Tellus SERIES B CHEMICAL AND PHYSICAL METEOROLOGY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

Tellus (2009), 61B, 361–371
Printed in Singapore. All rights reserved

TELLUS

Are response function representations of the global carbon cycle ever interpretable?

By SILE LI, ANDREW J. JARVIS* and DAVID T. LEEDAL, *Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK*

(Manuscript received 7 May 2008; in final form 22 September 2008)

ABSTRACT

Response function models are often used to represent the behaviour of complex, high order global carbon cycle (GCC) and climate models in applications which require short model run times. Although apparently black-box, these response function models need not necessarily be entirely opaque, but instead may also convey useful insights into the properties of the parent model or process. By exploiting a transfer function (TF) framework to analyse the Lenton GCC model, this paper attempts to demonstrate that response function representations of GCC models can sometimes also provide structural information on the parent model from which they are identified and calibrated. We take a fifth-order TF identified from the impulse response of the Lenton model atmospheric burden, and decompose this to show how it can be re-expresses in a generic five-box form in sympathy with the structure of the parent model.

1. Introduction

The global carbon cycle (GCC) would qualify under most definitions of a complex system. Because of the central role the GCC plays in determining the fate of anthropogenic CO_2, considerable effort has been committed to mapping out this complexity and representing it in the form of carbon cycle models of various kinds. Not surprisingly, as understanding of the GCC has increased, so has the complexity of the models with which we choose to represent it as the respective scientists strive for accuracy (e.g. Maier-Reimer et al., 1996; Sarmiento et al., 1998; Orr et al., 2001). However, in parallel to these model developments, certain applications have demanded simple representations of the GCC for, for example, detection-attribution studies (e.g. Enting and Trudinger, 2002), integrated assessments (e.g. Nordhaus and Boyer, 2000; den Elzen and Lucas, 2005); CO_2 emissions mitigation specification (Wigley, 1991; Wigley et al., 1996) and, recently, mitigation analysis (Jarvis et al., 2008). Often these simplified GCC analogues have taken the form of black-box response functions which capture the dominant perturbation dynamics of their large GCC model parents. In particular, sums of exponentials (SE) response functions have been widely used in this respect (e.g. Maier-Reimer and Hasselmann, 1987; Enting, 1990; Joos et al., 1996; Thompson and Randerson, 1999; Hooss et al. 2001).

In reviewing the use of response functions for these various GCC applications it appears that they invariably tend to be between fourth and sixth order, that is, the sum of between four and six first-order exponential response function terms. It has been known for many years that, extreme non-linearity aside, the dynamic response of complex systems can often be dominated by a relatively small number of dominant modes (see e.g. Moore, 1981; Dowell, 1996; Young et al., 1996; Young, 1999). However, what is less well known is that these modes can often represent fundamental dynamic states of a system. Indeed, systems identification methodologies and data-based mechanistic modelling studies, attempt to exploit such phenomena to resolve hidden processes from observed inputs and outputs (e.g. Godfrey, 1982; Tang et al., 2001; Garnier et al., 2003; Young and Garnier, 2006). Extending this philosophy to the GCC, it may be that the response functions identified from GCC model perturbation experiments are also a manifestation of the dominant modes of GCC models and offer insight into the emergent macroscale mechanisms within these models. For example, Meyer et al. (1999) derive a box-type model of the terrestrial biosphere from the SE response function of the Bern model. Also, Hooss et al., (2001) equate a layer cascade representation of the oceanic carbon cycle with response functions. However, in the main, response functions have only ever been viewed as empirical devices and Enting (2007) suggests that it is unlikely they relate to specific carbon reservoir dynamics. In this paper we will attempt to show that the dominant mode dynamics of a GCC model, as expressed in its impulse response can, under certain circumstances, reflect specific carbon reservoir turnovers,

*Corresponding author.
e-mail: A.Jarvis@lancaster.ac.uk
DOI: 10.1111/j.1600-0889.2008.00401.x

providing an appropriate decomposition of the black-box response function model is used.

The outline of this paper is as follows. Section 2 reviews the relationship between three representations of the GCC impulse response: SEs; transfer functions (TFs); and ordinary differential equations (ODEs). Section 3 shows the equivalence between a fifth-order TF response function and a generic five-box GCC model. Section 4 investigates the specific relationship between an identified response function, its 'grey-box' or mechanistic decomposition and the Lenton (2000) GCC model from which it was identified.

2. The representation of the GCC impulse response: sums of exponentials, transfer functions and ordinary differential equations

We start by outlining the relationship between familiar SE and TF response functions and their ODE parent structures. A generic, delay free deterministic single-input, single-output, linear, continuous-time ODE system can be written,

$$\frac{d^n x(t)}{dt^n} + a_1 \frac{d^{n-1}x(t)}{dt^{n-1}} + \cdots + a_n x(t) = b_0 \frac{d^{m-1}u(t)}{dt^{m-1}} + \cdots + b_{m-1}u(t), \quad (1)$$

where $u(t)$ is the input, $x(t)$ is the system output, a and b are coefficients and m and n define the system order. This expression could, for example, be the conservation equation for carbon in one or more elements of the GCC. Taking the Laplace transform of eq. (1) and simplifying by assuming all initial conditions are zero,[1] the generic linear time invariant continuous-time system is given by the following TF form,

$$x(s) = \frac{B(s)}{A(s)}u(s) = \frac{b_0 s^{m-1} + b_1 s^{m-2} + \cdots + b_{m-2}s + b_{m-1}}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}u(s), \quad (2)$$

where $u(s)$ and $x(s)$ are the Laplace transforms of the input $u(t)$ and output $x(t)$ respectively; $s^{m,n}$ is the Laplace derivative operator (i.e. $s^n = d^n/dt^n$) and the ratio of polynomials $B(s)/A(s)$ denotes the TF.

If $A(s)$ is of order n, then the characteristic equation [defined as $A(s) = 0$] will have n pole roots $p_1, p_2, \ldots, p_n$, which are the poles of the TF.[2] Providing $m = n$, the TF $B(s)/A(s)$ can be rearranged in a partial fraction expansion form:

$$\frac{B(s)}{A(s)} = \frac{r_1}{s - p_1} + \frac{r_2}{s - p_2} + \cdots + \frac{r_n}{s - p_n}, \quad (3)$$

where $r_1, r_2, \ldots r_n$ are the partial fraction expansion residue terms. In structural terms this is representative of a system comprised of a collection of first-order subsystems configured in parallel where the steady state gain of each first-order element in this expansion is $g_n = -r_n/p_n$ and the time constant or e-folding time is $\tau_n = {p_n}^{-1}$ for $p_n < 0$.

For the impulse response case considered here the relationship between the s domain and the t domain is given by

$$L^{-1}\left\{\frac{r}{s - p}\right\} = re^{pt}, \quad (4)$$

where L^{-1} denotes the inverse of Laplace transform. Therefore, the SE impulse response function is simply the inverse Laplace form of eq. (3):

$$L^{-1}\left\{\frac{B(s)}{A(s)}\right\} = r_1 e^{-\frac{t}{\tau_1}} + r_2 e^{-\frac{t}{\tau_2}} + \cdots + r_n e^{-\frac{t}{\tau_n}}. \quad (5)$$

Equations (1)–(5) show how the SE response function widely used in GCC analysis can be expressed as a subclass of the more generic TF response function, and how both of these relate to ODEs, and vice versa. In this study we will be exploiting TFs in order to facilitate a range of decompositions of response functions into first-order subsystem, as opposed to using the somewhat restrictive parallel decomposition inferred from the more familiar SE response functions. The following section will show how a generic GCC box structure based on ODEs can have a corresponding TF form.

3. A generic global carbon cycle structure

Figure 1 shows a generic structure for the GCC. Broadly, differences between various GCC models could be expressed as differences in the subdivisions and specifications of each compartment. In this section we will develop the TF response function for this system and show this in a block diagram form commonly used in control systems analysis.[3] This form provides a visualization of the GCC response function that is structurally related to Fig. 1.

[1] In this paper, the perturbational experiments analysed are initiated from a zero initial condition and so all related analysis assumes these conditions to be zero.

[2] Note that for a stable response (i.e., an impulse response that decays to zero) each pole must be < zero; for a conservative response (i.e., an impulse response that stabilises at a non-zero value) one or more poles must be equal to 0; if a pole is greater than zero then the impulse response is unstable and will grow exponentially.

[3] A key advantage of block diagram notation is that, while expressing the same information as the ODE form, it separates the input and output component from the system dynamics description facilitating algebraic manipulation of the dynamic components.

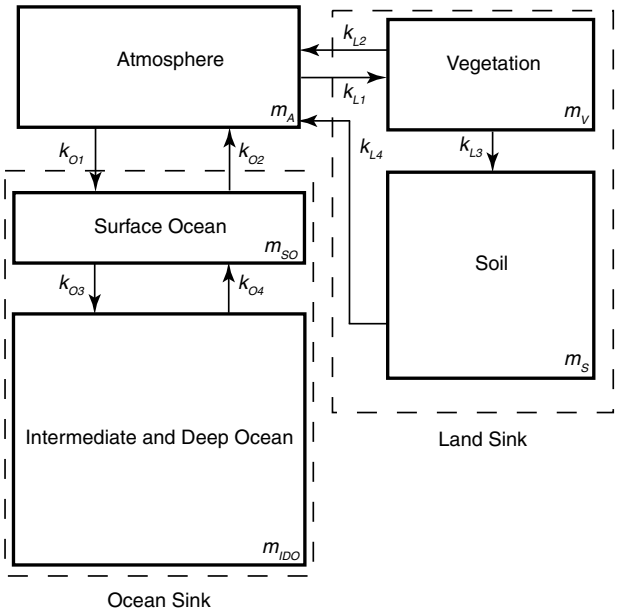

Fig. 1. A generic structure for the global carbon cycle showing the amount of carbon in the atmosphere (m_A), vegetation (m_V), soil (m_S), surface ocean (m_{SO}) and intermediate and deep ocean (m_{IDO}) reservoirs. The arrows represent the net carbon fluxes between boxes as governed by the exchange coefficients k (adapted from Joos et al., 1996 and Lenton, 2000).

For the small perturbations, locally linear case considered later, Fig. 1 gives the atmospheric mass balance ODE as,

$$\begin{aligned}\frac{dm_A(t)}{dt} &= u_E(t) - (k_{O1} + k_{L1})m_A(t) + k_{L2}m_V(t) \\ &\quad + k_{O2}m_{SO}(t) + k_{L4}m_S(t),\end{aligned} \tag{6a}$$

where $u_E(t)$ are exogenous CO_2 emissions (GtC yr^{-1}), $m_x(t)$ are the carbon burdens (GtC) of the various reservoirs, k's are exchange coefficients (yr^{-1}) and the subscripts relate to those shown in Fig. 1.

The four remaining reservoir mass balances are given by the following ODEs,

$$\begin{aligned}\frac{dm_{SO}(t)}{dt} &= k_{O1}m_A(t) - k_{O2}m_{SO}(t) - k_{O3}m_{SO}(t) \\ &\quad + k_{O4}m_{IDO}(t) \\ \frac{dm_{IDO}(t)}{dt} &= k_{O3}m_{SO}(t) - k_{O4}m_{IDO}(t) \\ \frac{dm_V(t)}{dt} &= k_{L1}m_A(t) - k_{L2}m_V(t) - k_{L3}m_V(t) \\ \frac{dm_S(t)}{dt} &= k_{L3}m_V(t) - k_{L4}m_S(t).\end{aligned} \tag{6b}$$

Assuming zero initial conditions, the Laplace transform of eq. (6) allows manipulation of this ODE system into the following collection of first-order TF response functions,

$$\begin{aligned}m_A(s) &= \frac{1}{s}[u_E(s) - (k_{O1} + k_{L1})m_A(s) + k_{L2}m_V(s) \\ &\quad + k_{O2}m_{SO}(s) + k_{L4}m_S(s)] \\ m_{SO}(s) &= \frac{1}{s + k_{O2} + k_{O3}}[k_{O1}m_A(s) + k_{O4}m_{IDO}(s)] \\ m_{IDO}(s) &= \frac{1}{s + k_{O4}}k_{O3}m_{SO}(s) \\ m_V(s) &= \frac{1}{s + k_{L2} + k_{L3}}k_{L1}m_A(s) \\ m_S(s) &= \frac{1}{s + k_{L4}}k_{L3}m_V(s).\end{aligned} \tag{7}$$

After some rearrangement of eq. (7), one gets the following system,

$$\begin{aligned}m_A(s) &= \frac{1}{s}\{u_E(s) + [k_{O2}m_{SO}(s) - k_{O1}m_A(s)] \\ &\quad + [k_{L2}m_V(s) - k_{L1}m_A(s)] + k_{L4}m_S(s)\} \\ k_{O2}m_{SO}(s) &= \frac{\frac{k_{O1}k_{O2}}{k_{O2}+k_{O3}}}{\frac{1}{k_{O2}+k_{O3}}s + 1}\left[m_A(s) + \frac{k_{O4}}{k_{O1}}m_{IDO}(s)\right] \\ \frac{k_{O4}}{k_{O1}}m_{IDO}(s) &= \frac{\frac{k_{O3}}{k_{O1}k_{O2}}}{\frac{1}{k_{O4}}s + 1}k_{O2}m_{SO}(s) \\ k_{L2}m_V(s) &= \frac{\frac{k_{L1}k_{L2}}{k_{L2}+k_{L3}}}{\frac{1}{k_{L2}+k_{L3}}s + 1}m_A(s) \\ k_{L4}m_S(s) &= \frac{\frac{k_{L3}}{k_{L2}}}{\frac{1}{k_{L4}}s + 1}k_{L2}m_V(s).\end{aligned} \tag{8}$$

Figure 2a shows this system in block diagram form. Equating eq. (8) with Fig. 2a yields the following identities

$$\begin{aligned}&\tau_{O1} = \frac{1}{k_{O2} + k_{O3}}; \quad g_{O1} = \frac{k_{O1}k_{O2}}{k_{O2} + k_{O3}}; \quad \tau_{O2} = \frac{1}{k_{O4}}; \\ &g_{O2} = \frac{k_{O3}}{k_{O1}k_{O2}} \\ &\tau_{L1} = \frac{1}{k_{L2} + k_{L3}}; \quad g_{L1} = \frac{k_{L1}k_{L2}}{k_{L2} + k_{L3}}; \quad \tau_{L2} = \frac{1}{k_{L4}}; \\ &g_{L2} = \frac{k_{L3}}{k_{L2}} \\ &g_L = k_{L1}; g_O = k_{O1}.\end{aligned} \tag{9}$$

Note from Fig. 2a how the block diagram form of this system makes explicit the fact that the exogenous emissions input $u_E(s)$ are moderated by two core feedbacks, one from the land, the other from the ocean, each representing the generation of sink/source fluxes to the atmosphere in response to perturbations in atmospheric carbon burden. These feedbacks are comprised of an instantaneous feedback effect that represents the atmospheric carbon loss to the terrestrial and oceanic sinks, and a lagged feedback effect, which represents the effects of adjustments in the respective carbon stock on the loss rate from the atmosphere.

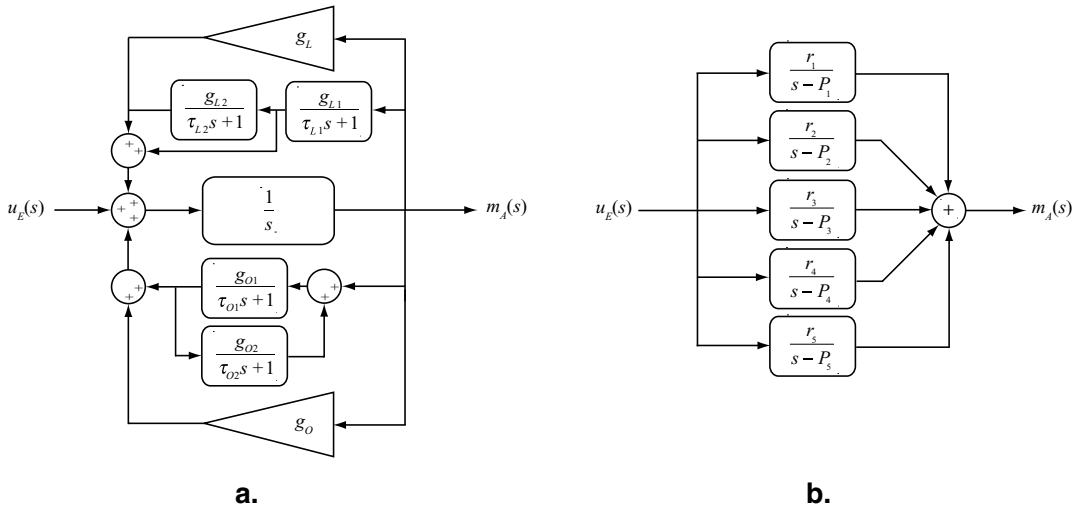

Fig. 2. (a) The block diagram equivalent to the generic GCC model structure shown in Fig. 1. The triangle blocks are instantaneous atmospheric carbon loss rates to the terrestrial (L)/oceanic (O) carbon sinks, and each rectangle block represents a terrestrial/oceanic carbon sink (see eqs (8) and (9)). (b) The block diagram of the SE decomposition of the TF eq. (10).

Using the rules of block diagram manipulation (see Appendix A) eqs (8) and (9) can be collapsed to form the overall TF response function relating $u_E(s)$ to $m_A(s)$.

$$m_A(s) = \frac{b_0 s^4 + b_1 s^3 + b_2 s^2 + b_3 s + b_4}{s^5 + a_1 s^4 + a_2 s^3 + a_3 s^2 + a_4 s + a_5} \cdot u_E(s), \qquad (10)$$

where the TF parameters a_1 to a_5 and b_0 to b_4 are comprised of combinations of the GCC system parameters in eq. (7) (see Appendix A).

The operation of moving from eqs (6) to (10) is obviously a bottom-up model building exercise. However, let us assume that we have arrived at eq. (10) through empirically describing the impulse response characteristics of a planned GCC model experiment. Then we could entertain the reverse operation of going from eqs (10) to (6) which would be the corresponding top-down systems identification process. Because, for this particular example, $m = n = 5$, eq. (10) could also be decomposed into its equivalent parallel structure, as shown in Fig. 2b. Then we would arrive at the familiar SE response function representation of the GCC yielding identical, albeit black-box, dynamics in terms of the overall relationship between the input $u_E(s)$ and the output $m_A(s)$.[4]

$$m_A(s) = \left[\frac{r_1}{s - p_1} + \frac{r_2}{s - p_2} + \frac{r_3}{s - p_3} + \frac{r_4}{s - p_4} + \frac{r_5}{s - p_5} \right] \cdot u_E(s). \qquad (11)$$

[4]Although the parallel SE structure is empirical in the present GCC application, there are numerous applications where a parallel system configuration is appropriate, for example the flow of rainfall inputs through catchments (see Young and Beven, 1994).

Because of the ability to interchange between the ODE, TF and the SE representations of this particular system one can see how SEs of order 4–6 regularly used to represent the GCC (e.g. Maier-Reimer and Hasselmann, 1987; Joos et al., 1996; Hasselmann et al., 1997; Hooss et al., 2001) could well be inadvertent expressions of the structural character of these GCC models, albeit in a very aggregated form. In essence the parallel SE form represents a translation of the axis of the systems state variables away from the mechanistically recognizable allocation of the ODEs to an equivalent albeit abstract space. This would not be at all obvious to those using SEs because it would be difficult to conceptualize how to make the transition to the relevant GCC structure without the necessary aggregation and disaggregation tools such as the Laplace transform and block diagram manipulation. A response function analysis based on TFs on the other hand naturally lends itself to this.

To summarize, the SE response function can be viewed as a specific subclass of the more general TF representation of linear perturbational dynamics and hence the authors would argue that response function representations of the GCC should move toward the use of TFs in future. There are two important reasons for this. First, the SE approach implicitly imposes the constraint $n = m$, whereas there are countless examples of systems, possibly including some GCC models, where this is not the case, especially where the system in question contains important serially organized subsystems, in which case $n > m$; or the time constants of elements of the system forward path are tending to zero, in which case $m > n$. Obviously, the TF makes no such restrictions, although it is worth bearing in mind that the lack of constraint can, in certain circumstances, be a disadvantage (see later). Second, TFs yield a unique input–output dependency which facilitates

the use of block diagrams in the representation of the response function. As we will see later, this not only greatly aids the communication of the mathematical interdependencies comprising a system such as the GCC in terms of mass (or energy) balance, it also greatly simplifies the top-down interpretation. Finally, eq. (2) is quasi-linear in the parameters when estimated from the input–output data whereas, unless one attempts to linearize the input–output data, p (or τ) in each exponential term in eq. (5) is non-linearly related to the response. Hence the TF estimation should, in principle, be more robust and able to exploit many of the advantages associated with linear estimation. In the following section we will apply TF response function estimation and mechanistic decomposition to analyse the impulse response of a specific GCC model in order to illustrate the points raised in this section.

4. Impulse response characteristics of the Lenton (2000) Global Carbon Cycle model

So far, we have formed the response function bottom-up having assumed a particular GCC box model structure a priori. However, in response function analysis we invariably start with input–output data collected from a planned experiment on a complex GCC model and attempt to fit to this in as parsimonious a way possible. The inference of the structural dynamics of the parent model then needs to be made from this identified response function, top-down, that is, classical systems identification.

In order to test whether such an analysis can capture the fundamental dynamic traits of the parent GCC model, in this section we apply a TF response function analysis to a GCC model and verify the estimates we obtain for the properties of the various compartments are consistent with the known values in the GCC model from which the TF was derived. Because response functions invariably involve some simplification it is worth keeping in mind that the TF decomposition will likely involve some aggregation of 'similar' dynamic elements. Also, given GCC models are non-linear in many of their processes (CO_2 fertilization; oceanic carbon chemistry; temperature feedbacks) the TF analysis will only provide a local linear approximation of these non-linearities.

For this verification exercise we will use the GCC model of Lenton (2000; LGCC) given it is simple enough to afford a relatively transparent comparison with its response function daughter. The LGCC model structure is shown in Fig. 3. This is a seven-box GCC model, with the terrestrial biosphere comprised of global vegetation and soil pools, the surface ocean comprised of high and low latitude reservoirs and the deep oceans comprised of intermediate and deep ocean reservoirs.

The unit (1 GtC) impulse response of the LGCC model is shown in Fig. 4 for pre-industrial background levels of atmospheric CO_2. The generic TF structure eq. (2) was fitted to this unit impulse response for the range of TF structures $n = 2{:}6$; $m = 2{:}6$ using linear least squares given the noise-free nature

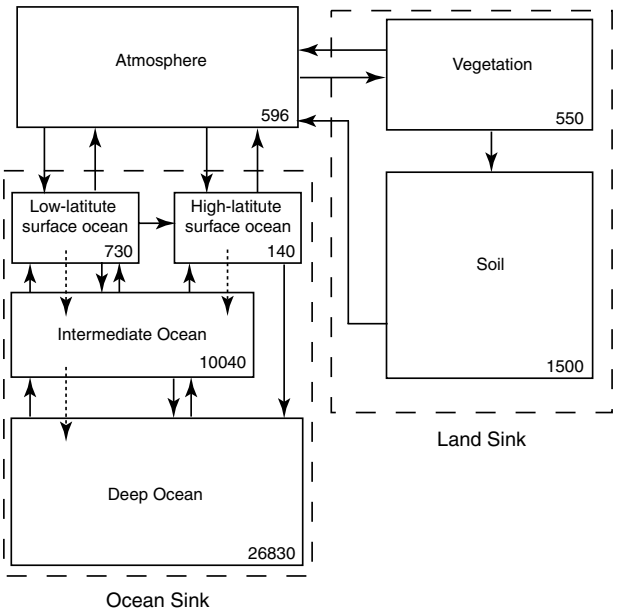

Fig. 3. The Lenton (2000) seven-box GCC model structure in pre-industrial steady state. Numbers represent carbon mass in reservoirs (GtC). Solid arrows are gas and liquid-phase fluxes (GtC yr^{-1}) and dotted lines are particulate-based fluxes (GtC yr^{-1}).

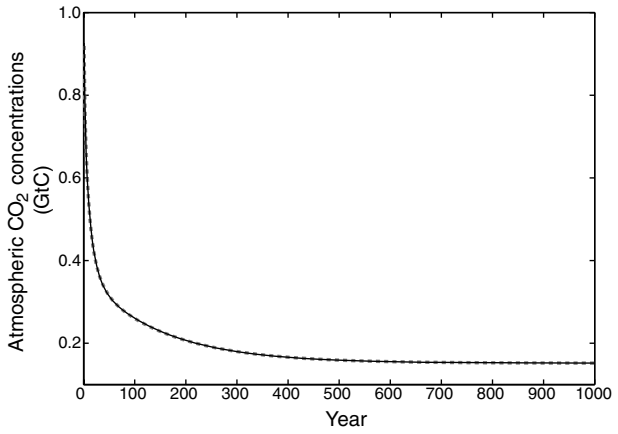

Fig. 4. The 1 GtC impulse response of the LGCC model (solid lines) applied under an equilibrium pre-industrial atmospheric CO_2 concentration initial condition. The dashed line is the fit of the TF eq. (10) to this response.

of $m_A(s)$ in this case. In this estimation we assume two physical constraints in order to aid the estimation. First, $b_0 = 1$, which simply states that one GtC of emissions adds one GtC to the atmosphere. This is shown in Fig. 2a as the unity integral gain (i.e. 1/s) in the forward path of the system. Second, $a_5 = 0$, which arises because the overall system is assumed to be perfectly conservative with respect to anthropogenic CO_2 emissions (i.e. there are no loss terms in Fig. 3). In systems terms, this gives

Table 1. Coefficients of determination (r^2) of five transfer function structures fitted to the data shown in Fig. 4

n	m	r^2
2	2	0.91549
3	3	0.99429
4	4	0.99994
5	5	0.99999
6	6	0.99999

Table 2. Parameter values associated with the TF response function eq. (10) fitted to the data in Fig. 4. The parameters a and b are parametrized through a least squares fit of $m_A(s)$ to the LGCC impulse response. The effects of noise bias on the parameter estimation are considered insignificant here given $m_A(s)$ is noise free and the model residuals are so small. The figures in parentheses denote two standard deviations on the parameter estimates

$$m_A(s) = \frac{b_0 s^4 + b_1 s^3 + b_2 s^2 + b_3 s + b_4}{s^5 + a_1 s^4 + a_2 s^3 + a_3 s^2 + a_4 s + a_5} \cdot u_E(s)$$

$b_0 = 1$	$a_1 = 1.06490$ (0.05317)
$b_1 = 0.89022$ (0.05175)	$a_2 = 0.24577$ (0.02616)
$b_2 = 0.15811$ (0.01877)	$a_3 = 0.01361$ (0.00183)
$b_3 = 0.00507$ (0.00069)	$a_4 = 0.0000811$ (0.0000112)
$b_4 = 0.0000123$ (0.0000017)	$a_5 = 0$

rise to a pole at 0.0 and hence an integrator (1/s) term in the overall system response rendering the system marginally stable. Interestingly, this integrator has profound implications for the design of mitigation strategies (see Jarvis et al., 2008).

As can be seen from Table 1, the trade-off of TF fit to the GCC impulse response against degree of TF parametrization falls off after $m = n = 5$. As a result, we conclude that all significant variation has been captured by the $m = n = 5$ TF response function, that is, eq. (10). Table 2 shows the corresponding parameter values of eq. (10) for the LGCC model pre-industrial case whilst Table 3 shows the equivalent SE parametrization for reference with previous studies.

Having identified and parametrized an appropriate TF, we can now consider the decomposition into an appropriate structure in sympathy with the LGCC model. Unfortunately, the decomposition does not simply involve reversing the operation detailed from eqs (7) to (10) because that process is a many-to-one operation and hence its inverse, one-to-many operation, is not unique. Therefore, we have developed a stepwise approach to decomposing eq. (10) back into the block diagram structure Fig. 2a. This is shown in Fig. 5 and detailed in Appendix B. The corresponding physical properties of the five box decomposition are given in Table 4.

Before offering an interpretation of the parameters derived from the response function, it is important to emphasis the fact

Table 3. SE response function inferred from the parallel decomposition of eq. (10). The figures in parentheses denote the 95% parameter confidence interval generated by means of 10^4 random draws from the covariance matrix structure associated with estimating the parameters in Table 2

$$m_a(t) = r_1 e^{(-t/\tau_1)} + r_2 e^{(-t/\tau_2)} + r_3 e^{(-t/\tau_3)} + r_4 e^{(-t/\tau_4)} + r_5 e^{(-t/\tau_5)}$$

$r_1 = 0.1517$ (0.1424 0.1621)	$\tau_1 = 1.3027$ (1.2474 1.3612)
$r_2 = 0.1505$ (0.1482 0.1527)	$\tau_2 = 4.5623$ (4.2355 4.9610)
$r_3 = 0.3330$ (0.3252 0.3406)	$\tau_3 = 14.023$ (13.867 14.176)
$r_4 = 0.2128$ (0.2126 0.2130)	$\tau_4 = 147.83$ (147.73 147.93)
$r_5 = 0.1517$ (0.1516 0.1517)	$\tau_5 = 0$

that the decomposition will involve some aggregation of the LGCC model properties. Starting with the land reservoir, the time constants of the vegetation and soil subcompartments are 5.36 and 26.57 yr, respectively, which compare favourably with actual values for LGCC of 5.40 and 27.80 yr. The vegetation and soil respiration rates k_{L2} and k_{L4}, and the litter fall rate k_{L3} also compare favourably with the Lenton (2000) pre-industrial values of 0.0920, 0.0375 and 0.0827 yr^{-1} (see Table 4). Although both plant and soil respiration rates are temperature-dependent in LGCC, this non-linearity appears to have not been excited by the 1 GtC perturbation.

For the response function decomposition the ocean dynamics are compartmentalized into the aggregated surface ocean and aggregated intermediate and deep ocean. In the LGCC model, the well mixed surface ocean consists of cold high-latitude and warm low-latitude compartments and the response function derived time constant of 1.5087 yr tallies with the cold, high latitude fraction of the surface ocean (see Table 5), highlighting the dominance of cold water CO_2 uptake in the LGCC model. The particulate flux is independent of CO_2 and hence is unperturbed in the unit impulse response (Caldeira et al., 2000; Harvey, 2000). The aggregated intermediate and deep ocean time constant of 337.78 yr compares favourably with the 349.14 yr specified in Kwon and Schnoor (1994) from which the LGCC model was derived. Again, the unit impulse appears insufficient to significantly alter the oceanic buffering of CO_2 partial pressure offered by the bicarbonate system. Given the nature of the aggregation in the response function derived decomposition it is not possible to compare the within-ocean exchange coefficients.

For the atmospheric component, the flux rates to land and ocean allied to the size of the atmosphere result in a time constant of 5.7235 yr which is marginally higher than the 3.3840 yr in the LGCC model. This slight mismatch arises from an inability to accurately capture the atmosphere–ocean exchange in the LGCC model. Better accuracy would be obtained if the oceanic component of the system was studied in isolation.

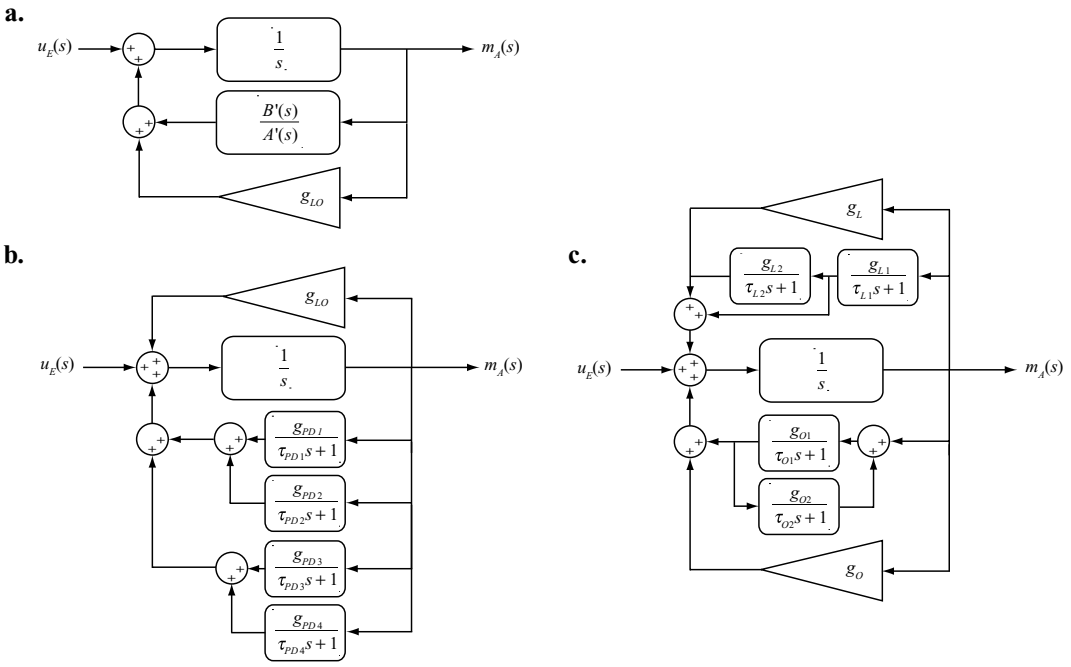

Fig. 5. Three-step decompositions of eq. (10) (for details see in Appendix B).

Table 4. Model parameters derived analytically from eq. (10). The properties of the TF decomposition shown in Fig. 2 are given by steady state gain g (GtC) and time constant τ (yr). The properties of five-box GCC are specified by effective carbon flux rates k (yr^{-1}) corresponding to the arrows in Figure 1. The figures in parentheses denote the 95% parameter confidence interval again generated from 10^4 random draws from the covariance matrix structure associated with estimating the parameters in Table 2

Model	Land component	Ocean component
TF decomposition	$g_{L1} = 0.0385$ (0.0355 0.0415)	$g_{O1} = 0.0916$ (0.0889 0.0942)
	$\tau_{L1} = 5.3610$ (4.9598 5.8305)	$\tau_{O1} = 1.5087$ (1.4280 1.5978)
	$g_{L2} = 0.8767$ (0.8306 0.9274)	$g_{O2} = 1.1652$ (1.0942 1.2433)
	$\tau_{L2} = 26.568$ (26.413 26.725)	$\tau_{O2} = 337.78$ (336.39 339.08)
	g_L = -0.0722 (-0.0684 -0.0761)	g_O = -0.1025 (-0.0994 -0.1056)
Five-box GCC	$k_{L1} = 0.0722$ (0.0684 0.0761)	$k_{O1} = 0.1025$ (0.0994 0.1056)
	$k_{L2} = 0.0994$ (0.0894 0.1097)	$k_{O2} = 0.5921$ (0.5605 0.6240)
	$k_{L3} = 0.0871$ (0.0818 0.0923)	$k_{O3} = 0.0708$ (0.0652 0.0766)
	$k_{L4} = 0.0376$ (0.0374 0.0379)	$k_{O4} = 0.00296$ (0.00295 0.00297)

5. Conclusions

From the preceding analysis we would conclude that, providing we acknowledge the affects of aggregation, then a macroscale interpretation of a parent model based on response function characterization is possible and that TF analysis is particularly promising in this regard because. However, it is important to stress some of the limitations of the approach.

First, effective TF estimation, identification and decomposition relies heavily on the information content of the input–output time series data. The unit impulse input is a useful excitation in the present context because the GCC system is rather

Table 5. Comparison of the corresponding reservoir time constants between LGCC and five-box GCC models (see Fig. 3 and Table 4)

Reservoir	Lenton (yr)	Five-box GCC (yr)
Atmosphere	∼3.3840	5.7235
Vegetation	∼5.4008	5.3610
Soil	∼27.800	26.568
High-latitude surface ocean	∼1.5305	1.5087
Intermediate and deep ocean	∼349.14	337.78

stiff[5] and the impulse excites the full range of dynamic modes. Sometimes it is not necessary to identify the full dynamic range of the parent model and a more parsimonious model structure may be formed by considering only behavioural modes that would be stimulated by an input with spectral properties achievable in the real world. For example, in mitigation analysis the specification of anthropogenic carbon emissions may not be interested in rapid variations in these emissions because they would violate certain socio-economic constraints (see Leedal, 2007 and Jarvis et al., 2008).

Second, some systems may be characterized by features such as pole-zero cancellation which would lead to an apparent reduction in the identified response function model order. Under these circumstances it would become very difficult if not impossible to retrieve an appropriate structural interpretation. Indeed, under many circumstances there are no structural interpretations of response functions (see e.g. Leedal, 2007).

Third, the framework we have employed is linear, whereas GCC models are only ever locally linear, although there are non-linear extensions to the linear TF framework that may be of particular use for further research in this area (e.g. Young, 2000).

In this paper, the approach used in Section 4 has been to infer the parent model properties from the identified response function which in turn is determined only by the data fitting process. This systems identification approach attempts to circumvent the need to impose a response function model order/structure a priori, hence attempting to make the process in some sense objective. Although this may actually be a hindrance when analysing LGCC given our prior knowledge of the model structure could be used to greatly enhance the parameter retrieval, when we move to more complex systems such as GCM's it will probably be the prudent approach given we will be far less confident about the nature of the aggregation in such models. Whether the aggregation of dynamic effects in GCMs is such that the response function says anything meaningful about GCMs remains to be seen, although we note that response functions do successfully capture global scale dynamics of GCMs remarkably well (Hasselman et al., 1993; 1997; Hooss, 2001; Lowe, 2003) and that simple models are frequently used to approximate the dynamics of GCMs (Raper et al., 2001; Meinshausen et al., 2008).

6. Acknowledgments

We would like to acknowledge two anonymous referees whose comments proved very helpful in the preparation of this paper.

[5]Refers to a system with dynamic modes distributed over a broad range of time constants.

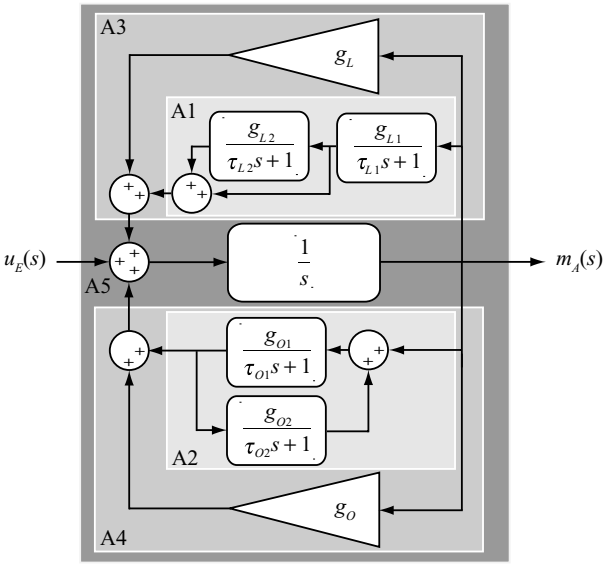

Fig. 6. The aggregation processes for forming eq. (10). The shaded areas show the derivations of eq. (A1)–(A6).

7. Appendix A: Deviation of the fifth-order TF based on a simplified generic GCC model structure

This appendix describes the procedure whereby the block diagram in Fig. 2a is aggregated to generate the generic form of fifth-order TF eq. (10). These steps are shown in Fig. 6.

Beginning with the shaded area (A1) in Figure A1,

$$\frac{g_{L1}}{\tau_{L1}s+1}+\frac{g_{L1}}{\tau_{L1}s+1}\cdot\frac{g_{L2}}{\tau_{L2}s+1}=\frac{\frac{g_{L1}}{\tau_{L1}}s+\left(\frac{g_{L1}+g_{L1}g_{L2}}{\tau_{L1}\tau_{L2}}\right)}{s^2+\left(\frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}}\right)s+\frac{1}{\tau_{L1}\tau_{L2}}} \tag{A1}$$

For the shaded area (A2) covering the ocean feedback,

$$\frac{\frac{g_{O1}}{\tau_{O1}s+1}}{1-\frac{g_{O1}}{\tau_{O1}s+1}\frac{g_{O2}}{\tau_{O2}s+1}}=\frac{\frac{g_{O1}}{\tau_{O1}}s+\frac{g_{O1}}{\tau_{O1}\tau_{O2}}}{s^2+\left(\frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}}\right)s+\left(\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}}\right)} \tag{A2}$$

Shaded area (A3) and (A4) are the net feedbacks from land and ocean respectively and give,

$$\frac{g_L s^2+\left(\frac{g_L\tau_{L1}+g_L\tau_{L2}+g_{L1}\tau_{L2}}{\tau_{L1}\tau_{L2}}\right)s+\frac{g_L+g_{L1}+g_{L1}g_{L2}}{\tau_{L1}\tau_{L2}}}{s^2+\left(\frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}}\right)s+\frac{1}{\tau_{L1}\tau_{L2}}} \tag{A3}$$

Shaded area (A4) produces,

$$\frac{g_O s^2+\left(\frac{g_O\tau_{O1}+g_O\tau_{O2}+g_{O1}\tau_{O2}}{\tau_{O1}\tau_{O2}}\right)s+\left(\frac{g_O-g_Og_{O1}g_{O2}+g_{O1}}{\tau_{O1}\tau_{O2}}\right)}{s^2+\left(\frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}}\right)s+\left(\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}}\right)} \tag{A4}$$

Finally, shaded area (A5) sums the net feedbacks eq. (A3) and (A4) and gives

$$m_A(s) = \frac{b_0 s^4 + b_1 s^3 + b_2 s^2 + b_3 s + b_4}{s^5 + a_1 s^4 + a_2 s^3 + a_3 s^2 + a_4 s + a_5} \cdot u_E(s) \qquad \text{(A5)}$$

where

$$\begin{aligned}
b_0 &= 1;\\
b_1 &= \frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}} + \frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}};\\
b_2 &= \frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}} + \frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}}\frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}} + \frac{1}{\tau_{L1}\tau_{L2}};\\
b_3 &= \frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}}\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}} + \frac{1}{\tau_{L1}\tau_{L2}}\frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}};\\
b_4 &= \frac{1}{\tau_{L1}\tau_{L2}}\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}};\\
a_1 &= \frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}} + \frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}} - (g_O + g_L);\\
a_2 &= \left(\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}} + \frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}}\frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}} + \frac{1}{\tau_{L1}\tau_{L2}}\right)\\
&\quad - \left[\frac{g_{L1}}{\tau_{L1}} + \frac{g_{O1}}{\tau_{O1}} + (g_O+g_L)\left(\frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}} + \frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}}\right)\right];\\
a_3 &= \left(\frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}}\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}} + \frac{1}{\tau_{L1}\tau_{L2}}\frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}}\right)\\
&\quad - \left[\frac{g_{L1}}{\tau_{L1}}\frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}} + \frac{g_{L1}+g_{L1}g_{L2}}{\tau_{L1}\tau_{L2}} + \frac{g_{O1}}{\tau_{O1}}\frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}}\right.\\
&\quad \left. + \frac{g_{O1}}{\tau_{O1}\tau_{O2}}\right]\\
&\quad - (g_O+g_L)\left(\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}} + \frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}}\frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}}\right.\\
&\quad \left.\left. + \frac{1}{\tau_{L1}\tau_{L2}}\right)\right];\\
a_4 &= \frac{1}{\tau_{L1}\tau_{L2}}\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}}\\
&\quad - \left[\frac{g_{L1}}{\tau_{L1}}\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}} + \frac{g_{L1}+g_{L1}g_{L2}}{\tau_{L1}\tau_{L2}}\frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}}\right.\\
&\quad \left. + \frac{1}{\tau_{L1}\tau_{L2}}\frac{g_{O1}}{\tau_{O1}} + \frac{g_{O1}}{\tau_{O1}\tau_{O2}}\frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}}\right]\\
&\quad - (g_O+g_L)\left(\frac{\tau_{L1}+\tau_{L2}}{\tau_{L1}\tau_{L2}}\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}}\right.\\
&\quad \left.\left. + \frac{1}{\tau_{L1}\tau_{L2}}\frac{\tau_{O1}+\tau_{O2}}{\tau_{O1}\tau_{O2}}\right)\right];\\
a_5 &= -\left[\frac{g_{L1}+g_{L1}g_{L2}}{\tau_{L1}\tau_{L2}}\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}} + \frac{g_{O1}}{\tau_{O1}\tau_{O2}}\frac{1}{\tau_{L1}\tau_{L2}}\right.\\
&\quad \left. + (g_O+g_L)\frac{1}{\tau_{L1}\tau_{L2}}\frac{1-g_{O1}g_{O2}}{\tau_{O1}\tau_{O2}}\right]; \qquad \text{(A6)}
\end{aligned}$$

8. Appendix B: Decompositions of a fifth-order TF into a GCC structure

8.1. Appendix B.1: Decomposition of Fig. 5a

Equation (10) gives,

$$m_A(s) = \frac{s^4 + b_1 s^3 + b_2 s^2 + b_3 s + b_4}{s^5 + a_1 s^4 + a_2 s^3 + a_3 s^2 + a_4 s} u_E(s) \qquad \text{(B1)}$$

Figure 5a yields,

$$m_A(s) = \frac{\frac{1}{s}}{1 - \frac{1}{s}\left[\frac{B'(s)}{A'(s)} + g_{LO}\right]} u_E(s) \qquad \text{(B2)}$$

where

$$\frac{B'(s)}{A'(s)} = \frac{b_{F0}s^3 + b_{F1}s^2 + b_{F2}s + b_{F3}}{(s+a_{F1})(s+a_{F2})(s+a_{F3})(s+a_{F4})} \qquad \text{(B3)}$$

Expanding eq. (B2) we find that

$$\begin{aligned}
a_1 &= a_{F1} + a_{F2} + a_{F3} + a_{F4} - g_{LO};\\
a_2 &= a_{F1}a_{F2} + a_{F1}a_{F3} + a_{F1}a_{F4} + a_{F2}a_{F3} + a_{F2}a_{F4}\\
&\quad + a_{F3}a_{F4} - g_{LO}(a_{F1} + a_{F2} + a_{F3} + a_{F4}) - b_{F0};\\
a_3 &= a_{F1}a_{F2}a_{F3} + a_{F1}a_{F2}a_{F4} + a_{F1}a_{F3}a_{F4} + a_{F2}a_{F3}a_{F4}\\
&\quad - g_{LO}(a_{F1}a_{F2} + a_{F1}a_{F3} + a_{F1}a_{F4} + a_{F2}a_{F3} + a_{F2}a_{F4}\\
&\quad + a_{F3}a_{F4}) - b_{F1};\\
a_4 &= a_{F1}a_{F2}a_{F3}a_{F4} - b_{F2} - g_{LO}(a_{F1}a_{F2}a_{F3} + a_{F1}a_{F2}a_{F4}\\
&\quad + a_{F1}a_{F3}a_{F4} + a_{F2}a_{F3}a_{F4});\\
a_5 &= -(b_{F3} + g_{LO}a_{F1}a_{F2}a_{F3}a_{F4}) = 0;
\end{aligned}$$

$$\begin{aligned}
b_1 &= a_{F1} + a_{F2} + a_{F3} + a_{F4};\\
b_2 &= a_{F1}a_{F2} + a_{F1}a_{F3} + a_{F1}a_{F4} + a_{F2}a_{F3} + a_{F2}a_{F4} + a_{F3}a_{F4};\\
b_3 &= a_{F1}a_{F2}a_{F3} + a_{F1}a_{F2}a_{F4} + a_{F1}a_{F3}a_{F4} + a_{F2}a_{F3}a_{F4};\\
b_4 &= a_{F1}a_{F2}a_{F3}a_{F4}; \qquad \text{(B4)}
\end{aligned}$$

It is interesting to note that a_{F1}, a_{F2}, a_{F3} and a_{F4} are derived by factoring the numerator of eq. (B1). This simplifies the process of deriving the analytical results from eq. (B4). For the physical interpretation, the current decomposition has not formally separated all the subsystems in the fourth-order TF eq. (B3) (see below).

8.2. Appendix B.2: Decomposition of Figure 5b

Looking at Fig. 5b, the sum of feedback terms of eq. (B1), that is, eq. (B3), is given by the following partial fraction expansion,

$$\frac{g_{PD1}}{\tau_{PD1}s+1} + \frac{g_{PD2}}{\tau_{PD2}s+1} + \frac{g_{PD3}}{\tau_{PD3}s+1} + \frac{g_{PD4}}{\tau_{PD4}s+1} + g_{LO}, \qquad \text{(B5)}$$

so eq. (B1) can be written as,

$$m_A(s) = \frac{\frac{1}{s}}{1 - \frac{1}{s}\left[\frac{g_{PD1}}{\tau_{PD1}s+1} + \frac{g_{PD2}}{\tau_{PD2}s+1} + \frac{g_{PD3}}{\tau_{PD3}s+1} + \frac{g_{PD4}}{\tau_{PD4}s+1} + g_{LO}\right]} u_E(s) \quad \text{(B6)}$$

8.3. Appendix B.3: Decomposition of Fig. 5c

Eq. (B5) can be reconstituted into its land and ocean components thus,

$$sys_L + sys_O + g_{LO} \quad \text{(B7)}$$

Substituting eq. (B7) into eq. (B6) one gets,

$$m_A(s) = \frac{\frac{1}{s}}{1 - \frac{1}{s}[sys_L + sys_O + g_{LO}]} u_E(s) \quad \text{(B8)}$$

where

$$sys_O = \frac{\left(\frac{g_{PD1}\tau_{PD4} + g_{PD4}\tau_{PD1}}{\tau_{PD1}\tau_{PD4}}\right)s + \left(\frac{g_{PD1}+g_{PD4}}{\tau_{PD1}\tau_{PD4}}\right)}{s^2 + \left(\frac{\tau_{PD1}+\tau_{PD4}}{\tau_{PD1}\tau_{PD4}}\right)s + \left(\frac{1}{\tau_{PD1}\tau_{PD4}}\right)} \quad \text{(B9)}$$

$$sys_L = \frac{\left(\frac{g_{PD2}\tau_{PD3} + g_{PD3}\tau_{PD2}}{\tau_{PD2}\tau_{PD3}}\right)s + \left(\frac{g_{PD2}+g_{PD3}}{\tau_{PD2}\tau_{PD3}}\right)}{s^2 + \left(\frac{\tau_{PD2}+\tau_{PD3}}{\tau_{PD2}\tau_{PD3}}\right)s + \left(\frac{1}{\tau_{PD2}\tau_{PD3}}\right)} \quad \text{(B10)}$$

Decomposing sys_O gives,

$$sys_O = \frac{\frac{g_{O1}}{\tau_{O1}s+1}}{1 - \frac{g_{O1}}{\tau_{O1}s+1}\frac{g_{O2}}{\tau_{O2}s+1}} \quad \text{(B11)}$$

Similarly, sys_L can be decomposed into

$$sys_L = \frac{g_{L1}}{\tau_{L1}s+1} + \frac{g_{L1}}{\tau_{L1}s+1} \cdot \frac{g_{L2}}{\tau_{L2}s+1} \quad \text{(B12)}$$

As can be seen from Fig. 6, eqs (B11) and (B12) represent the subsystems A1 and A2 whilst g_{LO} subdivides into g_L and g_O.

References

Caldeira, K., Hoffert, M. I. and Jain, A. 2000. Simple ocean carbon models. In: *The Carbon Cycle* (eds T. M. L. Wigley and D. S. Schimel).Cambridge University Press, Cambridge, 199–211.

Dowell, E. H. 1996. Eigenmode analysis in unsteady aerodynamics: reduced order models. *AIAA J.* **34**, 1578–1588.

den Elzen, M. G. J. and Lucas, P. 2005. The FAIR model: a tool to analyse environmental and costs implications of regimes of future commitments. *Environ. Modell. Assess.* **10**, 115–134.

Enting, I. G. 1990. Ambiguities in the calibration of carbon cycle models. *Inverse Problems* **6**, L39–L46.

Enting, I. G. 2007. Laplace transform analysis of the carbon cycle. *Environ. Modell. Software* **22**, 1488–1497.

Enting, I. G. and Trudinger, C. M. 2002. Modelling earth system change: I. Validating parameterisations for attribution calculations. CSIRO Atmospheric Research Technical Paper 56.

Garnier, H., Mensler, M. and Richard, A. 2003. Continuous-time model identification from sampled data: implementation issues and performance evaluation. *Int. J. Contr.* **76**, 1337–1357.

Godfrey, K. 1982. *Compartmental Models and Their Application*. Academic Press, London.

Harvey, L. D. D. 2000. Response of the carbon cycle and other biogeochemical cycles: translating emissions of GHGs and aerosols into concentrations and radiative forcing. In: *Climate and Global Environmental Change* (ed. L. D. D. Harvey). Pearson Education Ltd, Essex.

Hasselmann, K., Sausen, R., Maier-Reimer, E. and Voss, R. 1993. On the cold start problem in transient simulations with coupled atmosphere-ocean models. *Clim. Dyn.* **9**, 53–61.

Hasselmann, K., Hasselmann, S., Giering, R., Ocana, V. and Storch, H. V. 1997. Sensitivity study of optimal CO_2 emission paths using a simplied structural integrated assessment model (SIAM). *Clim. Change* **37**, 345–386.

Hooss, G. 2001. *Aggregate Models of Climate Change Development and Applications*. PhD Thesis. Max-Planck-Institut für Meteorologie, Hamburg, Germany.

Hooss, G., Voss, R., Hasselman, K., Maier-Reimer, E. and co-authors. 2001. A nonlinear impulse response model of the coupled carbon cycle-climate system (NICCS). *Clim. Dyn.* **18**, 189–202.

Jarvis, A. J., Young, P. C., Leedal, D. L. and Chotai, A. 2008. A robust sequential CO_2 emissions strategy based on optimal control of atmospheric CO_2 concentrations. *Clim. Change* **86**, 357–373

Joos, F., Bruno, M., Fink, R., Siegenthaler, U., Stocker, T. F. and co-authors. 1996. An efficient and accurate representation of complex oceanic and biospheric models of anthropogenic carbon uptake. *Tellus* **48B**, 397–417.

Kwon, O-Y. and Schnoor, J. L. 1994. Simple global carbon model: the atmosphere-terrestrial biosphere-ocean interaction. *Global Biogeochem. Cycles* **8**, 295–305.

Leedal, D. T. 2007. *Sequential Decision Making in Climate Change Mitigation: A Control System Perspective*. PhD Thesis.Lancaster University, Lancaster, UK.

Lenton, T. M. 2000. Land and ocean carbon cycle feedback effects on global warming in a simple Earth system model. *Tellus* **52B**, 1159–1188.

Lowe, J. 2003. Parameters for tuning a simple climate model. (accessed from http://unfccc.int/resource/brazil/climate.html).

Maier-Reimer, E. and Hasselmann, K. 1987. Transport and storage of CO_2 in the ocean – an inorganic ocean-circulation carbon cycle model. *Clim. Dyn.* **2**, 63–90.

Maier-Reimer, E., Mikolajewicz, U. and Winguth, A. 1996. Future ocean uptake of CO_2: interaction between ocean circulation and biology. *Clim. Dyn.* **12**, 711–721.

Meinshausen, M., Raper, S. C. B. and Wigley, T. M. L. 2008. Emulating IPCC AR4 atmosphere-ocean and carbon cycle models for projecting global-mean, hemispheric and land/ocean temperatures: MAGICC 6.0. *Atmos. Chem. Phys. Discuss.* **8**, 6153–6272.

Meyer, R., Joos, F., Esser, G., Heimann, M., Hooss, G. and co-authors. 1999. The substitution of high-resolution terrestrial biosphere models and carbon sequestration in response to changing CO_2 and climate. *Global Biogeochem. Cycles* **13**, 785–802.

Moore, B. 1981. Principle component analysis in linear systems: controllability, observability, and model reduction. *IEEE Trans. Auto. Contr.* **26**, 17–32.

Nordaus, W. D. and Boyer, J. 2000. *Warming the World: Economic Models of Global Warming*. MIT Press, Cambridge, MA.

Orr, J. C., Maier-Reimer, E., Mikolajewicz, U., Monfray, P., Sarmiento, J. L., and co-authors. 2001. Estimates of anthropogenic carbon uptake from four three-dimensional global ocean models. *Global Biogeochem. Cycles* **15**, 43–60.

Raper, S. C. B., Gregory, J. M. and Osborn, T. J. 2001. Use of an upwelling-diffusion energy balance climate model to simulate and diagnose A/OGCM results. *Clim. Dyn.* **17**, 601–613.

Sarmiento, J., Hughes, T., Stouffer, R. and Manabe, S. 1998. Simulated response of the ocean carbon cycle to anthropogenic climate warming. *Nature* **393**, 245–252.

Tang, D., Kholodar, D., Juang, J.-N. and Dowell, E. H. 2001. System identification and proper orthogonal decomposition applied to unsteady aerodynamics. *AIAA J.* **39**, 1569–1576.

Thompson, M. V. and Randerson, J. T. 1999. Impulse response functions of terrestrial carbon cycle models: method and application. *Global Change Biol.* **5**, 371–394.

Wigley, T. M. L. 1991. A simple inverse carbon cycle model. *Global Biogeochem. Cycles* **5**, 373–382.

Wigley, T. M.L, Richels, R. and Edmonds, J. A. 1996. Economic and environmental choices in the stabilization of atmospheric CO_2 concentrations. *Nature* **379**, 242–245.

Young, P. C. 1999. Data-based mechanistic modeling, generalized sensitivity and dominant mode analysis. *Comput. Phys. Commun.* **117**, 113–129.

Young, P. C. 2000. Stochastic, dynamic modeling and signal processing: time variable and state dependent parameter estimation. In: *Nonlinear and Nonstationary Signal Processing* (eds W. J. Fitzgerald, A. Walden, R. Smith and P. C. Young). Cambridge University Press, Cambridge, 74–114.

Young, P. C. and Beven, K. J. 1994. Data-based mechanistic modeling and the rainfall-flow nonlinearity. *Environmetrics* **5**, 335–363.

Young, P. C. and Garnier, H. 2006. Identification and estimation of continuous-time, data-based mechanistic models for environmental systems. *Environ. Model Software* **21**, 1055–1072.

Young, P. C., Parkinson, S. and Lees, M. J. 1996. Simplicity out of complexity: Occam's razor revisited. *J. Appl. Stat.* **23**, 165–210.