mean; Seasonality left after removal of running mean; Average annual seasonal cycle with dotted lines indicating one standard distribution (1σ) from the mean observational seasonal cycle.

D.2 R^2 Statistic

Two methods of fit estimation are used, both implementing the R^2 technique as in Eq. 16.

$$R^2 = 1 - \frac{\sum_m (\mathbf{d}_m - \bar{\mathbf{d}})^2}{\sum_m (\mathbf{d}_m - M(\mathbf{x}_p)_m)^2} \tag{16}$$

For the first, an estimate of the fit to the whole time series of observations is made relative to the running mean of the observations. In this case, $\bar{\mathbf{d}}$ is the running mean, m denotes a sum over all months in the calibration period (those remaining after calculation of the running mean), $\mathbf{d}$ is the observational data (atmospheric CO_2 concentration), $M(\mathbf{x}_p)$ is the model output of atmospheric CO_2 concentration after calibration (both in ppm).

In the second case, fit of the mean seasonal cycle is made. $\bar{\mathbf{d}}$ is the mean annual atmospheric CO_2 concentration over the whole calibration period, m denotes a sum over all months in the year, $\mathbf{d}$ is the mean seasonal cycle of the observational data (atmospheric CO_2 concentration, with the running mean removed), $M(\mathbf{x}_p)$ is the mean monthly model output of atmospheric concentration after calibration (both in ppm).

If the fit is good, the R^2 value should be near 1. If the R^2 value is negative, there is a significant lag in the seasonal cycle of the model output relative to observations. A higher negative value suggests a phase shift in the seasonal cycle away from that in observations, but any improvement towards a positive value (or change to a positive value) indicates a phase shifting towards that in observations.

D.3 % Improvement in Fit

Any improvement in the fit of the mean seasonal cycle is quantified by a percentage increase from the fit from the **hv** output towards 1 (a perfect fit), as in Eq. 17, where R^2_{hv} and R^2_{ff} are the R^2 fits for the **hv** and **ff** model output respectively.

$$\Gamma = \frac{(1 - R^2_{\mathbf{hv},s}) - (1 - R^2_{\mathbf{ff},s})}{1 + |R^2_{\mathbf{hv},s}|} \times 100 \tag{17}$$

Any small value for Γ is likely to be encompassed in the uncertainties in the model output ($M(\mathbf{x}_p)$), and the distribution around the mean, and so is disregarded.

E Additional Results

E.1 Regional Flux

Table 5: Mean annual fossil fuel CO_2 flux per region (and uncertainty representing one standard deviation from the mean) as output from **ff** model, and from CDIAC data set [38]. Percentage contribution of each region to the total emissions of the model output.

Region	Flux ($GtCyr^{-1}$)	% of total emissions	CDIAC flux ($GtCyr^{-1}$
1	1.9848 ± 0.2541	34.73	1.9834 ± 0.26
2	1.7503 ± 0.4855	30.63	1.7525 ± 0.49
3	0.2364 ± 0.0544	4.14	0.2368 ± 0.05
4	1.6265 ± 0.1587	28.46	1.6279 ± 0.16
5	0.1166 ± 0.0203	2.04	0.1169 ± 0.02

E.2 Fit to Data

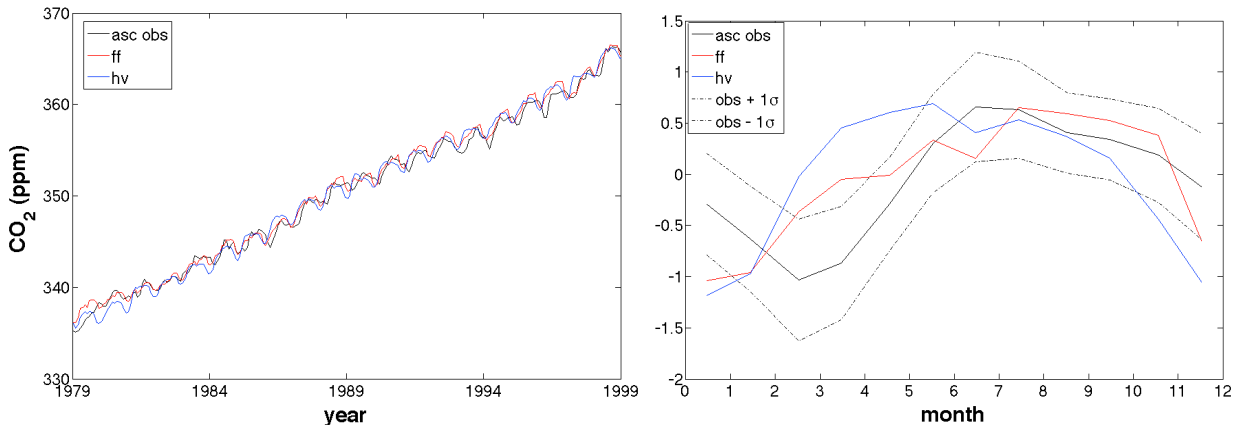


Figure 24: Station = ASC (Ascension Island). Black = observational data, blue = **hv** model output, red = **ff** model output. Left: monthly atmospheric CO_2 concentration, over model calibration period. Right: running mean (inter-annual trend) removed, and mean seasonal cycle constructed over the central 19 years of calibration. Dashed lines signify 1σ of distribution either side of the mean.

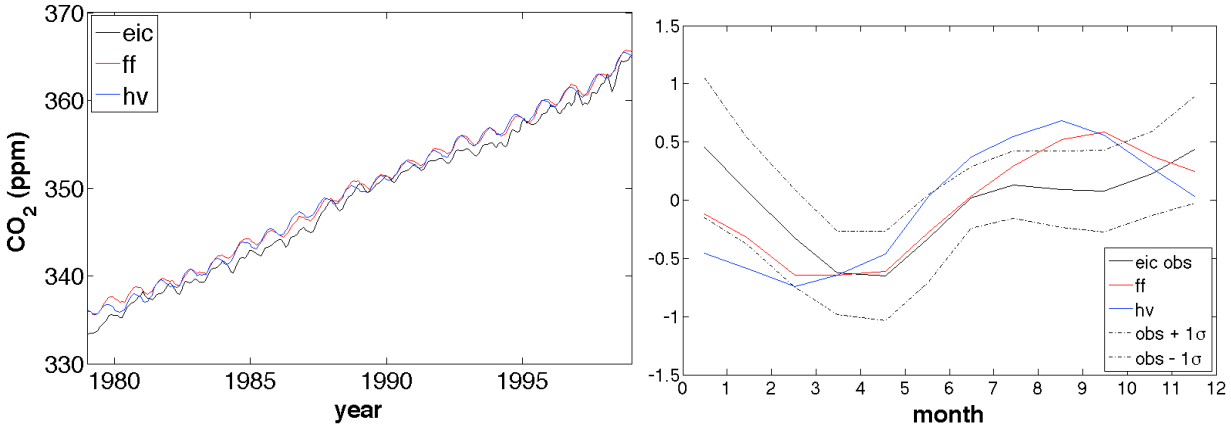


Figure 25: Station = EIC (Easter Island). Black = observational data, blue = **hv** model output, red = **ff** model output. Left: monthly atmospheric CO_2 concentration, over model calibration period. Right: running mean (inter-annual trend) removed, and mean seasonal cycle constructed over the central 19 years of calibration. Dashed lines signify 1σ of distribution either side of the mean.

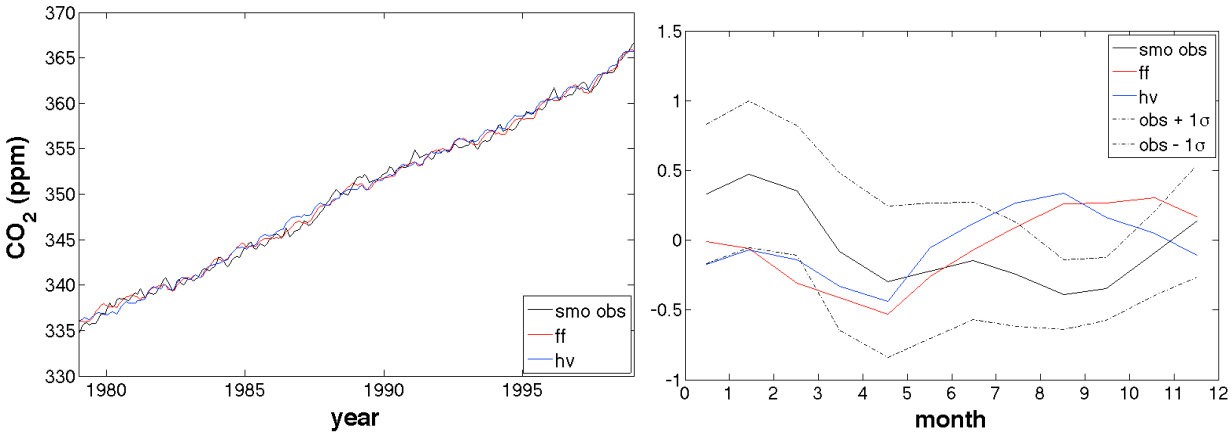


Figure 26: Station = SMO (Tutuila, American Samoa). Black = observational data, blue = **hv** model output, red = **ff** model output. Left: monthly atmospheric CO_2 concentration, over model calibration period. Right: running mean (inter-annual trend) removed, and mean seasonal cycle constructed over the central 19 years of calibration. Dashed lines signify 1σ of distribution either side of the mean.

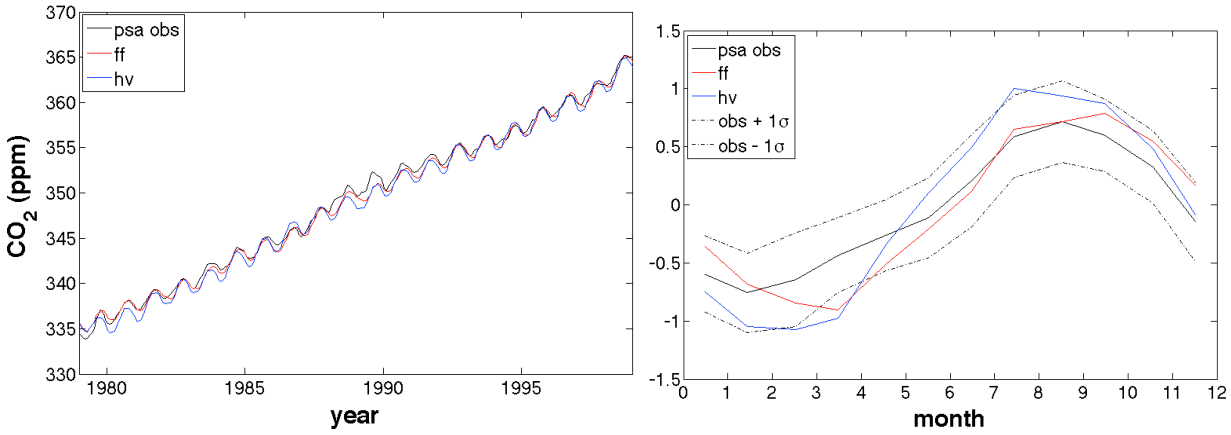
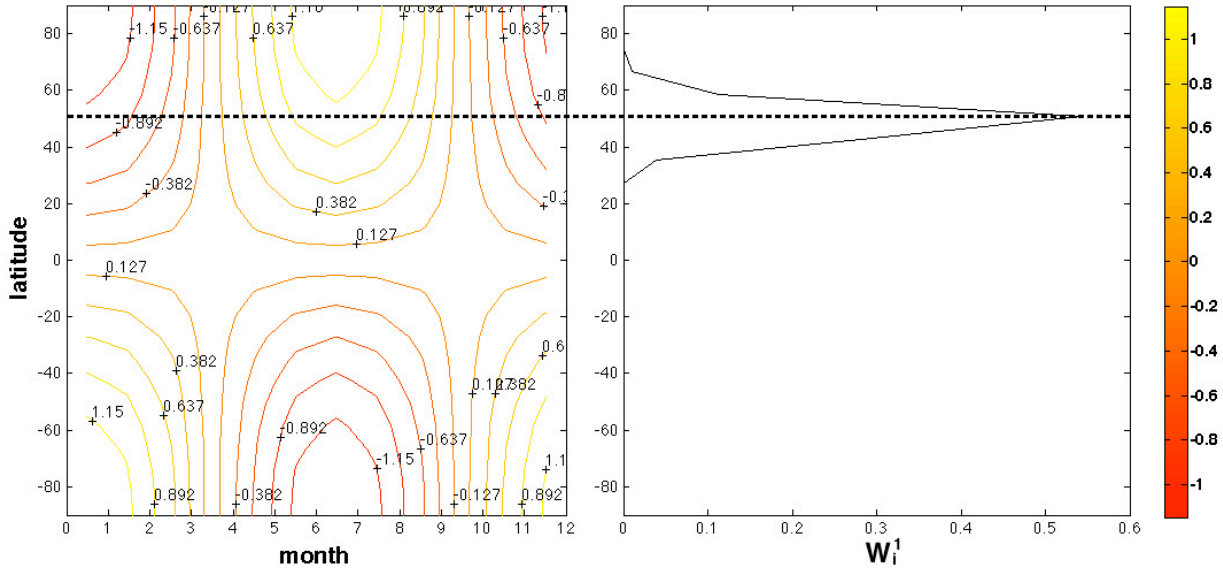
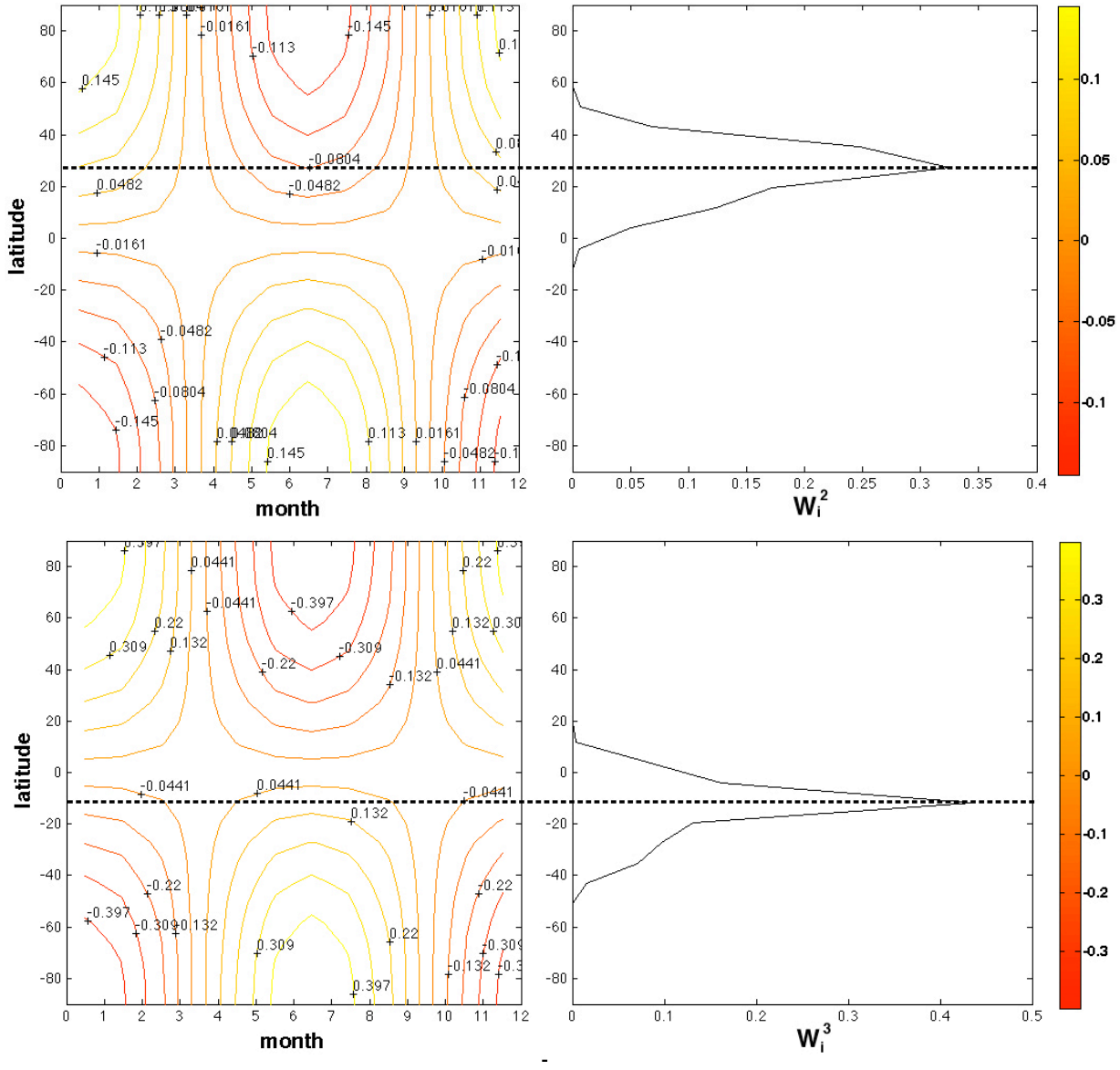


Figure 27: Station = PSA (Palmer Station, Antarctica). Black = observational data, blue = **hv** model output, red = **ff** model output. Left: monthly atmospheric CO_2 concentration, over model calibration period. Right: running mean (inter-annual trend) removed, and mean seasonal cycle constructed over the central 19 years of calibration. Dashed lines signify 1σ of distribution either side of the mean.

E.3 Seasonality and Population Weighting




latitude
month
W_i^2
latitude
month
W_i^3

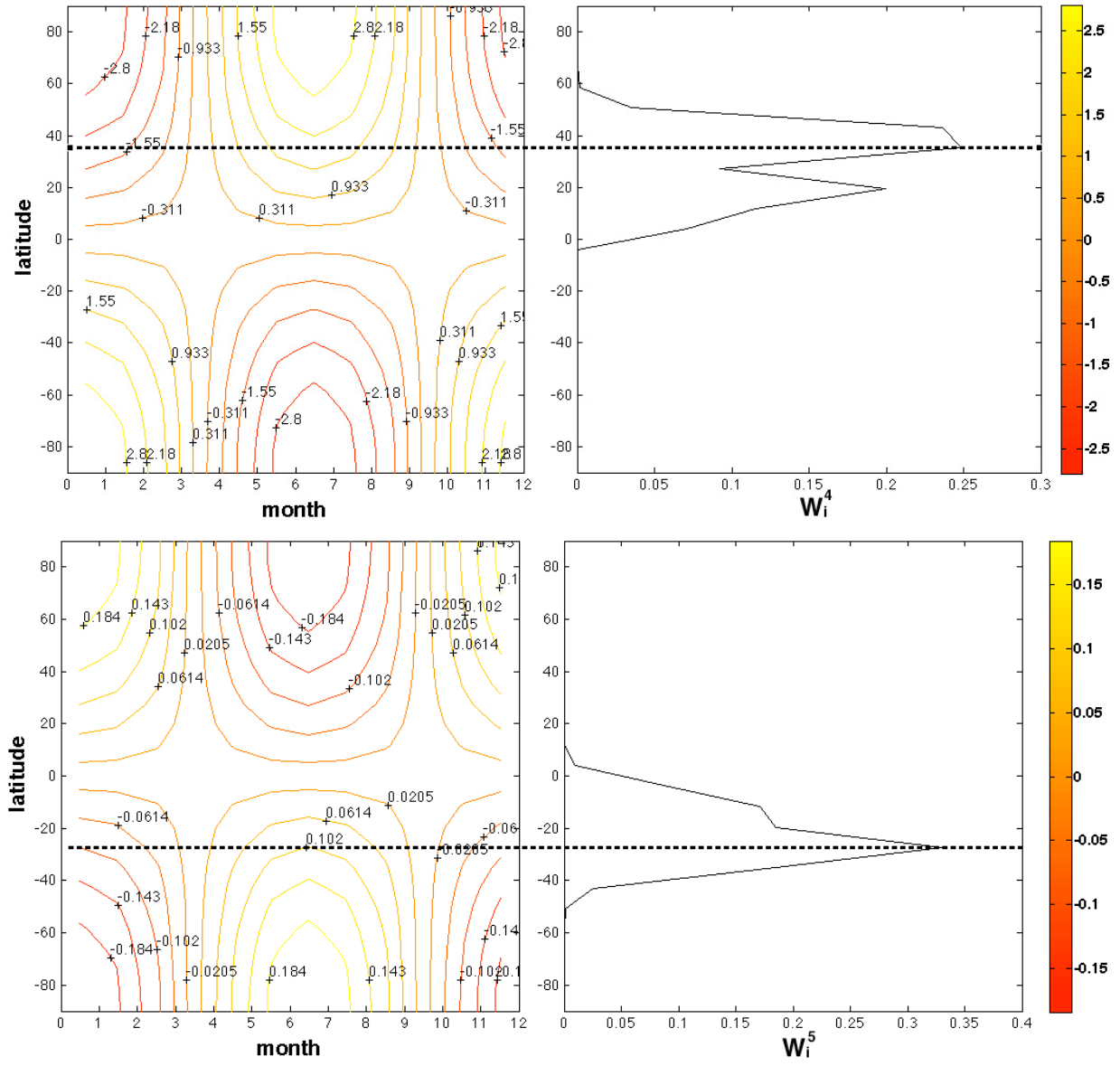


Figure 28: Contour plots of latitudinal and monthly function dependence of seasonality ($F^r_{i,j,m}$) and corresponding latitudinal distribution of population fraction ($\sum_j W^r_{i,j}$) for regions 1 to 5 (top to bottom). Dotted lines represents the latitudinal value taken to be *most influential* to the seasonal cycle of that region (i.e. the overall seasonal cycle will be weighted with regards to the latitudinal distribution of the population fraction).