SPECIAL
C4MIP
COLLECTION

Sources of Uncertainty in Future Projections of the Carbon Cycle

ALAN J. HEWITT, BEN B. B. BOOTH, CHRIS D. JONES, EDDY S. ROBERTSON,
AND ANDY J. WILTSHIRE

Met Office Hadley Centre, Exeter, United Kingdom

PHILIP G. SANSOM AND DAVID B. STEPHENSON

Exeter Climate Systems, University of Exeter, Exeter, United Kingdom

STAN YIP

Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, China

ABSTRACT

The inclusion of carbon cycle processes within CMIP5 Earth system models provides the opportunity to explore the relative importance of differences in scenario and climate model representation to future land and ocean carbon fluxes. A two-way analysis of variance (ANOVA) approach was used to quantify the variability owing to differences between scenarios and between climate models at different lead times. For global ocean carbon fluxes, the variance attributed to differences between representative concentration pathway scenarios exceeds the variance attributed to differences between climate models by around 2025, completely dominating by 2100. This contrasts with global land carbon fluxes, where the variance attributed to differences between climate models continues to dominate beyond 2100. This suggests that modeled processes that determine ocean fluxes are currently better constrained than those of land fluxes; thus, one can be more confident in linking different future socioeconomic pathways to consequences of ocean carbon uptake than for land carbon uptake. The contribution of internal variance is negligible for ocean fluxes and small for land fluxes, indicating that there is little dependence on the initial conditions. The apparent agreement in atmosphere–ocean carbon fluxes, globally, masks strong climate model differences at a regional level. The North Atlantic and Southern Ocean are key regions, where differences in modeled processes represent an important source of variability in projected regional fluxes.

1. Introduction

The global carbon cycle is a crucial component of future climate change, closely linking anthropogenic CO_2 emissions with future changes in atmospheric CO_2 concentration and hence climate (Denman et al. 2007; Ciais et al. 2013). Inclusion of the carbon cycle as an interactive component in comprehensive Earth system models (ESMs) has grown since early coupled studies (Cox et al. 2000) and intercomparisons such as the Coupled Carbon Cycle–Climate Model Intercomparison Project (C4MIP; Friedlingstein et al. 2006) and is now a

mainstream component of coordinated climate simulations like phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012).

Such coupled climate–carbon cycle ESMs simulate the natural exchange of carbon by the land and ocean with the atmosphere and thus provide a predictive link between emissions and atmospheric concentrations of CO_2. They can be used to compute the emissions required to follow a prescribed concentration pathway (Jones et al. 2006; Plattner et al. 2008). This method has become widespread and was recommended by Hibbard et al. (2007) as the experimental design for CMIP5 and has subsequently been used to present compatible emissions from the CMIP5 multimodel ensemble (Jones et al. 2013).

The natural uptake of carbon by land and ocean biospheres is sensitive to both changes in climate and

Corresponding author address: Alan J. Hewitt, Met Office Hadley Centre, Fitzroy Road, Exeter EX1 3PB, United Kingdom. E-mail: alan.j.hewitt@metoffice.gov.uk

the concentration of atmospheric CO_2 and the balance between these large but offsetting effects on decade–century time scales is not well constrained by observations. Comparison between the C^4MIP ESMs showed large quantitative uncertainty in the future projections of carbon uptake (Friedlingstein et al. 2006), similarly seen in perturbed parameter simulations (Booth et al. 2012). This spread of results across global climate models (GCMs) has not reduced substantially in CMIP5 (Arora et al. 2013; Jones et al. 2013).

To improve the understanding of future land–atmosphere and ocean–atmosphere exchange of carbon, it is imperative to attribute the variation in these fluxes to their component sources. "GCM variability" originates from an incomplete understanding of physical processes including both climate and ecosystem processes, involved in air–surface carbon exchange and from the limitation of GCMs to represent known behavior. "Scenario variability" arises from uncertainty in future human activity; socioeconomic storylines of population and technology growth are produced by integrated assessment models (IAMs; see van Vuuren et al. 2011b) to provide plausible scenarios of future anthropogenic activity such as energy use (and hence fossil fuel emissions) and land-use change. The "internal variability" of a given GCM represents the natural variability of the climate at daily to multidecadal time scales (Karoly and Wu 2005), owing to the chaotic and nonlinear nature of the carbon flux processes; this variability has been long observed even in a stationary climate (Madden 1976). There also exists the possibility of "GCM–scenario interaction" if the differences in simulated climate between GCMs vary between scenarios.

The aim of this study is to quantify the relative importance with time of GCM and scenario variability and to estimate the future time beyond which scenario variability dominates GCM variability. The importance of the GCM–scenario interaction term will be quantified as a tool to understanding the response of GCMs to different scenarios.

Analysis of variance (ANOVA; Gelman 2005) has been used in several climate studies for quantifying sources of variability (von Storch and Zwiers 2001; Tingley 2012) and more recently has been effectively used to diagnose variability in multimodel ensembles (Yip et al. 2011; Sansom et al. 2013; Hingray and Saïd 2014). The opportunistic nature of the analysis of the CMIP5 carbon fluxes has resulted in a varying number of runs for each GCM–scenario pair (some of which have zero runs)—thus an unbalanced factorial design. An ANOVA method appropriate for this data (Northrop and Chandler 2014) has been used to partition the different sources of variability.

a. RCP scenarios

A set of four representative concentration pathways (RCPs) were developed to provide a common set of future climate scenarios to the scientific community, which would allow for better comparisons between ESM/GCM studies and for ease of communication of GCM results (van Vuuren et al. 2011b). The four RCP scenarios—RCP8.5 (Riahi et al. 2011), RCP6.0 (Masui et al. 2011), RCP4.5 (Thomson et al. 2011), and RCP2.6 (van Vuuren et al. 2011a)—lead to an approximate increase in global radiative forcing by the year 2100 of 8.5, 6.0, 4.5, and 2.6 W m^{-2}, respectively. The scenarios are sufficiently separated in terms of the radiative forcing pathways to provide distinguishable climate results at the global scale (Moss et al. 2010). The RCP scenarios have a harmonized historical period, assumptions for carbon emissions and concentrations, and land-use change.

The dominant driver of future radiative forcing for each RCP is the CO_2 concentration pathway (Fig. 1a) along with the fossil fuel emissions associated with that pathway (Fig. 1b) from each IAM that generated the scenario. According to the concentration-driven experiment design in CMIP5 (Taylor et al. 2012), each GCM performs simulations from a preindustrial state (typically representative of 1850) up to 2100, using these CO_2 concentrations as a boundary condition to force the GCM.

Such results are of relevance to policy decisions that aim to achieve a given climate target, but they are subject to large uncertainty (Jones et al. 2013). We want to understand the causes of this uncertainty in compatible emissions. The scenarios are designed to be different—they were selected from hundreds of possible scenarios and approximately span the 10th to 90th percentiles of future radiative forcing across published scenarios. They represent very different societal choices around climate targets and how to achieve them. Hence, it is desirable that the consequences of these choices can be distinguished in terms of their impacts on global and regional climate and ecosystems. We might therefore (by the year 2100) expect the differences between scenarios to be bigger than the differences between GCMs whose aim is to represent the same processes (Cox and Stephenson 2007). For example, the variability in global average temperature by 2100 was greater across scenarios than between GCMs for the CMIP3 ESMs running the SRES family of scenarios (Hawkins and Sutton 2009). However, at regional scales this is not always true (e.g., over the British Isles by 2100; Hawkins and Sutton 2009). Our study has explored the variability in carbon cycle behavior between different GCMs and scenarios, along with their
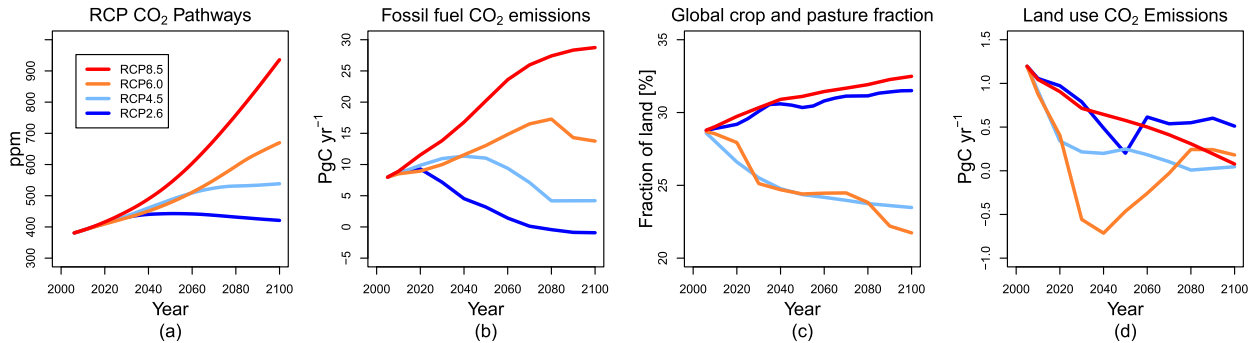
FIG. 1. Shown for each RCP scenario are future (a) atmospheric CO_2 concentration pathway, (b) fossil fuel CO_2 emissions, (c) crop and pasture land fraction (i.e., the anthropogenic land-use change), and (d) anthropogenic land-use change CO_2 emissions [(a)–(c) courtesy of Jones et al. (2013)].

interaction, and how this varies by process, by region and through time.

b. Land-use change in the RCP scenarios

While both ocean and land carbon fluxes respond to change in climate and CO_2, terrestrial carbon storage is additionally influenced by direct anthropogenic activity to modify vegetation cover. The RCPs include changes in anthropogenic land use during the twenty-first century, and the ESMs attempt to simulate this although with varying degrees of complexity and completeness of process representation (Hurtt et al. 2011; Jones et al. 2013). Differences in land-use change contribute to the spread of results across scenarios, and differences in the representation of land use contribute to the spread between ESMs.

The degree of land-use change in the scenarios is not closely related to global radiative forcing, and so the relative importance of land use in the global carbon balance changes in time differently between scenarios. Figure 1c shows the global-scale evolution of land use in the scenarios in terms of fraction of land given over to agriculture (crop and pasture). The scenarios differentiate between agricultural land for crops and for pasture, whereas GCMs differ in how they treat these two classes of land. A common feature is that conversion of forest land to agricultural land involves the removal of large amounts of carbon (as tree biomass). Subsequent changes of soil carbon and the response of carbon storage to management practices differ between ESMs. This might be expected to lead to a strong GCM–scenario interaction and will be explored explicitly in the results section.

Global time series up to 2100 hide the time evolution of land use at regional scales, which can differ markedly between regions. Unlike global climate and CO_2, land use in the RCPs changes most markedly in the early decades of the twenty-first century. We expect the carbon cycle impact from CO_2 and climate change to continue to increase under most scenarios throughout the twenty-first century and perhaps to be more similar across scenarios during the early decades. For prescribed land-use change, however, it is the case that the scenarios diverge rapidly and differ most in the early decades with reduced land-use forcing by the end of the century (Fig. 1d).

2. Data

The monthly fields of carbon mass flux out of atmosphere due to net biospheric production on land (output variable name nbp) and surface downward CO_2 flux (output variable name fgco2) from seven GCMs participating in CMIP5 have been extracted for this work (Taylor 2013; CCCma 2015; Dunne et al. 2014; Sanderson et al. 2014; Denvil et al. 2016; JAMSTEC et al. 2015; Giorgetta et al. 2012; Bentsen et al. 2012). These GCMs were selected as they had good coverage of the four future scenarios and/or multiple runs for each GCM–scenario pair. Where a particular GCM has been designed with and without coupled atmospheric chemistry (e.g., MIROC-ESM and MIROC-ESM-CHEM), only one version of the GCM was included; it was felt that including multiple versions of a particular GCM could artificially alter the GCM variance. The seven GCMs included in the analysis are listed in Table 1.

TABLE 1. Number of runs of each GCM–scenario pair, made available for download. (Acronym expansions are available online at http://www.ametsoc.org/PubsAcronymList.)

GCM	RCP scenario			
	2.6	4.5	6.0	8.5
CanESM2	5	3	0	4
GFDL-ESM2G	1	1	1	1
HadGEM2-ES	4	4	4	4
IPSL-CM5A-LR	4	2	1	4
MIROC-ESM	1	1	1	1
MPI-ESM-LR	3	3	0	2
NorESM1-ME	1	1	1	1

In the analysis we have aggregated the monthly fields into decadal atmosphere–ocean carbon fluxes ($GtC\,decade^{-1}$) and decadal atmosphere–land carbon fluxes ($GtC\,decade^{-1}$); the rational for the choice of time step is given in the discussion (section 5). The first (e.g., decadal atmosphere–land carbon flux) time step was obtained by summing all the monthly (e.g., fgco2) fields over the area of interest for the period 2006–15. The decadal atmosphere–land carbon fluxes can also be thought of as the decadal change in stored carbon. In the analysis we have used raw projection carbon flux data as opposed to change projection data; this is justified in section 5.

The data will be analyzed for the period 2006–95 where all of the simulations listed in Table 1 contain data (the historical period in the GCMs covers the years 1861–2005, and the future period continues to at least 2100 and beyond for some of the GCMs).

3. Methodology

The unbalanced design of the CMIP5 experiment, having different numbers of runs for each combination of GCM i and scenario j, complicates the application of classical ANOVA techniques (e.g., Yip et al. 2011). Unbalanced designs can be addressed using multiple regression (Searle 1987; Sansom et al. 2013). However, the variance attributed to each component depends on the order that the components are entered in to the regression (Davison 2003, chapter 8.5). Therefore, we follow Northrop and Chandler (2014) and use a random effects ANOVA to accommodate the unbalanced design.

Let Y_{ijk} represent a climate variable of interest simulated in run k of scenario j by GCM i. The CMIP5 data analyzed here contain results from seven GCMs and four scenarios. The number of runs k of each scenario by each GCM varies between 0 and 5 (see Table 1). The ANOVA model has the following form:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}, \tag{1}$$

where μ is the mean carbon flux over all GCMs and scenarios. The effect α_i represents the expected difference between the flux simulated by GCM i and the mean μ, over all scenarios. Similarly, β_j represents the expected difference between the flux simulated in scenario j and the mean μ, over all GCMs. The term γ_{ij} represents any GCM-specific response to a particular scenario; that is, the response to a particular scenario might vary between GCMs. The residual ε_{ijk} represents variations between different runs k of the same scenario by the same GCM (i.e., internal variability).

The GCM differences α_i are modeled as independent, identically distributed normal random variables with mean

zero and variance σ_G^2—that is, $\alpha_i \overset{\text{iid}}{\sim} N(0, \sigma_G^2)$. The scenario differences β_j, interaction terms γ_{ij}, and departures owing to internal variability ε_{ijk} are all modeled similarly—that is, $\beta_j \overset{\text{iid}}{\sim} N(0, \sigma_S^2)$, $\gamma_{ij} \overset{\text{iid}}{\sim} N(0, \sigma_{GS}^2)$, and $\varepsilon_{ijk} \overset{\text{iid}}{\sim} N(0, \sigma_R^2)$.

The interpretation of this framework is that the GCM effects α_i, scenario effects β_j, interaction effects γ_{ij}, and internal variability ε_{ijk} are samples from some larger superpopulations (Stephenson et al. 2012). Therefore, the variances σ_G^2, σ_S^2, σ_{GS}^2, and σ_R^2 are referred to as the superpopulation variances. These variances provide the desired partitioning of variability by quantifying the variability attributed to each component (i.e., GCMs, scenarios, internal variability, and GCM–scenario interaction).

It is also possible to compute variances for the specific sample of GCMs and scenarios being analyzed—for example, $s_G^2 = (M-1)^{-1} \sum_i (\alpha_i - \bar{\alpha})^2$, where M is the number of models. These are known as "finite population" variances.

We are primarily interested in the superpopulation variances since there are many other climate models and scenarios we could consider if data were available. However, the finite-population variances provide a useful alternative view from the current sample.

We follow Northrop and Chandler (2014) and take a Bayesian approach to estimating the population variances. In a Bayesian analysis, we are required to specify our prior beliefs about the quantities of interest. We then update those beliefs after observing the data (i.e., the CMIP5 runs). A vague normal prior (normal with large variance) was specified for the overall mean μ of each flux.

Vague inverse-gamma priors are a common choice for variance parameters in normal models. However, with only seven GCMs and four scenarios in the ensemble, there is limited information to quantify the population variances σ_G^2, σ_S^2, and σ_{GS}^2. Specifying a "vague" inverse-gamma prior for the population variances may lead to distorted inferences owing to a high concentration of probability mass near zero, which the small sample size may be insufficient to override (Gelman 2006).

Half-Cauchy priors were specified for each of the population standard deviations σ_G, σ_S, and σ_{GS}. The half-Cauchy distribution with scale parameter A has the following form:

$$p(\sigma) = \frac{2}{\pi A}\left(1 + \frac{\sigma^2}{A^2}\right)^{-1}, \quad \text{for} \quad \sigma > 0. \tag{2}$$

The half-Cauchy distribution has the advantage of concentrating less of the prior mass close to zero than the inverse gamma distribution (Gelman 2006). By controlling the scale parameter A, we can spread the prior mass
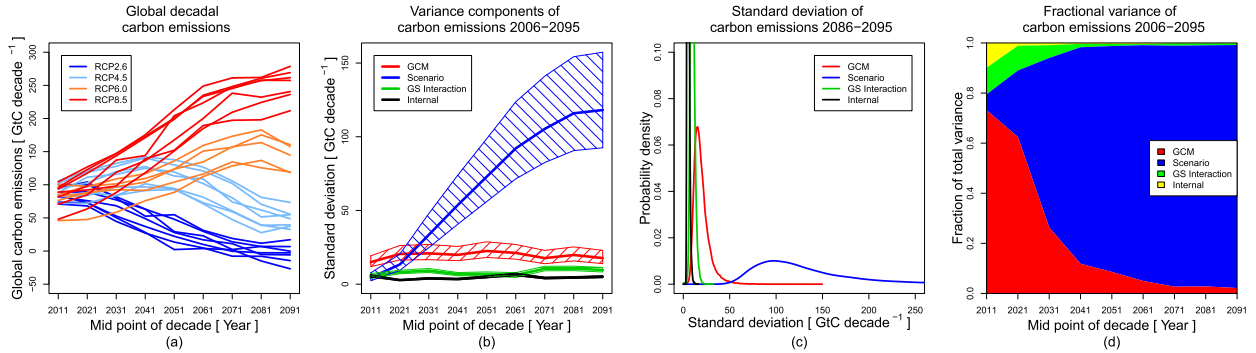
FIG. 2. (a) The decadal CO_2 fossil fuel emissions for all GCMs and scenarios; (b) the corresponding standard deviation of σ_G, σ_S, σ_{GS}, and σ_R with median values denoted by the thick lines and hatched interquartile range; (c) the posterior distribution of σ_G, σ_S, σ_{GS}, and σ_R for the decade 2086–95; and (d) the fractional variance (based on the estimated posterior medians of the superpopulations σ_G^2, σ_S^2, σ_{GS}^2, and σ_R^2) for the period 2006–95.

of the population standard deviations over a plausible range, limiting the possibility of either unreasonably small or large estimates of the population standard deviations (Gelman 2006).

The prior scale parameter A was approximated from the range in decadal fluxes in the year 2100 in Figs. 2a, 3a, 3e, 4a, 4e, 5a, 5e, and 5i. Following Northrop and Chandler (2014), the scale parameter was set to one-quarter of the approximated range (since in the Gaussian approximation 95% of the data should be within two standard deviations of the mean). Values for figures are given in Table 2.

The same scale parameter A was used for all three standard deviations σ_G, σ_S, and σ_{GS}, but different values of A were specified for each flux. Once the priors were specified, the posterior distributions of the population standard deviations were estimated by Markov chain Monte Carlo (MCMC) techniques (Gilks et al. 1996; Gelman and Rubin 1992) using the just another Gibbs sampler (JAGS) software (Plummer 2015), and the code provided by Northrop and Chandler (2014).

Estimating the prior scale parameter from the data is a double use of the data. However, no independent source was available, and the prior is designed only to provide a mild constraint on plausible values of the population standard deviations, so the compromise is acceptable.

4. Results

We follow the convention for graphically comparing variance components using stacked fractional variance plots (Hawkins and Sutton 2009). The fractional variances are based on the posterior medians of the superpopulation variances σ_G^2, σ_S^2, σ_{GS}^2, and σ_R^2.

a. Variability in CO_2 emissions

Four CO_2-concentration-driven scenarios (RCP2.6, RCP4.5, RCP6.0, and RCP8.5) were used, from which the decadal rate of change ΔCO_2 was determined. Compatible fossil fuel emissions can be determined as follows (shown in Fig. 2a):

$$\text{Emission}(t) = \Delta CO_2(t) + \text{decadal atmosphere–land carbon fluxes}(t)$$
$$+ \text{decadal atmosphere–ocean carbon fluxes}(t). \tag{3}$$

Note that Emission(t) is to the atmosphere. When we consider variability in CO_2 emissions (Fig. 2b), the scenario variance overtakes the GCM variance in the late 2020s and completely dominates by midcentury.

From the terms in Eq. (3), the compatible emissions are strongly influenced by the change in atmospheric CO_2, which is prescribed as a common forcing to all GCMs and is larger, by a factor of approximately 2–3, than the change in land and ocean carbon storage. Hence, the dominant term in the compatible emissions

does not vary between GCMs. This leads to the striking result that the variability in emissions between GCMs is much smaller than the variability between scenarios. However, there is still substantial GCM spread in the compatible emissions for each scenario, and it is desirable to investigate the components of this.

b. Variability in global CO_2 fluxes

Subjectively, it is clear that the multimodel ensemble of global land and ocean carbon fluxes behave
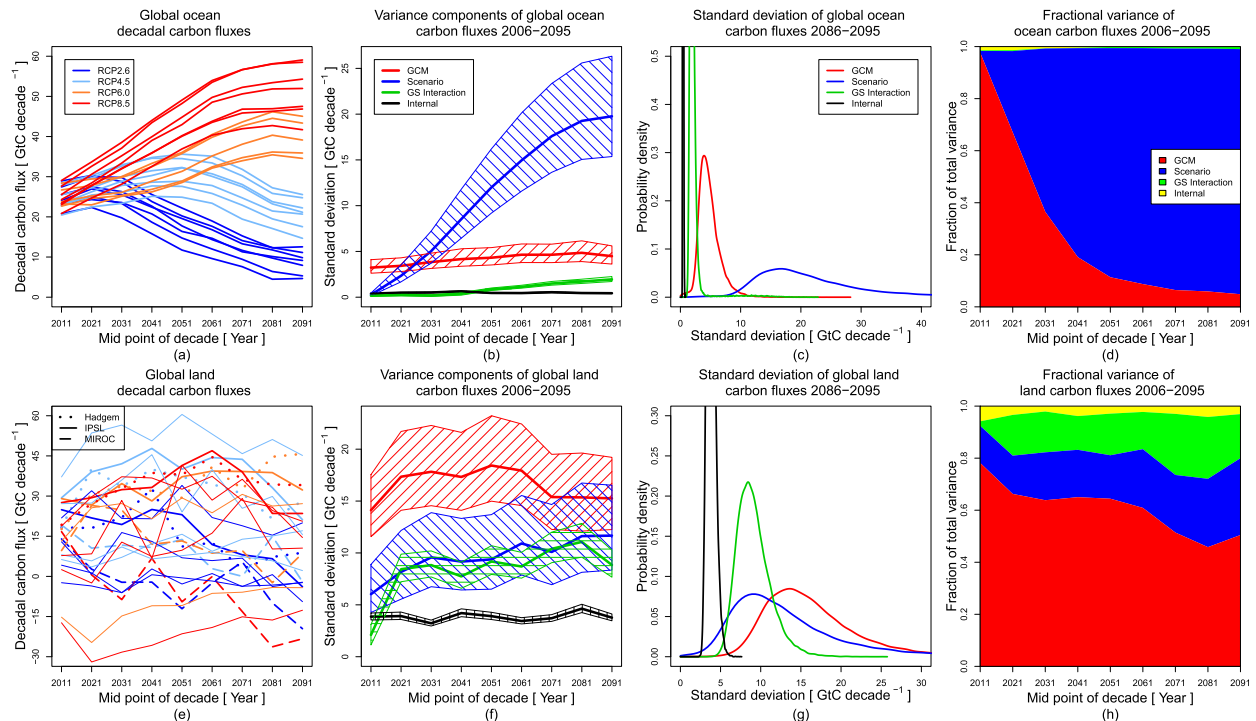
FIG. 3. As in Fig. 2, but for decadal (a)–(d) ocean and (e)–(h) land carbon fluxes for the period 2006–95. A subset of GCMs is highlighted for the global atmosphere–land carbon fluxes in (e).

differently (Fig. 3). For decadal atmosphere–ocean carbon fluxes, scenario variance σ_S^2 overtakes GCM variance σ_G^2 around 2025 and by the end of the century is contributing about 90% of the total variance across simulated ocean carbon fluxes (Fig. 3d). For decadal atmosphere–land carbon fluxes, σ_G^2 remains the dominant term throughout the twenty-first century (Fig. 3h).

It is interesting to see that the posterior median of the scenario standard deviation σ_S of the land carbon fluxes does not grow as quickly or as large as for the ocean (Figs. 3b,f) and in magnitude is about half that of the ocean by 2095. The reason for this is not explored further in this paper but may be due to competing effects of CO_2 and climate on land carbon (Arora et al. 2013). Unlike for ocean carbon fluxes, the fractional scenario variance of land carbon fluxes is nonzero in the near future (Fig. 3h), owing in part to the rapid divergence of the land-use scenarios. The GCM–scenario interaction term is an important source of variability in global land carbon fluxes. There are several contributing factors to the interaction term, and these will be investigated in section 4d.

c. Variability in regional ocean CO_2 fluxes

Two key ocean regions of CO_2 uptake were considered: the North Atlantic Ocean and the Southern Ocean. The regional extents of these ocean regions are given in Table 3. The main uptake of atmospheric CO_2 in the

Southern Ocean takes place in the latitudes 60°–40°S (Le Quéré et al. 2000, 2007; Takahashi et al. 2002), but there is much greater variability in the simulated fluxes below 60°S (Lenton et al. 2013). The noticeable difference in the fractional variance plots between Southern Ocean (Fig. 4d) and global ocean (Fig. 3d) carbon fluxes is that σ_S^2 rapidly overtakes σ_G^2 in the global case, whereas σ_G^2 remains a major source of variability at the end of the century in the Southern Ocean. The Southern Ocean circulation is known to be poorly simulated (Russell et al. 2006), and the CO_2 fluxes in this region are not well constrained by observations (Monteiro et al. 2009); these factors help to explain the large contribution made by σ_G^2.

The North Atlantic Ocean has been identified as a key sink of atmospheric CO_2 (Schuster and Watson 2007; Takahashi et al. 2002). The North Atlantic Ocean has unusually large differences between simulated carbon fluxes, as can be seen in the bunching by GCM in these fluxes (Fig. 4e) compared with the bunching by scenario in the global ocean fluxes (Fig. 3a). This explains why σ_G^2 remains the major variance term at the end of the twenty-first century (Fig. 4h). Having outlying GCMs (Fig. 4e) greatly increases σ_G^2; IPSL-CM5A-LR is one outlier here whose anomalous behavior may result from the parameterization of ice calving, which produces a flux of freshwater from the polar ice sheets (Marti et al. 2010; Roy et al. 2011).
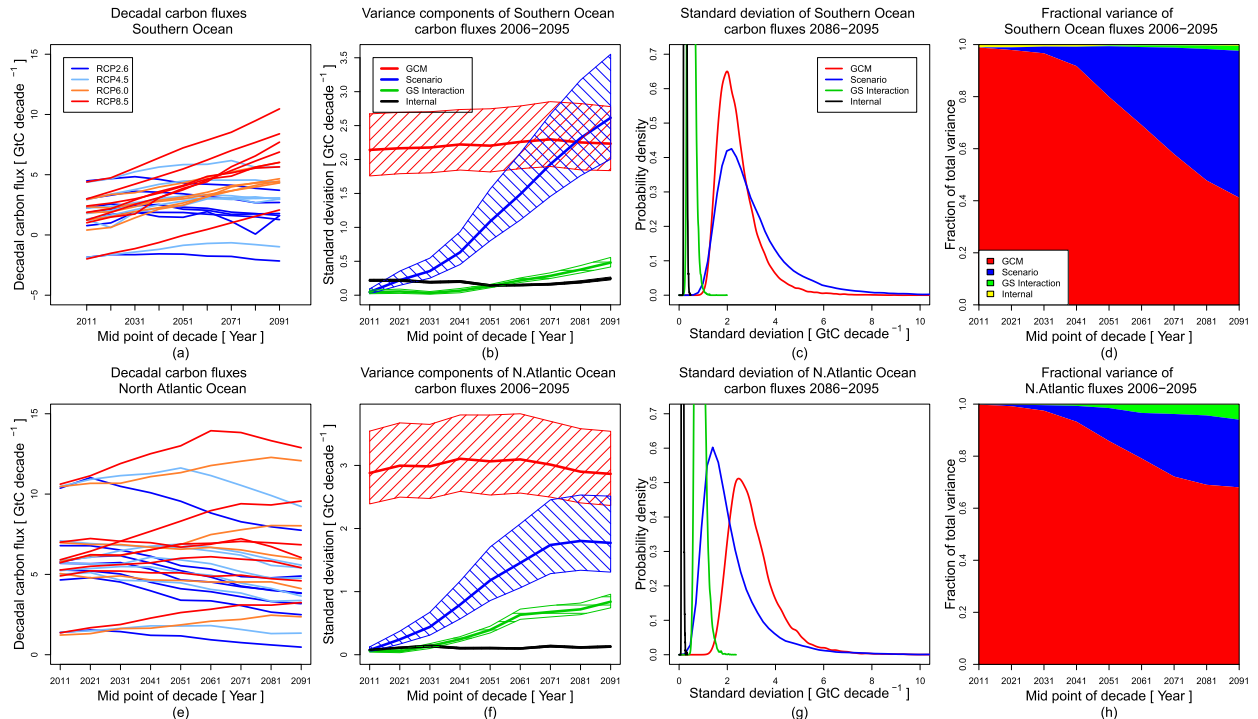
FIG. 4. As in Fig. 2, but for decadal (a)–(d) Southern Ocean and (e)–(h) North Atlantic Ocean carbon fluxes for the period 2006–95.

d. Variability in regional land CO$_2$ fluxes

We performed ANOVA for the three land regions defined in Table 3. The behavior of the northern temperate and northern high latitudes is similar. In the early decades, the carbon fluxes bunch together closely by GCM (represented by the line styles in Figs. 5a,e). Later in the century, the contribution of the GCM–scenario interaction σ^2_{GS} becomes progressively more important (Figs. 5b,h). The interaction term σ^2_{GS} appears because the spread of fluxes for GCMs run under RCP2.6 remains small but grows larger with time for the other scenarios (Figs. 5a,e). A plausible explanation for this behavior is that RCP2.6 is closest to the current climate, so the GCM spread is better constrained by present-day carbon flux observations than the other three scenarios. The scenario variance σ^2_S is negligible in the first decade, suggesting that anthropogenic land-use change is not an important early source of variability in the northern-temperate and northern-high-latitude carbon fluxes.

In the tropical land region, different scenarios are already displaying different CO$_2$ flux behaviors in the first decade (Fig. 5i) with σ^2_S being an important source of variability in the first decade (Fig. 5l), strongly suggesting that anthropogenic land-use change is an important early source of variance in tropical land region carbon fluxes. It is interesting to note that the progression of the RCP scenario radiative forcing from low to high is not repeated in the degree of carbon uptake in the tropics (Fig. 5i), with RCP8.5 appearing to have a lower carbon uptake than RCP4.5 and RCP6.0.

5. Discussion

We made a decision to aggregate the monthly fields into decadal atmosphere–ocean carbon fluxes and decadal atmosphere–land carbon fluxes (see section 2). We are not interested in trying to predict carbon fluxes for individual years; instead, we want to know how long-term cumulative changes in natural stores of carbon affect anthropogenic emissions and/or climate. For this we want to smooth out annual- and subannual-scale variability, but we still care about how much the decadal-scale fluxes are dependent on the initial conditions. The ocean uptake looks much the same on annual (not shown) and decadal time steps. However, the internal variance of annual land carbon fluxes (not shown) is very large, in agreement with expectations, owing to the tropical land response to El Niño–Southern Oscillation (ENSO)-induced climate fluctuations (Jones et al. 2001; Cox et al. 2013). Overall, more robust estimates of the variance components of land carbon fluxes were obtained from the decadal time-step.

We have chosen to use raw projection carbon flux data in the analysis since model differences at 2006 represent genuine uncertainty. For ocean uptake, observational
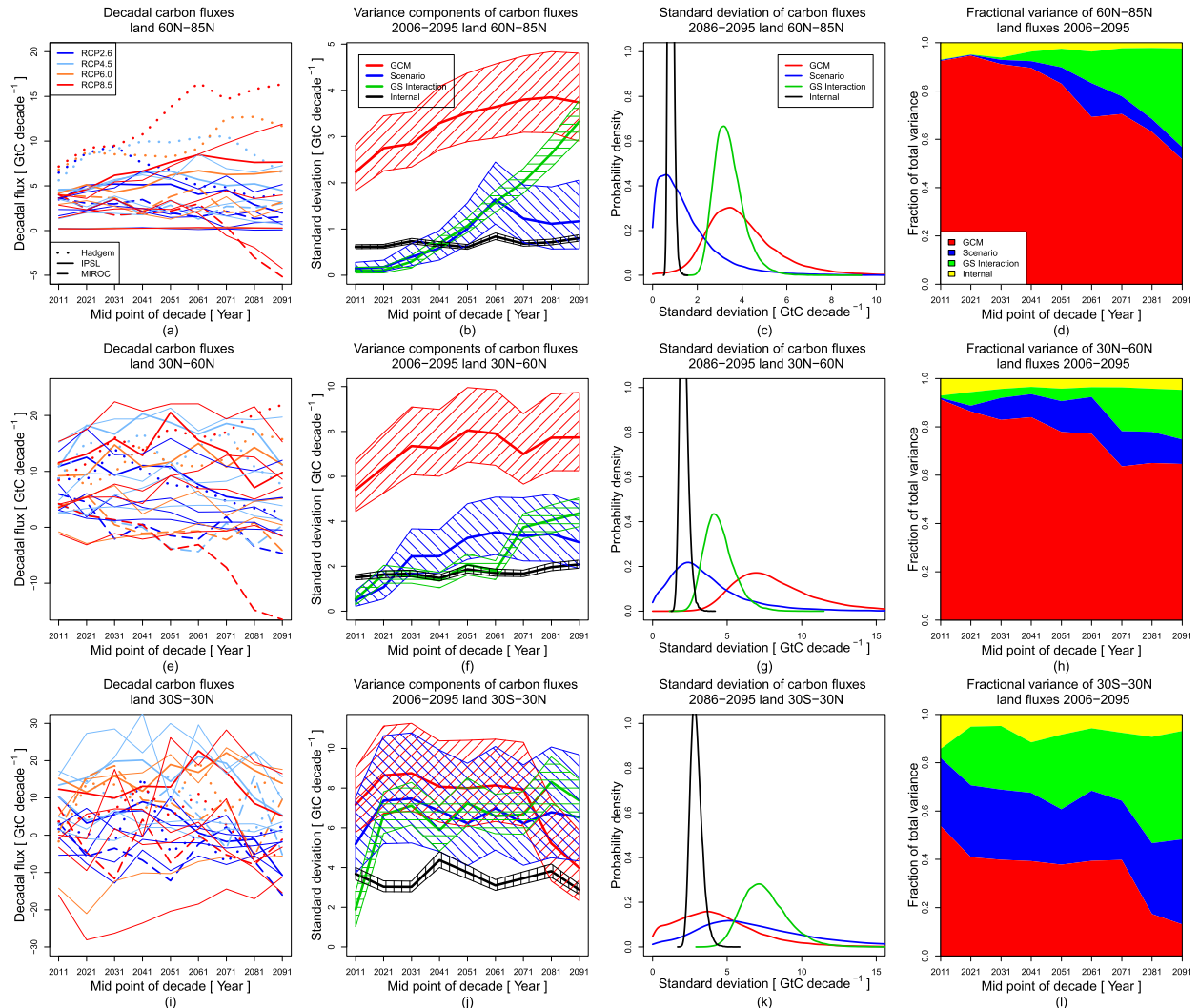
FIG. 5. As in Fig. 2, but for decadal carbon fluxes of (a)–(d) northern high-latitude land, (e)–(h) northern temperate land, and (i)–(l) tropical land for the period 2006–95. A subset of GCMs is highlighted for the decadal carbon fluxes in (a),(e),(i).

estimates in the 1990s suggest a mean ocean CO_2 sink of $2.2 \pm 0.4 \, GtC \, yr^{-1}$, which has a very similar uncertainty to the spread of GCM global ocean carbon fluxes of $2.0 \pm 0.4 \, GtC \, yr^{-1}$ given by Le Quéré et al. (2015). There are no observational constraints on land uptake of carbon in the present; these were instead estimated from the residual of CO_2 emissions, which are neither absorbed by the ocean nor remain in the atmosphere (Le Quéré et al. 2015).

Our simple ANOVA model treats each time point separately, so the results may not be robust in the presence of large internal variability; that is, the fractional variance plots may be very noisy rather than varying smoothly. In that case, it might be more appropriate to use a time series ANOVA approach such as the heuristic polynomial method of Hawkins and Sutton

(2009) or the quasi-ergodic approach of Hingray and Saïd (2014). In the present work, such a time series approach would likely have produced robust estimates of the variance terms if using variables with annual rather than decadal time steps.

TABLE 2. Values of prior scale parameter A for figures.

Figure	A
Fig. 2a	90.0
Fig. 3a	15.0
Fig. 3e	22.5
Fig. 4a	3.5
Fig. 4e	4.0
Fig. 5a	5.0
Fig. 5e	10.0
Fig. 5i	15.0

TABLE 3. Extent of regions where carbon fluxes were analyzed.

Region	Latitude	Longitude
Southern Ocean	90°–60°S	All
North Atlantic Ocean	35°–80°N	80°W–10°E
Tropical land	30°S–30°N	All
Northern temperate land	30°–60°N	All
Northern high latitudes	60°–85°N	All

In Eq. (1) the GCMs are assumed to be independent. However, there are likely to be similarities in the way certain processes are parameterized. In that case, σ_G^2 will underestimate the variability in future carbon fluxes owing to differences in the representation of carbon cycle processes and parameterizations in GCMs. Additionally, the CMIP5 multimodel ensemble is an "ensemble of opportunity"; that is, it is not designed to span the range of possible behaviors of the Earth system but is a collection of best guesses. This will also lead to σ_G^2 underestimating the variability in future carbon fluxes owing to GCM differences. For a review of the difficulties associated with interpreting multimodel ensembles, see Stephenson et al. (2012).

In contrast to the GCMs, the RCP scenarios are designed to span the 10th and 90th percentiles of likely future radiative forcing. They do not constitute a random sample from the space of possible future scenarios (i.e., they are unlikely to be independent). As a result, the superpopulation variance σ_S^2 may tend to overestimate the variability in future carbon fluxes owing to scenario uncertainty. The results when using finite-population variance (not shown) are much the same as those obtained through using superpopulation variance. The σ_S grows a little slower when using finite-population variance, and the point when scenario variance overtakes GCM variance is a few years later. The interquartile range of σ_G and σ_S are much smaller when using finite-population variance.

6. Conclusions

a. Interpretation of results

An ANOVA method was used to quantitatively partition the compatible fossil fuel emissions and their component carbon fluxes in terms of their variability across scenarios and between GCMs. For the compatible emissions, the spread of results is dominated by spread across scenarios (scenario variability overtakes GCM variability in the late 2020s), at least partly owing to the large role of prescribed CO_2 concentration in the compatible emissions. The CO_2 pathways in the RCPs are sufficiently distinct that the scenarios are more different from each other than the differences between GCMs. However, there is still variability between GCMs and therefore uncertainty in how to achieve a given pathway owing to natural land and ocean carbon uptake.

Scenario variability overtakes GCM variability in global ocean fluxes around 2025, while GCM variability remains dominant over scenario variability beyond 2100 for global land fluxes. There are large regional variations in fractional variance (e.g., GCM variability dominating beyond 2100 in the North Atlantic Ocean).

Although the focus here is on simulated carbon fluxes, this explicitly includes the GCM representation of climate processes that will differ regionally and by quantity (temperature, precipitation, etc.); that is, the simulated carbon fluxes may differ owing to different simulated climate as well as owing to differences in carbon cycle representation. Studies that have run vegetation models offline forced by prescribed climate inputs have found substantial differences from both multiple land surface models driven by the emulated climate of a single common GCM (Sitch et al. 2008) and emulated climate changes of multiple GCMs (Huntingford et al. 2013). Similarly, changes to the land surface within a common climate model framework have been shown to drive a broad spread in carbon responses (Booth et al. 2012).

At the moment land carbon flux output from different GCMs are not sufficient to reliably distinguish the consequences of different CO_2 pathways, with the disagreement between GCMs growing increasingly large from RCP2.6 to RCP8.5. The reasonable agreement in land carbon fluxes between GCMs in the present day is likely the result of constraining carbon fluxes to observations. On the other hand, the global intramodel agreement suggests that we can be more confident in linking different future socioeconomic pathways to consequences of ocean carbon uptake (such as ocean acidification) than we can for consequences of land carbon uptake.

To aid decision making, GCM output is of use if it allows a distinction to be seen between different courses of action (such as land-use choices or emissions reductions). If GCM variance is greater than scenario variance it reduces confidence in statements around the difference between scenarios and reduces the utility of the GCMs and the simulations. Where the scenarios are more different than the GCM variance this implies we have some confidence in the differences in climate outcomes owing to socioeconomic choices. The smaller we can make the GCM variance, the finer the distinction we can make between policy options.

b. Ideas for further work

In the RCP scenarios there is a progression of radiative forcing from low to high, but the degree of land-use

change does not follow this with different rates and even different signs of changes between scenarios and in particular regions. Land-use change storylines also diverge more quickly than CO_2 concentration and climate across the RCPs and are likely therefore to be responsible for the early twenty-first-century changes. It has not been possible to sample the range of land-use change scenarios independently of future CO_2 scenarios in this analysis; it would be of great use if a designed set of GCM runs could sample this range.

REFERENCES

Arora, V. K., and Coauthors, 2013: Carbon–concentration and carbon–climate feedbacks in CMIP5 Earth system models. *J. Climate*, **26**, 5289–5314, doi:10.1175/JCLI-D-12-00494.1.

Bentsen, M., and Coauthors, 2012: CMIP5 output1 NCC NorESM1-M rcp85. World Data Center for Climate at DKRZ, accessed 1 January 2016, doi:10.1594/WDCC/CMIP5.NCCNMr8.

Booth, B. B. B., and Coauthors, 2012: High sensitivity of future global warming to land carbon cycle processes. *Environ. Res. Lett.*, **7**, 024002, doi:10.1088/1748-9326/7/2/024002.

CCCma, 2015: CanESM2 model output prepared for CMIP5 rcp85. World Data Center for Climate at DKRZ, accessed 1 January 2016, doi:10.1594/WDCC/CMIP5.CCE2r8.

Ciais, P., and Coauthors, 2013: Carbon and other biogeochemical cycles. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 465–570.

Cox, P., and D. Stephenson, 2007: Climate change—A changing climate for prediction. *Science*, **317**, 207–208, doi:10.1126/science.1145956.

——, R. Betts, C. Jones, S. Spall, and I. Totterdell, 2000: Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, **408**, 184–187, doi:10.1038/35041539.

——, D. Pearson, B. B. Booth, P. Friedlingstein, C. Huntingford, C. D. Jones, and C. M. Luke, 2013: Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature*, **494**, 341–344, doi:10.1038/nature11882.

Davison, A. C., 2003: *Statistical Models*. Cambridge University Press, 726 pp.

Denman, K., and Coauthors, 2007: Couplings between changes in the climate system and biogeochemistry. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 499–587.

Denvil, S., and Coauthors, 2016: IPSL-CM5A-LR model output prepared for CMIP5 RCP8.5. World Data Center for Climate at DKRZ, accessed 1 January 2016, doi:10.1594/WDCC/CMIP5.IPILr8.

Dunne, J. P., and Coauthors, 2014: NOAA GFDL GFDL-ESM2G, rcp85 experiment output for CMIP5 AR5. World Data Center for Climate at DKRZ, accessed 1 January 2016, doi:10.1594/WDCC/CMIP5.NGEGr8.

Friedlingstein, P., and Coauthors, 2006: Climate–carbon cycle feedback analysis: Results from the C^{4}MIP model intercomparison. *J. Climate*, **19**, 3337–3353, doi:10.1175/JCLI3800.1.

Gelman, A., 2005: Analysis of variance—Why it is more important than ever. *Ann. Stat.*, **33**, 1–31, doi:10.1214/009053604000001048.

——, 2006: Prior distributions for variance parameters in hierarchical models (Comment on article by Browne and Draper). *Bayesian Anal.*, **1**, 515–534.

——, and D. B. Rubin, 1992: Inference from iterative simulation using multiple sequences. *Stat. Sci.*, **7**, 457–472, doi:10.1214/ss/1177011136.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter, Eds., 1996: *Markov Chain Monte Carlo*. 2nd ed. Chapman and Hall, 512 pp.

Giorgetta, M., and Coauthors, 2012: CMIP5 simulations of the Max Planck Institute for Meteorology (MPI-M) based on the MPI-ESM-LR model: The rcp85 experiment. World Data Center for Climate at DKRZ, accessed 1 January 2016, doi:10.1594/WDCC/CMIP5.MXELr8.

Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.*, **90**, 1095–1107, doi:10.1175/2009BAMS2607.1.

Hibbard, K. A., G. A. Meehl, P. Cox, and P. Friedlingstein, 2007: A strategy for climate change stabilization experiments. *Eos, Trans. Amer. Geophys. Union*, **88**, 217–221, doi:10.1029/2007EO200002.

Hingray, B., and M. Saïd, 2014: Partitioning internal variability and model uncertainty components in a multimember multimodel ensemble of climate projections. *J. Climate*, **27**, 6779–6798, doi:10.1175/JCLI-D-13-00629.1.

Huntingford, C., and Coauthors, 2013: Simulated resilience of tropical rainforests to CO_2-induced climate change. *Nat. Geosci.*, **6**, 268–273, doi:10.1038/ngeo1741.

Hurtt, G. C., and Coauthors, 2011: Harmonization of land-use scenarios for the period 1500 to 2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic Change*, **109**, 117–161, doi:10.1007/s10584-011-0153-2.

JAMSTEC, AORI, and NIES, 2015: MIROC-ESM model output prepared for CMIP5 rcp85. World Data Center for Climate at DKRZ, accessed 1 January 2016, doi:10.1594/WDCC/CMIP5.MIMEr8.

Jones, C. D., M. Collins, P. M. Cox, and S. A. Spall, 2001: The carbon cycle response to ENSO: A coupled climate–carbon cycle model study. *J. Climate*, **14**, 4113–4129, doi:10.1175/1520-0442(2001)014<4113:TCCRTE>2.0.CO;2.

——, P. M. Cox, and C. Huntingford, 2006: Climate–carbon cycle feedbacks under stabilization: Uncertainty and observational constraints. *Tellus*, **58B**, 603–613, doi:10.1111/j.1600-0889.2006.00215.x.

——, and Coauthors, 2013: Twenty-first-century compatible CO_2 emissions and airborne fraction simulated by CMIP5 Earth system models under four representative concentration pathways. *J. Climate*, **26**, 4398–4413, doi:10.1175/JCLI-D-12-00554.1.

Karoly, D. J., and Q. Wu, 2005: Detection of regional surface temperature trends. *J. Climate*, **18**, 4337–4343, doi:10.1175/JCLI3565.1.

Lenton, A., and Coauthors, 2013: Sea–air CO_2 fluxes in the Southern Ocean for the period 1990–2009. *Biogeosciences*, **10**, 4037–4054, doi:10.5194/bg-10-4037-2013.

Le Quéré, C., J. C. Orr, P. Monfray, O. Aumont, and G. Madec, 2000: Interannual variability of the oceanic sink of CO_2 from 1979 through 1997. *Global Biogeochem. Cycles*, **14**, 1247–1265, doi:10.1029/1999GB900049.

——, and Coauthors, 2007: Saturation of the Southern Ocean CO_2 sink due to recent climate change. *Science*, **316**, 1735–1738, doi:10.1126/science.1136188.

——, and Coauthors, 2015: Global carbon budget 2015. *Earth Syst. Sci. Data*, **7**, 349–396, doi:10.5194/essd-7-349-2015.

Madden, R. A., 1976: Estimates of the natural variability of time-averaged sea-level pressure. *Mon. Wea. Rev.*, **104**, 942–952, doi:10.1175/1520-0493(1976)104<0942:EOTNVO>2.0.CO;2.

Marti, O., and Coauthors, 2010: Key features of the IPSL ocean atmosphere model and its sensitivity to atmospheric resolution. *Climate Dyn.*, **34**, 1–26, doi:10.1007/s00382-009-0640-6.

Masui, T., and Coauthors, 2011: An emission pathway to stabilize at 6 W m^{-2} of radiative forcing. *Climatic Change*, **109**, 59–76, doi:10.1007/s10584-011-0150-5.

Monteiro, P. M. S., and Coauthors, 2009: A global sea surface carbon observing system: Assessment of changing sea surface CO_2 and air–sea CO_2 fluxes. *Proc. OceanObs'09: Sustained Ocean Observations and Information for Society*, Venice, Italy, European Space Agency, ESA Publ. WPP-306, doi:10.5270/OceanObs09.cwp.64.

Moss, R. H., and Coauthors, 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 747–756, doi:10.1038/nature08823.

Northrop, P. J., and R. E. Chandler, 2014: Quantifying sources of uncertainty in projections of future climate. *J. Climate*, **27**, 8793–8808, doi:10.1175/JCLI-D-14-00265.1.

Plattner, G. K., and Coauthors, 2008: Long-term climate commitments projected with climate–carbon cycle models. *J. Climate*, **21**, 2721–2751, doi:10.1175/2007JCLI1905.1.

Plummer, M., 2015: JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. R JAGS 3.4.0. [Available online at http://mcmc-jags.sourceforge.net.]

R Core Team, 2013: R: A language and environment for statistical computing. R Foundation for Statistical Computing. [Available online at http://www.R-project.org/.]

Riahi, K., and Coauthors, 2011: RCP8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change*, **109**, 33–57, doi:10.1007/s10584-011-0149-y.

Roy, T., and Coauthors, 2011: Regional impacts of climate change and atmospheric CO_2 on future ocean carbon uptake: A multimodel linear feedback analysis. *J. Climate*, **24**, 2300–2318, doi:10.1175/2010JCLI3787.1.

Russell, J. L., R. J. Stouffer, and K. W. Dixon, 2006: Intercomparison of the Southern Ocean circulations in IPCC coupled model control simulations. *J. Climate*, **19**, 4560–4575, doi:10.1175/JCLI3869.1.

Sanderson, M., J. Hughes, and C. Jones, 2014: HadGEM2–ES model output prepared for CMIP5 RCP8.5. World Data Center for Climate at DKRZ, accessed 1 January 2016, doi:10.1594/WDCC/CMIP5.MOGEr8.

Sansom, P. G., D. B. Stephenson, C. A. T. Ferro, G. Zappa, and L. Shaffrey, 2013: Simple uncertainty frameworks for selecting weighting schemes and interpreting multimodel ensemble climate change experiments. *J. Climate*, **26**, 4017–4037, doi:10.1175/JCLI-D-12-00462.1.

Schuster, U., and A. J. Watson, 2007: A variable and decreasing sink for atmospheric CO_2 in the North Atlantic. *J. Geophys. Res.*, **112**, C11006, doi:10.1029/2006JC003941.

Searle, S. R., 1987: *Linear Models for Unbalanced Design*. Wiley, 536 pp.

Sitch, S., and Coauthors, 2008: Evaluation of the terrestrial carbon cycle, future plant geography and climate–carbon cycle feedbacks using five dynamic global vegetation models (DGVMs). *Global Change Biol.*, **14**, 2015–2039, doi:10.1111/j.1365-2486.2008.01626.x.

Stephenson, D. B., M. Collins, J. C. Rougier, and R. E. Chandler, 2012: Statistical problems in the probabilistic prediction of climate change. *Environmetrics*, **23**, 364–372, doi:10.1002/env.2153.

Takahashi, T., and Coauthors, 2002: Global sea–air CO_2 flux based on climatological surface ocean ρCO_2, and seasonal biological and temperature effects. *Deep-Sea Res. II*, **49**, 1601–1622, doi:10.1016/S0967-0645(02)00003-6.

Taylor, K. E., 2013: CMIP5 Coupled Model Intercomparison Project: Standard output document. [Available online at http://cmip-pcmdi.llnl.gov/cmip5/docs/standard_output.pdf.]

——, R. Stouffer, and G. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Thomson, A. M., and Coauthors, 2011: RCP4.5: A pathway for stabilization of radiative forcing by 2100. *Climatic Change*, **109**, 77–94, doi:10.1007/s10584-011-0151-4.

Tingley, M. P., 2012: A Bayesian ANOVA scheme for calculating climate anomalies, with applications to the instrumental temperature record. *J. Climate*, **25**, 777–791, doi:10.1175/JCLI-D-11-00008.1.

van Vuuren, D. P., and Coauthors, 2011a: RCP2.6: Exploring the possibility to keep global mean temperature increase below 2°C. *Climatic Change*, **109**, 95–116, doi:10.1007/s10584-011-0152-3.

——, and Coauthors, 2011b: The representative concentration pathways: An overview. *Climatic Change*, **109**, 5–31, doi:10.1007/s10584-011-0148-z.

von Storch, H., and F. Zwiers, 2001: *Statistical Analysis in Climate Research*. Cambridge University Press, 496 pp.

Yip, S., C. A. T. Ferro, D. B. Stephenson, and E. Hawkins, 2011: A simple, coherent framework for partitioning uncertainty in climate predictions. *J. Climate*, **24**, 4634–4643, doi:10.1175/2011JCLI4085.1.