



# Extreme Sensitivity and Climate Tipping Points

Peter Ashwin<sup>1</sup> · Anna S. von der Heydt<sup>2</sup>

Received: 10 May 2019 / Accepted: 29 October 2019 / Published online: 13 November 2019  
© The Author(s) 2019

## Abstract

A climate state close to a tipping point will have a degenerate linear response to perturbations, which can be associated with extreme values of the equilibrium climate sensitivity (ECS). In this paper we contrast linearized (‘instantaneous’) with fully nonlinear geometric (‘two-point’) notions of ECS, in both presence and absence of tipping points. For a stochastic energy balance model of the global mean surface temperature with two stable regimes, we confirm that tipping events cause the appearance of extremes in both notions of ECS. Moreover, multiple regimes with different mean sensitivities are visible in the two-point ECS. We confirm some of our findings in a physics-based multi-box model of the climate system.

**Keywords** Climate sensitivity · Tipping point · Energy balance model · Stochastic climate model

## 1 Introduction

The equilibrium climate sensitivity (ECS) is widely used as a measure for expected future global warming. Following Charney’s definition [7], the ECS is the increase in global mean surface temperature (GMST) per radiative forcing change after the fast-acting feedback processes in the earth system reach equilibrium. Fast-acting means here that those processes are faster than the time-horizon for global mean temperature evolution that interests us, typically taken to be 100 years [32].

The value of ECS remains not very well constrained, as the expected warming per doubling of atmospheric CO<sub>2</sub> still contains a considerable uncertainty of 1.5–4.5 °C [22]. In fact, this range has not changed much since first ECS estimates based on energy balance argu-

---

Communicated by Valerio Lucarini.

---

✉ Peter Ashwin  
P.Ashwin@exeter.ac.uk  
Anna S. von der Heydt  
A.S.vonderHeydt@uu.nl

<sup>1</sup> Centre for Systems, Dynamics and Control, Department of Mathematics, University of Exeter, Exeter EX4 4QF, UK

<sup>2</sup> Institute for Marine and Atmospheric Research, Department of Physics, also Centre for Complex Systems Studies, Utrecht University, Princetonplein 5, 3584CC Utrecht, The Netherlands

ments, despite enormous developments in climate modelling, and improved observational methods [23]. In particular, large temperature changes and dangerous climate change as a consequence of increased atmospheric CO<sub>2</sub> cannot be excluded. For example, climate observations from the instrumental period have not narrowed down the range of expected climate change mainly because of uncertainties in quantification of the forcing [33]. Recently there have been concerted attempts to constrain ECS using emergent constraints in climate models [9].

Some sources of uncertainty for ECS lie in the classical measurement or model uncertainty, although in particular for observations the quantification of the applied forcing generally contains the largest uncertainties. Furthermore, by going back in time further than the instrumental period (e.g. last millennium, glacial cycles or even millions of years of palaeoclimate data) the uncertainty in both the forcing and the global mean temperature response becomes significant [23]. There has been a debate as to the cause of the uncertainty and extremes of the ECS distribution, in particular the long tail towards high sensitivity values. On the one hand, [31] suggest it is an inevitable consequence of nonlinear transformation of normally distributed feedbacks that appear in the denominator when calculating ECS. On the other hand, [41] suggest they are a sign of ‘tipping points’ owing to nonlinearities in the system—this has generated a lively debate. In this paper we highlight (a) the notion of ECS can usefully be generalised to a truly nonlinear geometric notion: the two-point sensitivity and (b) the distribution of ECS values is a valuable tool for characterising both state-dependent (feedback) dynamics and tipping of the climate system.

The climate system exhibits both internal and forced variability on many timescales. The consequence is that any ‘equilibrium’ is only relative to fixing part of the feedback processes that are internal to the climate system, in particular the ‘slower’ part. This requires an assumption that a time scale separation (into fast and slow processes) exists and the time scale of interest sits between fast and slow. For climate model simulations and observations of the last century there might be a time horizon where this is a reasonable assumption [33], but as we include palaeoclimate data and model simulations into the estimate of ECS this assumption needs to be carefully evaluated. In particular, methods to estimate ECS from palaeoclimate data or models differ from those of (short) climate model simulations; the latter generally derive ECS from the decay of the energy imbalance at the top of the atmosphere induced by a instantaneous doubling or quadrupling of CO<sub>2</sub> [18]; palaeoclimate reconstructions instead make the assumption that the reconstructed climate is in (energetic, short time scale) equilibrium and compare different of these ‘equilibria’ to each other for estimating ECS. Without compensating for slow feedback processes, palaeoclimate records give the so-called earth system sensitivity (ESS) that includes the effect of slow processes and boundary conditions (e.g. geography, vegetation and land ice) [29]. If estimates of these slow processes are available then ECS can be estimated from the ESS under an assumption of time scale separation [32,38].

Note that the ECS is usually thought of as a linearized response of the GMST to perturbations in the radiative balance of the earth. Next to incoming (short-wave) and outgoing (long-wave) radiation, feedback processes in the climate system play an important role in determining the ECS. In their sum these feedback processes tend to enhance ECS (net positive feedback) and associated time scales vary from fast to very slow. Examples of fast feedback processes include cloud feedbacks, water vapour feedback and sea-ice processes. The strength of each of these feedback processes depends on the background (long-term mean) climate state [39] and it is therefore not surprising that the sum of the fast feedback processes varies over time in particular when considering climate states far back in time and under very different boundary conditions. For example, from palaeoclimate records together

with ice sheet modelling, it has been found that ECS varies considerably between glacial and interglacial states [15,24,25]. Both present-day and Palaeogene climate model simulations suggest state-dependence of ECS due to feedback processes [6,30]. For example, Transient Climate Response (TCR) considers the deep ocean warming as a slow process.

Abrupt climate shifts have occurred in the past climate system and therefore seem likely to occur in the future for a variety of reasons [27,35]. In the present climate system, potential tipping elements have been identified some of which may have a considerable impact on future values of GMST [14,27]. Even those tipping elements that have little effect on the GMST may cause significant regional damage and/or contribute to global mean climate change by triggering cascades of transitions involving other tipping elements [11]. Across such an abrupt transition there is a breakdown of the assumption of a linear response to perturbations, suggesting that the ECS does not adequately represent the temperature response to radiative perturbations [4,40]. In practice, when deriving ECS from palaeoclimate time series, which include abrupt transitions, these shifts may lead to extreme values of ECS.

In this paper we show that more general notions of ECS can be useful in understanding the response of a climate state to changes in radiative forcing—in addition to an ‘instantaneous’ linearised notion of ECS we explore ‘incremental’ and ‘two-point’ climate sensitivities that are distributions related to dynamic properties of the climate system: they characterise the geometry of the dynamics and are not simply estimators of a ‘mean ECS’. In fact the distributions of sensitivities reflect the intrinsic uncertainty due to climate system dynamics. The paper is organized as follows: Sect. 2 introduces these notions of ECS and relates them to the underlying climate dynamics. We illustrate these concepts using a global energy balance model in Sect. 3. In particular, we relate properties of extremes of the ECS to the presence of tipping points and multistability. Section 4 finishes with a discussion of conclusions and some challenges for the future. Appendix extends results of [38] and examines extremes of this sensitivity associated with tipping points in a more realistic physics-based multi-box model of the glacial cycles by Gildor and Tziperman [17].

## 2 Sensitivities and the Climate Attractor

In order to understand variability, abrupt transitions and response to perturbations we consider the climate system as a high-dimensional multiscale complex dynamical system whose evolving trajectories form a *climate attractor*. The ECS can be defined on this attractor and regimes or states may be identified where a linear approximation of the response may be reasonable. Tipping points visible in the GMST will show up as large but occasional shifts between different ‘climate regimes’ of the attractor, or indeed different attractors. We visualise the attractor by projection onto climate observables relevant for determining ECS, i.e. the GMST  $T$  and the radiative forcing  $R$  per unit area [38]. Consider the energy balance model

$$c_T \frac{dT}{dt} = R_{\text{forcing}} + R_{\text{slow}} + R_{\text{fast}} - R_{\text{OLW}}, \quad (1)$$

where the left hand side represents the rate of change of the global mean surface temperature  $T$  (with specific heat capacity  $c_T$ ) and on the right hand side  $R_{\text{forcing}}$  is the (external) radiative forcing (including changes in  $\text{CO}_2$ ),  $R_{\text{slow}}$  ( $R_{\text{fast}}$ ) is the radiative perturbation due to all slow (fast) feedback processes within the climate system and  $R_{\text{OLW}}$  is the outgoing longwave radiation, respectively. Following the formalism of [32], the specific climate sensitivity is

$$S_{\text{forcing,slow}} = \frac{\Delta T}{\Delta R_{\text{forcing}} + \Delta R_{\text{slow}}} \approx \frac{dT}{d(R_{\text{forcing}} + R_{\text{slow}})}, \tag{2}$$

which equals the Charney sensitivity  $S$  if  $\Delta R_{\text{slow}}$  is the sum of all slow feedback processes contributing to the ECS (and under the assumption of time scale separation). In practise, only some of the slow processes are accessible from palaeoclimate records (e.g. only land ice), in which case the specific climate sensitivity is only an approximation of the Charney sensitivity [32] (e.g.  $S_{[CO_2,LI]}$  is the specific climate sensitivity considering only land ice changes as slow feedback). This ECS gives a linear prediction for change in temperature:

$$T' = T + S_{\text{forcing,slow}} (\Delta R_{\text{forcing}} + \Delta R_{\text{slow}}). \tag{3}$$

For a specific energy balance model including regime shifts we can explicitly calculate ECS for the different regimes, see Sect. 3. We note that several other authors have highlighted the need to improved notions of ECS: this includes [8] who propose to use a measure-based approach to understand climate sensitivity and [12] who consider conditional climate sensitivities constrained by temperature, coupled with resilience measures for switching to other regimes.

### 2.1 Observation of the Climate Attractor

We consider the climate system as a high dimensional dynamical system that evolves along trajectories  $x(t)$  according to a smooth flow

$$x(t) = \varphi_t(x_0) \tag{4}$$

where  $x \in X$  represents the instantaneous state of the climate system in some high dimensional state space and  $\varphi_t(x_0)$  evolves the initial state  $x_0$  along by a time  $t$ . The global mean temperature  $T : X \rightarrow \mathbb{R}$  and radiative forcing  $R : X \rightarrow \mathbb{R}$  are considered to be observables of the underlying dynamical system on  $X$ . We assume that the dynamics of  $x$  are stationary, i.e. that there is a natural probability measure  $M$  on  $X$  such that typical trajectories  $x(t)$  of (4) satisfy

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_{s=0}^t p(x(s)) ds = \int_X p(x) dM(x). \tag{5}$$

for typical  $x_0$  and any integrable observable  $p : X \rightarrow \mathbb{R}$ , i.e. the long-time average of  $p$  can be computed using an ergodic hypothesis, by averaging over the measure  $M$  in phase space. This implies that, for any open set  $\mathcal{A} \subset X$ , the long-term average proportion of time a typical trajectory spends in  $\mathcal{A}$  is  $M(\mathcal{A})$ . For small enough perturbations, linear response theory suggests a linear change in mean observables: see for example [28].

In [38] it is supposed there is a stationary measure  $\mu$  of points in the  $(\Delta R_{[CO_2,LI]}, T)$ -plane according to how often they are visited over asymptotically long times, i.e. for any measurable subset  $\mathcal{A} \subset \mathbb{R}^2$  we define

$$\mu(\mathcal{A}) := \lim_{t \rightarrow \infty} \frac{1}{t} \int_{s=0}^t \chi_{\mathcal{A}}(\Delta R_{[CO_2,LI]}(s), T(s)) ds. \tag{6}$$

for typical initial condition, where  $\chi_{\mathcal{A}}$  is the indicator function,  $\chi_{\mathcal{A}}(\Delta R, T) = 1$  if  $(\Delta R, T) \in \mathcal{A}$  and  $= 0$  otherwise. Note that applying (5) with  $p(x) = \chi_{\mathcal{A}}(x)$  (where  $\chi_{\mathcal{A}}(x) = 1$  if  $x \in \mathcal{A}$  and 0 otherwise) gives

$$\mu(\mathcal{A}) = M(\{x : (R_{[CO_2,LI]}(x), T(x)) \in \mathcal{A}\}). \tag{7}$$

In other words, the measure  $\mu$  is simply a projection of a natural measure  $M$  on the ‘climate attractor’ onto the two observables  $(R_{[CO_2, LI]}, T)$ . In general we will consider throughout

$$\Delta R_{[CO_2, LI]} = R_{[CO_2, LI]} - \tilde{R}_{[CO_2, LI]}$$

i.e. the change in radiative forcing relative to some fixed reference level  $\tilde{R}_{[CO_2, LI]}$  usually chosen as the level during the pre-industrial climate.

### 2.2 Incremental and Two-Point Sensitivities

In order to predict the temperature at some fixed future time-horizon  $\Delta t$  in response to a change in radiative forcing, we consider the quotient (2) for fixed changes in time. In this case we can view the distribution of what we call *incremental sensitivities* as the spread of trajectories from the current estimated values of instantaneous  $(\Delta R, T)$ , where from now on we write the radiative forcing corrected for slow feedback as  $R$ . On the other hand we can consider a time-independent choice of pairs of points on the climate attractor to obtain *two-point sensitivities*.

Let us assume the current state at time  $t = 0$  of the climate system is given by a measure  $\sigma_0$  on some high dimensional phase space  $X$  (that projects onto a point  $A$  in Fig. 1). Note that this will always be a measure rather than a point because of lack of knowledge of sub-grid parametrized processes (e.g. [37]) but it will project onto the current values  $(\Delta R_0, T_0) = (\Delta R(x), T(x))$  for all  $x$  in the support of  $\sigma_0$ . As time progresses, this state will spread to give a measure at time  $t$  that is

$$\sigma_t(A) = \sigma(\varphi_{-t}(A))$$

for any  $A \subset X$  (Fig. 1 shows a trajectory in black and others from the ensemble starting at  $A$  in grey). The *incremental sensitivity* for a time interval  $\Delta t$  is then

$$S_0^{\Delta t}(x) = \frac{T(\varphi_{\Delta t}(x)) - T_0}{\Delta R(\varphi_{\Delta t}(x)) - \Delta R_0} \tag{8}$$

with distribution

$$\mathbb{P}(S_0^{\Delta t} \in \mathcal{A}) = \sigma(\{x : S_0^{\Delta t}(x) \in \mathcal{A}\}).$$

Over long time, if there is decay of correlations and mixing of trajectories on the climate attractor [28,36,37] then  $\sigma_{\Delta t} \rightarrow M$  in the weak sense as  $\Delta t \rightarrow \infty$ , and so we expect the distribution of long-term incremental sensitivities for  $\Delta t \rightarrow \infty$  becomes time-independent for typical trajectories within the attractor:

$$\mathbb{P}(S_0^\infty \in \mathcal{A}) = M(\{x : S_0^\infty(x) \in \mathcal{A}\}). \tag{9}$$

where

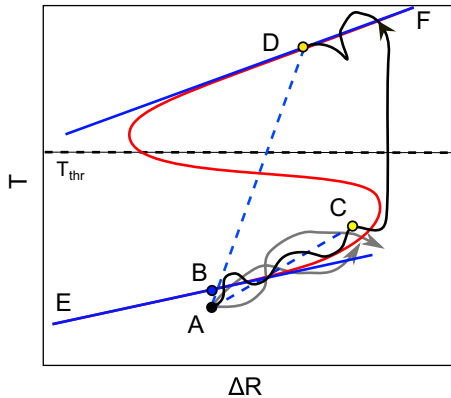
$$S_0^\infty(x) = \frac{T(x) - T_0}{\Delta R(x) - \Delta R_0}.$$

Note that (7) means that the distribution of long-term sensitivities starting at  $(\Delta R_0, T_0)$  can be written in terms of the geometry of the projected measure  $\mu$

$$\mathbb{P}(S_{0,1}^\infty \in \mathcal{A}) = \mu(\{(\Delta R_1, T_1) : S_{0,1}^\infty \in \mathcal{A}\}) \tag{10}$$

where we define the *two-point sensitivity* as

$$S_{0,1}^\infty = \frac{T_1 - T_0}{\Delta R_1 - \Delta R_0}. \tag{11}$$



**Fig. 1** Schematic diagram to demonstrate the instantaneous, incremental and two-point sensitivities for an ensemble of trajectories starting at point  $A$  that evolves on the climate attractor. The equilibrium of a mean energy balance model is shown as a red curve. For a specific trajectory (shown in black) in the ensemble the slopes of  $AC$  and  $AD$  correspond to incremental sensitivities (for fixed  $\Delta t$ ) or two-point sensitivities (for varying  $\Delta t$ ). The instantaneous sensitivity is the slope of the tangent to the closest solution in the equilibrium model at  $B$ . Note that the cold regime  $T \leq T_{\text{thr}}$  and the warm regime  $T > T_{\text{thr}}$  have different asymptotic (instantaneous) sensitivities corresponding to slopes  $EB$  and  $DF$  respectively (Color figure online)

The distribution of long-term incremental sensitivities (10) for a generic choice of the initial climate state suggests [38] a time-independent notion of climate sensitivity that can be found by picking pairs of points  $(R_{0,1}, T_{0,1})$  independently and identically distributed according to  $\mu$  and evaluating (2).

This means that for any  $\mathcal{A} \subset \mathbb{R}$  we can use  $\mu$  to assign a probability to the sensitivity being in  $\mathcal{A}$ :

$$\mathbb{P}(S_{0,1}^\infty \in \mathcal{A}) := \mu \times \mu (\{(\Delta R_0, T_0), (\Delta R_1, T_1) : S_{0,1}^\infty \in \mathcal{A}\}). \tag{12}$$

with  $S_{0,1}^\infty$  defined as in (11). This gives, in some sense, a maximal set of possibilities for the sensitivities in that it compares the observables  $T$  and  $\Delta R$  over all possible time points and possible trajectories of the system. This is comparable to the conditional climate sensitivity of [12] except rather than dividing into regimes, they restrict to deviations of temperature at most  $\delta_T$  from  $T_0$ . In the case that the sensitivity is fixed at  $S_0$ , note that  $S_{0,1}^\infty$  is a Dirac  $\delta$ -distribution centred at  $S_0$ .

### 2.3 Sensitivities and Climate Regimes

By partitioning the climate attractor into a number of regimes, we can condition the sensitivities on staying within a regime, or undergoing a transition between regimes. By making an optimal partition of the attractor projected into  $(\Delta R, T)$  space we can hope to find localised distributions of sensitivities for pairs in the same regime. As in [38] we consider these sensitivities conditional on climate regime by partitioning  $\mu$  into two distributions

$$\mu = \mu_C + \mu_W$$

corresponding to being in a cold (C) or warm (W) state. In our case we set

$$\begin{aligned} \mu_C(\mathcal{A}) &= \mu(\mathcal{A} \cap \{(\Delta R, T) : T \leq T_{thr}\}), \\ \mu_W(\mathcal{A}) &= \mu(\mathcal{A} \cap \{(\Delta R, T) : T > T_{thr}\}) \end{aligned}$$

for any measurable  $\mathcal{A} \subset \mathbb{R}^2$  and some threshold temperature  $T_{thr}$ . As in [38] we define distributions of conditional sensitivities by

$$\mathbb{P}(S^{WW} \in \mathcal{A}) := \mu_W \times \mu_W (\{(\Delta R_0, T_0), (\Delta R_1, T_1) : S_{0,1}^\infty \in \mathcal{A}\}) \tag{13}$$

Conditional sensitivities for changes of regime are for example

$$\mathbb{P}(S^{CW} \in \mathcal{A}) := \mu_C \times \mu_W (\{(\Delta R_0, T_0), (\Delta R_1, T_1) : S_{0,1}^\infty \in \mathcal{A}\}) \tag{14}$$

The distribution of sensitivities (12) is then a sum of the four conditional sensitivities

$$\mathbb{P}(S_{0,1}^\infty \in \mathcal{A}) = \mathbb{P}(S^{CC} \in \mathcal{A}) + \mathbb{P}(S^{CW} \in \mathcal{A}) + \mathbb{P}(S^{WC} \in \mathcal{A}) + \mathbb{P}(S^{WW} \in \mathcal{A}).$$

Moreover, (14) means that we have a symmetry

$$\mathbb{P}(S^{CW} \in \mathcal{A}) = \mathbb{P}(S^{WC} \in \mathcal{A}).$$

The distributions of  $S^{CW}$  and  $S^{WC}$  correspond to choices of pairs across the two regimes: these distributions are associated with ‘tipping between regimes’. Even though the two-point sensitivities may measure states very far apart in time, we will see that extreme values of the sensitivity are usually associated with choice of points from two different regimes.

### 3 Sensitivity and Tipping in Climate Models

To illustrate the notions of instantaneous and two-point sensitivities, we consider a conceptual energy balance model: a more complex model is briefly discussed in Appendix. We consider a variant of the Budyko–Ghil–Sellers energy balance model [5,16,34] for GMST. This model builds on [12,41] and has multiple regimes with state-dependent sensitivity in each. It is a special case of (1) for global mean surface temperature  $T(t)$  with atmospheric CO<sub>2</sub> concentration  $C(t)$  as a parameter:

$$c_T \frac{dT}{dt} = F(T, C) := \left[ Q_0(1 - \alpha(T)) + A \ln \left( \frac{C}{C_0} \right) - \epsilon(T)\sigma T^4 \right] \tag{15}$$

For this equation,  $Q_0$  represents the solar input modulated by the temperature-dependent albedo  $\alpha(T)$ . The change in radiative forcing due to atmospheric greenhouse gases is

$$\Delta R_{[CO_2]} = A \ln(C/C_0).$$

where  $A = 5.35 \text{ Wm}^{-2}$  is the direct forcing effect of CO<sub>2</sub> and  $C_0$  represents pre-industrial CO<sub>2</sub> levels. Finally, the outgoing long wave radiation  $\sigma T^4$  is modified by a temperature-dependent emissivity  $0 < \epsilon(T) < 1$ .

We consider a temperature-dependent emissivity decreasing from one plateau to a lower one because of changes in water vapour and cloud feedbacks. There are other choices [41], but for simplicity we assume here

$$\epsilon(T) = \epsilon_1 + \frac{\epsilon_2 - \epsilon_1}{2} \left[ 1 + \tanh \left( \frac{T - T_0}{T_\epsilon} \right) \right].$$

where  $\epsilon_1$  and  $\epsilon_2$  are the limit emissivities for low and high temperatures respectively,  $T_0$  is the threshold and  $T_\epsilon > 0$  corresponds to the range of temperatures over which there is variation (see Fig. 2b).

Note that water vapour feedback is sometimes included in the  $CO_2$  term, resulting in an additional constant and modified  $A$  [12,20]. Here we separate radiative forcing due to  $CO_2$  and temperature-dependent water feedbacks in emissivity. As in [12], we assume that the albedo varies with temperature due to changes in land-ice feedback processes: we assume there are threshold temperatures  $T_1 < T_2$  associated with changes of albedo  $\alpha(T)$  and define a function

$$\Sigma(T) = \frac{(T - T_1)}{T_2 - T_1} H(T - T_1) H(T_2 - T) + H(T - T_2) \tag{16}$$

that switches from 0 for  $T < T_1$  to 1 for  $T > T_2$ :  $H(T)$  is approximately a Heaviside unit step function and we use a smooth approximation

$$H(T) = (1 + \tanh(T/T_\alpha))/2$$

as in [12]. As in that paper, we write the albedo

$$\alpha(T) = \alpha_1(1 - \Sigma(T)) + \alpha_2\Sigma(T)$$

so that it changes smoothly from a higher albedo  $\alpha_1$  in the presence of more ice surface ( $T < T_1$ ) to a lower  $\alpha_2$  in the presence of more ocean surface ( $T > T_2$ ), see Fig. 2a. Note that [12] consider a global transition from ice-covered to ocean-covered earth - here we model a large but regional change in ice cover resulting in a smaller contrast in global albedo between the two states; our choice of parameters might be more realistic for albedo variations between glacial and interglacial states.

We add a stochastic term to (15) that represents unresolved subgrid processes with a fixed amplitude  $\eta_T$ :

$$c_T dT = F(T, C)dt + \eta_T dW_T. \tag{17}$$

The motivation for noise on the  $T$  variable is to model internal variability of climate heat transport processes. The parameters listed in Table 1 are used, except where specified. Note that the deterministic equilibria of (15) are at  $F(T, C) = 0$ , which gives

$$C = \Gamma(T) := C_0 \exp \left[ \frac{\epsilon(T)\sigma T^4 - Q_0(1 - \alpha(T))}{A} \right]. \tag{18}$$

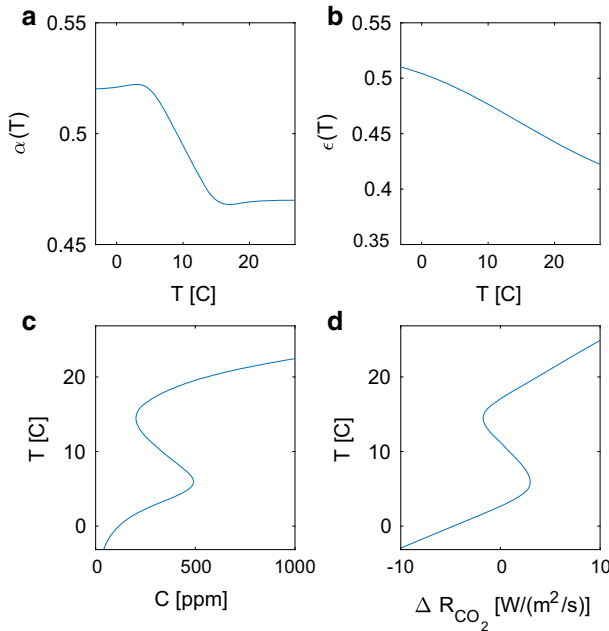
From (18), this means we have equilibria at

$$\Delta R_{[C O_2]} = A \ln(\Gamma(T)/C_0) = \epsilon(T)\sigma T^4 - Q_0(1 - \alpha(T)) \tag{19}$$

Figure 2 illustrates temperature dependence of albedo and emissivity as well as the resulting equilibrium forcing  $\Delta R = A \ln(\Gamma(T)/C_0)$  needed to give this temperature. Note there is a unique equilibrium for each  $T$ , but not necessarily for each  $C$ : as discussed in [12,41] there are three branches of equilibria for a range of  $C$ : for the parameters used there is bistability in the region

$$- 1.744 \text{ W m}^{-2} < \Delta R < 3.004 \text{ W m}^{-2}, \quad 202 \text{ ppm} < C < 490 \text{ ppm}, \tag{20}$$





**Fig. 2** Behaviour of the equilibrium energy balance model (15) with parameters as in Table 1. **a** temperature-dependence of albedo  $\alpha(T)$  and **b** emissivity  $\epsilon(T)$ ; **c** CO<sub>2</sub> and **d** radiative forcing levels necessary to give temperature equilibria, corresponding to (18) and (19), respectively. Note the region of multistability, and temperature-dependence of the sensitivity corresponding to slopes in the bottom right figure

denoted using the red lines in Figs. 3 and 5. We can define the instantaneous sensitivity<sup>1</sup> as  $S = 1/\lambda$ , where

$$\lambda = \frac{d}{dT} \Delta R_{[CO_2]} = [\epsilon'(T)T + 4\epsilon]\sigma T^3 + Q_0\alpha'(T) \tag{21}$$

is the total feedback factor in this model:  $S$  corresponds to the slope of the tangent of the equilibrium (non-stochastic) model (see Fig. 1, point B).

The sensitivities on the stable branches differ due to both nonlinearity of black body radiation and change in emissivity. Due to the choice of albedo and emissivity changes in this model,  $\alpha'$  is nonzero only in the bistable regime and  $\epsilon'$  is nonzero only in the temperature range where it varies ( $T_0 - T_\epsilon/2 \leq T \leq T_0 + T_\epsilon/2$ ): see Fig. 2. State-dependence between glacial and interglacial states has been detected in estimates of specific climate sensitivities from different palaeo-data, suggesting lower sensitivity during cold periods than during warm periods (e.g. [15,39] who estimate a close approximation of the Charney sensitivity and find warm (interglacial) climate states to be about 60% more sensitive than cold (glacial) states).

We can compute the curvature of (19) as

$$\frac{d^2}{dT^2} \Delta R_{[CO_2]} = (\epsilon''T^2 + 8\epsilon T + 12\epsilon')T^2 - Q_0\alpha''$$

Note that this is small except near the folds at  $T \approx T_1$  and  $T \approx T_2$  with maximum absolute value  $\frac{d^2}{dT^2} \Delta R$  of order  $T_\alpha^{-1}$ . This confirms that the saddle-node becomes non-smooth in the

<sup>1</sup> This is referred to as *local slope* sensitivity in [24].

**Table 1** Parameters for the energy balance model (15, 17) adapted from [12] to include state-dependent emissivity

$\alpha_1$	0.52	–	$\alpha_2$	0.47	–
$T_1$	278	K	$T_2$	288	K
$\epsilon_1$	0.53	–	$\epsilon_2$	0.39	–
$A$	5.35	$\text{W m}^{-2}$	$T_0$	288	K
$T_\alpha$	5	K	$T_e$	20	K
$C_0$	280	ppm	$Q_0$	342	$\text{W m}^{-2}$
$C_T$	$5 \times 10^8$	$\text{J m}^{-2} \text{K}^{-1}$	$\sigma$	$5.67 \times 10^{-8}$	$\text{W m}^{-2} \text{K}^{-4}$
$\eta_T$	$5 \times 10^{-6}$	$\text{K s}^{-1/2}$	$\eta_C$	$2 \times 10^{-6}$	$\text{s}^{-1/2}$

Note that [12] consider a global transition and so use different values:  $\alpha_1 = 0.7$   $\alpha_2 = 0.2$ ,  $T_1 = 263$  K,  $T_2 = 293$  K,  $T_\alpha = 0.273$  K  $A = 20.5 \text{ W m}^{-2}$  and have an additional constant  $150 \text{ W m}^{-2}$

limit  $T_\alpha \rightarrow 0$ . The second derivative gives the size of the quadratic correction  $a$  in Zaliapin et al. [41]; very large values of the slope near the saddle nodes correspond to the run-away climate observed in [4].

For the model (17), the atmospheric  $\text{CO}_2$  concentration is a parameter for the energy balance dynamics. We explore this by considering a ‘wandering’  $\text{CO}_2$  profile such that  $\gamma(t) := \ln(C(t))$  undertakes a Brownian motion with growth in variance  $\eta_\gamma$  per unit time between reflecting limit values. More precisely, we consider  $\text{CO}_2$  dynamics governed by soft reflecting boundary conditions at  $\ln C_{min}$  and  $\ln C_{max}$ :

$$d\gamma = K\theta(\gamma) dt + \eta_C dW_\gamma \quad (22)$$

where the noise in the  $C$  variable (with fixed amplitude  $\eta_C$ ) represents variability in  $\text{CO}_2$  forcing. We assume

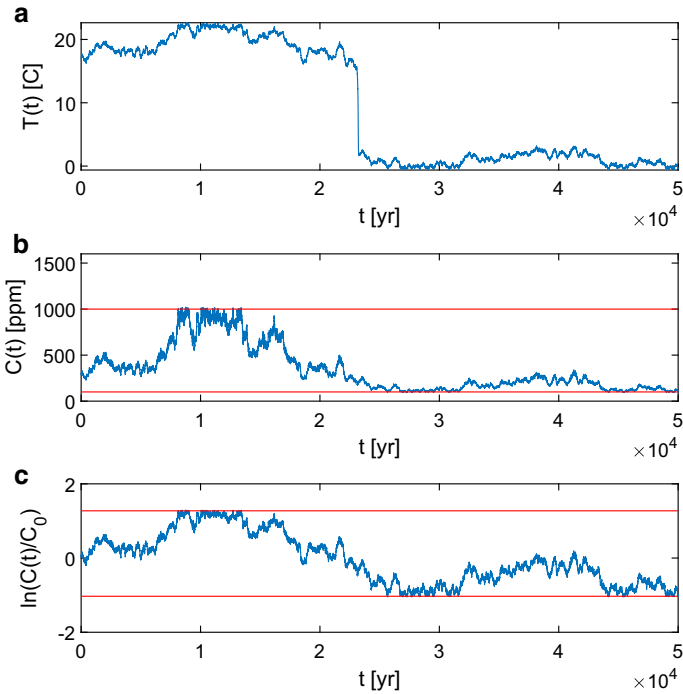
$$\theta(\gamma) := H(\ln C_{min} - \gamma)(\ln C_{min} - \gamma) + H(\gamma - \ln C_{max})(\ln C_{max} - \gamma) \quad (23)$$

and we use parameters

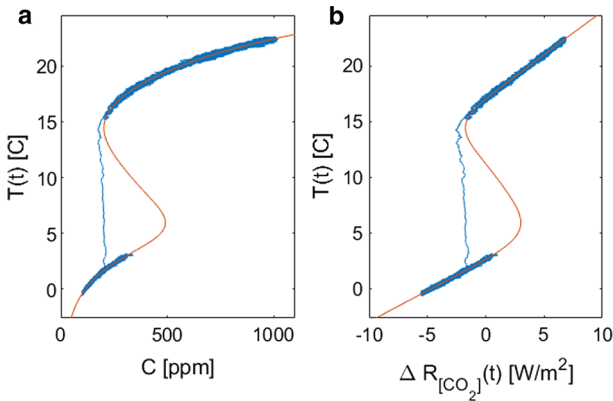
$$K = 10^{-7} \text{ s}^{-1}, \quad C_{min} = 10^2 \text{ ppm}, \quad C_{max} = 10^3 \text{ ppm}, \quad \eta_C = 2 \times 10^{-6} \text{ s}^{-1/2}. \quad (24)$$

Clearly there are common causes of variability of temperature and  $\text{CO}_2$  and so in general there will be strong correlation between the noise terms  $W_T$  and  $W_\gamma$ ; for convenience we assume here that they are uncorrelated. In most studies of climate sensitivity, carbon cycle processes are not treated as feedbacks but rather as forcing (external to the system); this assumption corresponds in our model to  $\text{CO}_2$  driving temperature changes with no direct impact of temperature on  $\text{CO}_2$ . The parameter  $\eta_C$  determines the timescale of wandering of the  $\text{CO}_2$ : we consider cases where this is slower than, or comparable to the timescale of evolution of  $T$ .

Figure 3b shows a time series for a typical simulation of (17) with wandering  $\text{CO}_2$  (22) and parameters as in Table 1, while the corresponding time series of  $T$  is shown in Fig. 3a; we see as  $C$  crosses thresholds (for the bistable regime as calculated in (20)) the state of the system tips between warm and cold states. Global mean temperature  $T$  vs  $C$  and relative radiative forcing  $\Delta R_{[CO_2]}$  are shown in Fig. 2c, d for the non-stochastic model and in Fig 4 for the stochastic model, respectively. Observe the region of bistability around  $\Delta R_{[CO_2]} = 0$  (Fig. 4b) that switches rapidly between cool high albedo and warm low albedo states via saddle-node bifurcations. There is approximate linearity away from these tipping points, but with different mean slopes.



**Fig. 3** Time series for the energy balance model (1) with wandering CO<sub>2</sub> (22). **a** Global mean temperature  $T(t)$  and **b, c** atmospheric CO<sub>2</sub>  $C$  in original **b** and logarithmic coordinates (**c**). The red lines indicate the limits imposed on the randomly wandering CO<sub>2</sub>. See text for details



**Fig. 4** **a** Temperature versus **a** atmospheric CO<sub>2</sub> concentration and **b**  $\Delta R$  of radiative forcing by CO<sub>2</sub>, for the time series in Fig. 3. The red line corresponds to the steady solution of (15) with dependence on CO<sub>2</sub> (Color figure online)

### 3.1 Extreme Sensitivities, Early Warning Signals and Tipping Between Regimes

The energy balance model (17) with wandering CO<sub>2</sub> (22) gives a framework in which one can test correlation between extreme values of sensitivity and tipping between climate regimes,

as well as testing other possible early warning signals for a tipping event. Indeed, we find a rise in instantaneous sensitivity seems to act as a good precursor in cases of slow variation of  $\text{CO}_2$ .

Figure 5(left) shows the variation of instantaneous sensitivity and two early warning signals for tipping between regimes for the wandering variation of  $\text{CO}_2$  concentration  $\ln C(t)$  relative to the timescale of evolution of  $T$ , with  $\eta_C = 2 \times 10^{-6} \text{ s}^{-1/2}$ ; this means that  $\text{CO}_2$  variations are comparable in timescale to  $T$  fluctuations. By considering the nearest equilibrium point on  $C = \Gamma(T)$  for fixed  $C$  (18), we evaluate the instantaneous sensitivity  $S$  using (21) and plot this in middle panel. Observe there is a qualitative change in  $S$  before and after tipping events. There is a clear precursor and then singularity as the tipping point is crossed. Note that in this case the instantaneous sensitivity depends only on the current  $\text{CO}_2$  level and the nearest branch—fluctuations in  $T$  around the branch do not affect  $S$ , while fluctuations in  $C$  do. There are also apparent ‘false alarms’: for example, the fluctuations of  $S$  around 3 kyr and 18 kyr. Figure 5(left, d, e) show detrended estimates of  $sd(t)$  and  $AR1(t)$  [42] using moving averages with length  $\tau = 500$  years. Neither show any clear precursors before tipping events. By contrast, Fig. 5(right) shows early warning signals for tipping between regimes for slower variation of  $\text{CO}_2$  concentration with  $\eta_C = 5 \times 10^{-7} \text{ s}^{-1/2}$ , where  $T$  evolves faster than  $C$ . Unlike the case on the left, this slower switching gives a visible precursor of increasing  $AR1$ . In both cases there are increasing fluctuations of instantaneous sensitivity  $S$ . Note however, that the instantaneous  $S$  we consider here uses the model equations, and hence will be more difficult to access from complex model realisations or observations.

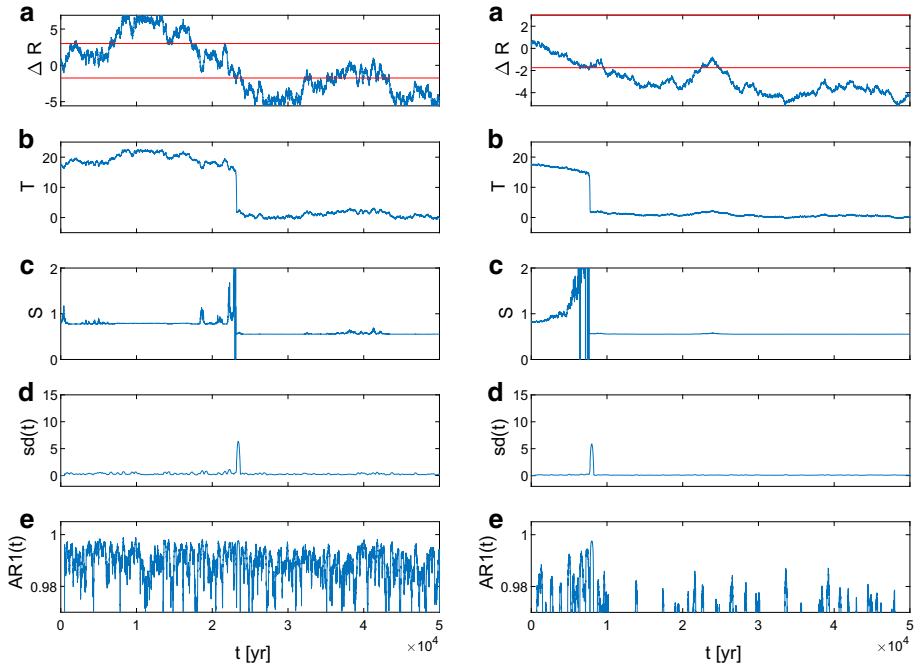
### 3.2 Two-Point Sensitivities and Tipping

An approximation of the stationary density of the global attractor for the system (17, 22) is shown in Fig. 6. We classify the system regime as one of:

$$\begin{cases} \text{Warm (W)} & \text{if } T > T_{\text{thr}} \\ \text{Cold (C)} & \text{if } T \leq T_{\text{thr}} \end{cases}$$

where we choose a threshold  $T_{\text{thr}} = 10 C$  between the two stable branches: see Fig. 1. We simulate a single very long trajectory ( $5 \times 10^5$  years) of the energy balance model (15, 17) with wandering  $\text{CO}_2$  and use this to create a density plot of the climate attractor projected onto  $T$  versus  $\Delta R$ , as shown in Fig. 6. This is used to consider the two-point sensitivities and probabilities of tipping as in Fig. 7. Panels (a–d) are computed by sampling incremental sensitivities (8) from points for increments up to 20 kyr. Panels (e–h) are computed by sampling  $10^7$  pairs of points from the distribution in Fig. 6 using the two-point sensitivity (11). We observe:

- There is good qualitative agreement between the incremental sensitivities averaged over long delays and the two-point sensitivities sampled independently from the attractor. Indeed, the autocorrelation of the timeseries for  $T$  (not shown) has substantially decayed and has its first zeros around 20 kyr.
- High probability of tipping (see Fig. 7b, d) corresponds mostly to extremes of  $S$  that may be positive or negative  $S$ .
- Within the W and C regimes, the sensitivities are closely clustered but have different means for the W and C state. We can estimate these using average temperatures and (21) as  $S \approx 0.79 \text{ K} [\text{W m}^2]^{-1}$  for the W and  $S \approx 0.55$  for the C state, respectively.
- Note that there are relatively low-probability ‘shoulders’ of the distributions within-regime. These are due to the classification of regimes also including states that are in



**Fig. 5** Examples of tipping events for  $T$  for the energy balance model (15, 17) with wandering  $CO_2$  (22) with (left) parameters as in Fig. 3, i.e. with  $(\eta_T, \eta_C) = (5 \times 10^{-6}, 2 \times 10^{-6})$  and (right) slower variation of  $C$   $(\eta_T, \eta_C) = (5 \times 10^{-6}, 5 \times 10^{-7})$ . **a, b** Time series of  $\Delta R$  and  $T$ . The red lines on **a** show the locations of the saddle-node bifurcations that bound the region of bistability (20). **c** shows estimated instantaneous sensitivity from the nearest equilibrium of (17). Note the gradual rise and fluctuations in  $S$  on approach to the tipping point, and the two levels of  $S$  corresponding to the differing sensitivities of the two stable branches. **d, e** Standard deviation and AR1 coefficient: note that the AR1 coefficient seems to have predictive power only for the right column

transitions: although the system is in transition, both points are still classified as the same regime.

- A hysteresis-like effect may appear even in the absence of bistability. If the dynamic change in a system parameter is fast enough then an apparent hysteresis may arise due to different lags for a rising or falling parameter: see for example [19]. However, Fig. 4 suggests this is not the case here.

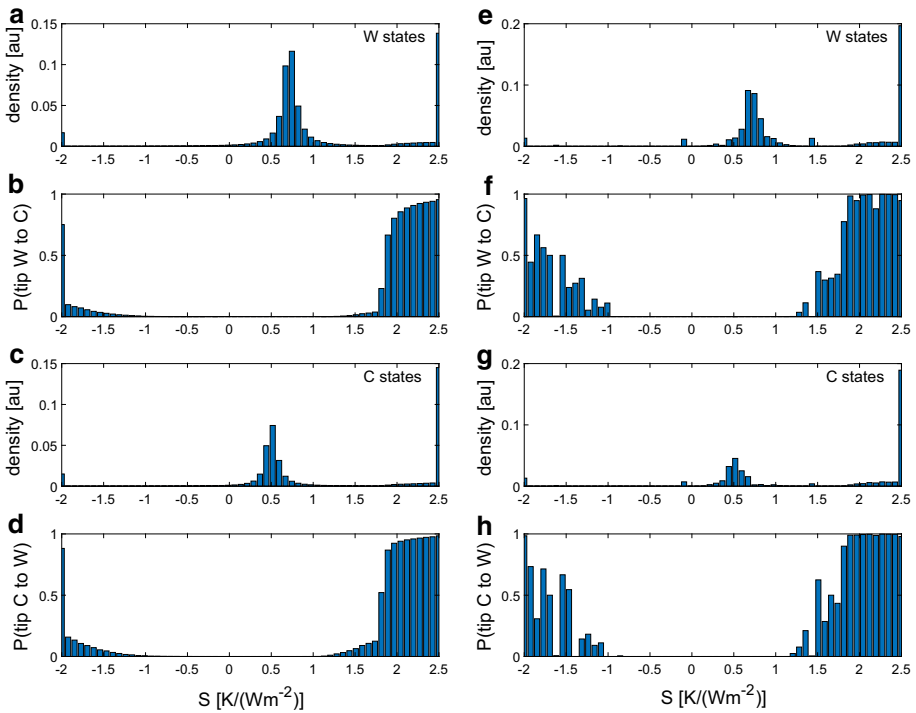
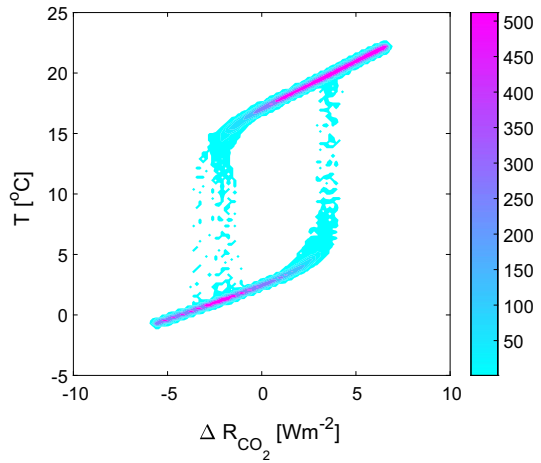
### 3.3 Sensitivities in the Absence of Tipping

Considering the same model (15, 17) with wandering  $CO_2$  (22), we use different parameters to contrast the results in the previous sections with cases where there is no bistability. In particular we consider parameters as in Table 1 except for:

- Default albedo contrast:  $\alpha_1 = 0.52, \alpha_2 = 0.47$  (i.e. also as in Table 1).
- Low albedo contrast:  $\alpha_1 = 0.50, \alpha_2 = 0.48$ .
- No albedo contrast:  $\alpha_1 = \alpha_2 = 0.495$ .

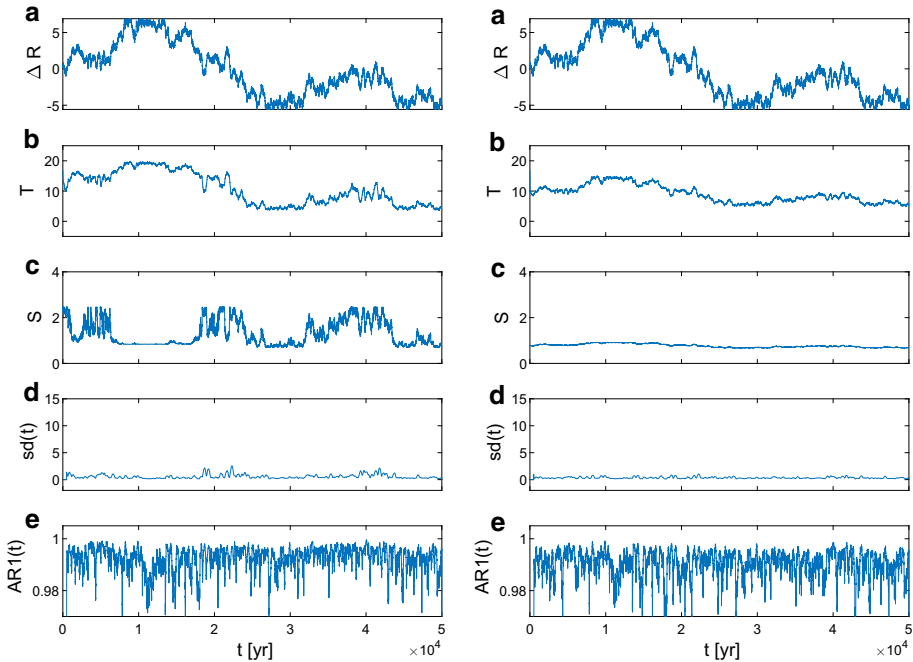
Figure 8 shows the low and no albedo contrast cases, comparing to the default case Fig. 5(left). For the low albedo contrast case there is no longer a region of bistability, but there is nontrivial

**Fig. 6** Density of  $T$  versus  $\Delta R = A \ln(C/C_0)$  of radiative forcing by  $\text{CO}_2$  for a longer time series similar to that in Fig. 4b. Observe the two regions with approximately linear relationship (cf. Fig. 1): the colour scale indicates density [arbitrary units] (Color figure online)



**Fig. 7** **a, c, e, g** Conditional two-point sensitivities and **b, d, f, h** probabilities of tipping from warm (**a, b, e, f**) and cold (**c, d, g, h**) states, for the energy balance model (17) with wandering  $\text{CO}_2$  (22). **a–d** Are computed using a range of delays up to 20 kyr while **e–h** are computed using  $10^7$  independently sampled pairs of points from the distribution in Fig. 6. Note that sensitivities outside the horizontal range are rounded to the last bin

variation of the instantaneous sensitivity along the attractor. For the no albedo contrast case, the instantaneous sensitivity is close to constant. The projection of the climate attractor into the  $(\Delta R, T)$  plane is shown in Fig. 9a–c, while the corresponding distribution of two-



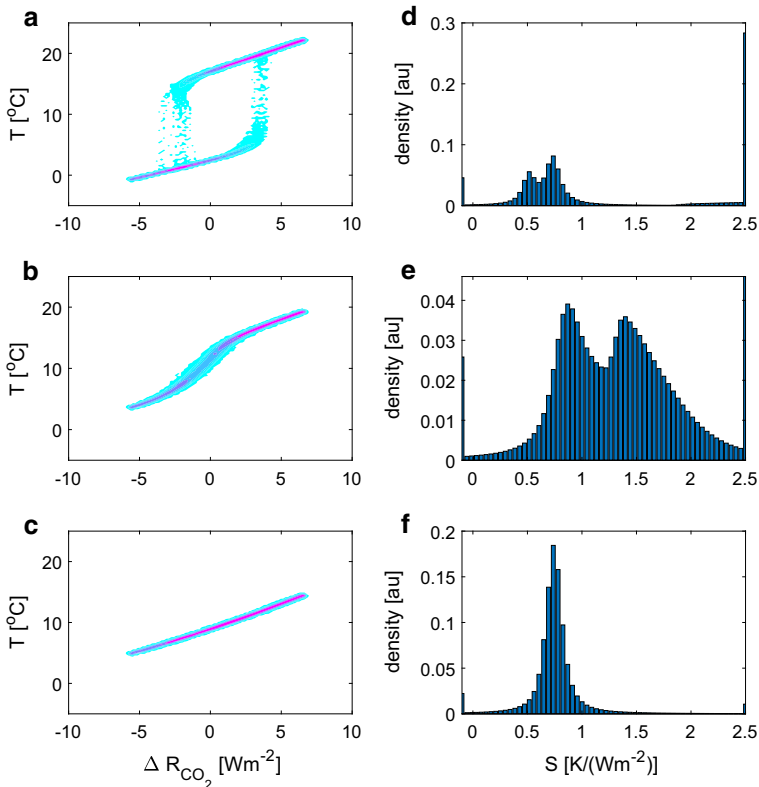
**Fig. 8** Response of global mean temperature to the same realisation of wandering  $\text{CO}_2$  for the energy balance model (17) with (left) low albedo contrast and (right) no albedo contrast; **a–e** as in Fig. 5. In both cases there is no region of bistability. Observe there are large but bounded fluctuations of the instantaneous sensitivity in case (left), indicating state-dependency but no tipping. In case (right) there is comparatively little fluctuation of sensitivity; compare with Fig. 5(left) for the default parameters with a region of bistability

point sensitivities in Fig. 9d–f. Observe the presence of non-unimodal distributions for (d, e) associated with regions with different two-point sensitivity, and clear skewness and tails again associated with the geometry of the measure in (a, b). Note the higher average in (e) corresponds to there being only a single regime in (b) that runs over a wide range of temperatures.

In physical terms, the skewness (and long tails) in (d, e) originate from the state-dependence and nonlinearity of feedbacks (i.e. non-constant feedback factors). The bistability of the two regimes with different feedbacks gives the two peaks in the distribution of Fig. 9d. However, Fig. 9e still has two peaks: these originate from state dependence on the same attractor (Fig. 9b). For this low-albedo-contrast case, there is no ‘tipping’ but we still find very non-Gaussian distribution of  $S$  that comes from nonlinearities in the system that, in this case, do not produce tipping. Note that only in the no albedo contrast case (c) is there a plausible fit to Gaussian.

### 4 Discussion

We demonstrate that state-dependence and the presence of tipping points produces signatures in the distribution of instantaneous and two-point notions of ECS. We explore this using a global energy balance model where state-dependence and multistability originate from the dependence of both albedo and emissivity on temperature.



**Fig. 9** **a–c** Plots of  $T$  versus  $\Delta R$  for a long timeseries of the energy balance model (17) with wandering  $\text{CO}_2$ , parameters as in Fig. 6. **d–f** Plots of distributions of two-point sensitivities corresponding to **a–c** respectively. **a, d** Correspond to the default scenario with bistability, **b, e** to the low albedo contrast and **c, f** to the no albedo contrast scenario; cf. Fig. 8. Observe the presence of two peaks (corresponding to state dependent sensitivities on the two branches) and a large mass of extremes for **(d)**. **e** A range of state-dependent variation but relatively few extremes; while **f** resembles a normal distribution. Observe that **d** can be decomposed into the conditional sensitivities shown in Fig. 7

For the deterministic version of our model (15) with fixed  $\text{CO}_2$  the changes in albedo mean there can be bistability between regimes, while the changes in emissivity contribute to different sensitivities within these regimes. The distribution of ECS comes from several sources—nonlinearities that result in tipping points and/or state-dependence of the feedbacks and sub-grid variability that we model here as stochastic perturbations. Such regime-dependent sensitivity and extremes associated with tipping points are also visible in the more complex Gildor and Tziperman model [17,38], as outlined in Appendix.

For the stochastic model (17) with wandering  $\text{CO}_2$ , regime-dependent sensitivity is visible as differences in slope of the stable regimes for the  $T$  versus  $\Delta R$  plots (see Fig. 4b). The densities of the stable regimes for the  $T$  versus  $\Delta R$  plots (see Fig. 6) show varying slopes and so conditional two-point sensitivities for the two regimes (see e.g. Fig. 7a, c) can have peaks at different sensitivities. We compare several notions of sensitivity. These are the instantaneous sensitivity associated with the slope of the equilibrium branch, incremental sensitivity associated with a fixed delay, and two-point sensitivity that compares arbitrary points on the climate attractor. The presence of tipping points gives extremes of sensitivities



in that (i) there are large fluctuations of instantaneous sensitivities for the nearest equilibrium just before a tipping point (see Fig. 5); and (ii) in the distributions of conditional two-point sensitivities that cross regimes (see Fig. 7d, f). It is remarkable that the two-point sensitivities are so informative, given that they compare points that may be very far from each other in time.

There remains much work to be done to understand the relation between and limitations of these notions of ECS, and indeed ECS calculated in other ways, for example from instantaneous doubling of CO<sub>2</sub>: this will involve transient non-equilibrium processes due to ocean thermal inertia. Note that determining ECS from palaeoclimate time-series [40], the two-point notion clearly has an advantage that we are not limited by the time-resolution of the time series.

The energy balance model can be criticised as being very simple and hard to parametrize in terms of the various physical processes that contribute to albedo and emissivity. Moreover, we consider CO<sub>2</sub> in (17) purely as a forcing term which ignores known land surface and ocean processes where temperature is known to affect CO<sub>2</sub> balance. However, the model is complex enough to confirm that extremes in ensembles of computed climate sensitivities can indicate nearby tipping points. Computations presented in Appendix confirm this picture in a more complex box-model for the glacial cycles, where the CO<sub>2</sub> is modelled dynamically.

#### 4.1 Future Perspectives

When (and how) extremes of sensitivity can be effective precursors of a tipping event will depend on a number of factors. In particular, the timescale of dynamics of the climate response needs to be faster than the timescale of changes in forcing. Figure 5(left) shows that as CO<sub>2</sub> variability is rapid, this results in tipping points with little precursor visible in changes to *AR1*, though it is visible in the instantaneous sensitivity. There may be ‘rate-induced’ tipping points [1,2] that appear when the timescale of the forcing interacts with that of the system. The size of the region of effective nonlinearity can also vary. For (17) this is affected by the temperature scales  $T_\alpha$  and  $T_\epsilon$  over which the albedo and emissivity changes occur. Note also that although tipping points do give rise to extremes in the distribution of ECS, extremes do not necessarily indicate a tipping point.

Translating these results to a time-dependent setting and to more complex models will be difficult: the possible states in the ‘climate attractor’ and the associated invariant measure  $\mu$  is harder to define in the presence of non-stationary temporal variation of forcing, or for large recurrence times and a variety of nonlinear multiscale processes. In such cases, interpreting transitions as tipping points is a challenge; nonetheless, the palaeoclimate record does show a variety of large and sudden transitions [26]. For example, ice core/ocean core records indicate repeated sudden changes in (regional) surface temperature associated with glacial cycles [20] or Dansgaard–Oeschger [13] events as well as global transitions, for example the greenhouse-icehouse transition at the Eocene–Oligocene transition [10]. Although glacial cycles can be found in models such as [17] as relaxation oscillation with clear regimes, for climate reconstruction data these regimes are not so clear (e.g. [15,24]).

Recent work [35] suggests we are at a crossroads in terms of the future earth system state. On the one hand, looking at the palaeoclimate record for the last 1 million years suggests that we are overdue descent into an ice age. On the other hand, comparison of anthropogenic CO<sub>2</sub> emissions with the palaeo record suggest the next tipping point may be to much warmer ‘hothouse’ earth. A better understanding of improved indicators such as two-point ECS and what they say about the climate response to changes in greenhouse gases, together with a

better understanding of hothouse earth climate states that may have existed in the past (e.g. the Palaeocene climate [3,21]) should help our understanding and guide future generations in their need to avoid dangerous climate change.

**Acknowledgements** We thank Peter Cox, Henk Dijkstra, Valerio Lucarini, Michael Ghil and Georg Gottwald for interesting comments and suggestions concerning this work. We thank the EU ITN network CRITICS and CliMathNet for opportunities to discuss and develop this research. PA thanks the EPSRC for support under grant number EP/M008495/1. AvdH thanks the program of the Netherlands Earth System Science Centre (NESSC), financially supported by the Dutch Ministry of Education, Culture and Science (OCW). This project is TILES contribution #7: this project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820970. MATLAB code for the Energy Balance Model is available from <https://github.com/peterashwin/ashwin-heydt-2019>.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

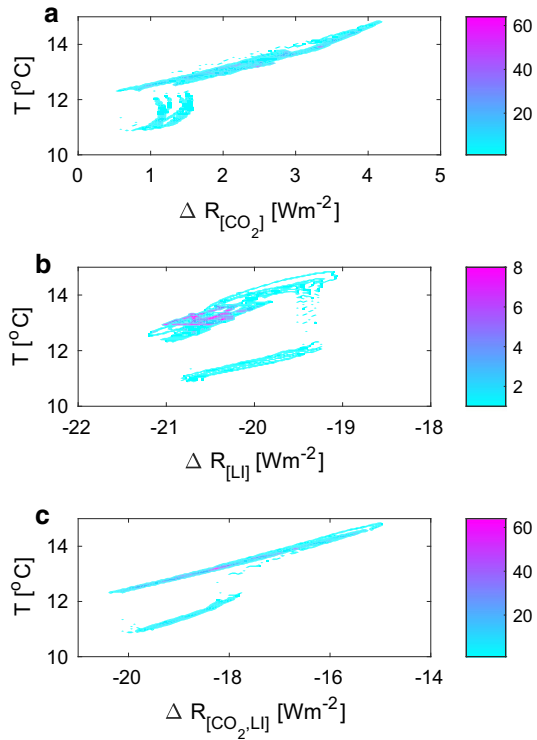
**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Appendix: Tipping in a Multi-box Climate Model with Ocean Biogeochemistry

To investigate the notion of two-point sensitivity in a more complex and physics based earth system model, we explore here the Gildor–Tziperman model [17] with Milankovich forcing and biogeochemistry in the ocean. In this model the ocean and atmosphere consists of 4 latitudinal (zonally averaged) boxes each, with two layers in the ocean and one in the atmosphere. Within the boxes a variety of thermodynamic quantities (e.g. temperature  $T$ ) and species (e.g. ocean salinity  $S$  and atmospheric  $\text{CO}_2$ ) are modelled. In addition there is dynamic land ice (slowly evolving) and sea ice (rapidly evolving) as well as fluxes that join these boxes. The system is sufficiently complex to allow modelling of the Pleistocene ice-age oscillations of land-ice in response to Milankovich forcing. The glacial-interglacial cycles appear in this model as internally generated self-sustained oscillations, which are then modified by the Milankovich forcing. More details are given in [17,38].

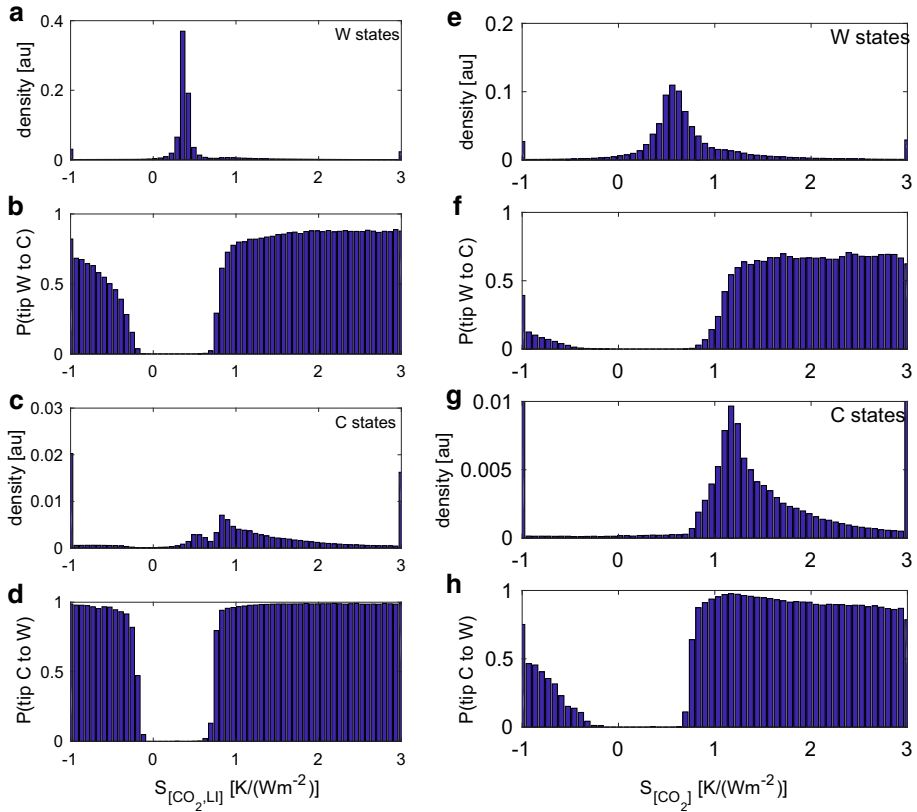
Figure 10 shows projections of a long trajectory (500 kyr) on the climate attractor of this more complex climate model onto the plane of global mean temperature  $T$  against (a)  $\Delta R_{[\text{CO}_2]}$  (b)  $\Delta R_{[LI]}$  and (c)  $\Delta R_{[\text{CO}_2, LI]}$  [38]. Observe that all three projections clearly show two climate regimes, a lower ‘cold’ state (corresponding to large amounts of sea ice and  $T < 12.28\text{C}$ ) and an upper ‘warm’ state (corresponding to  $T > 12.28\text{C}$ ). The projection on the combined radiative forcing of  $\text{CO}_2$  and the slow feedback in land ice changes  $\Delta R_{[LI]}$  shows a clear slope and hence ‘mean’ sensitivities in both regimes (Fig. 10c). Following the formalism of [32], the slopes in (a) should reflect the Earth System sensitivity, while (c) should give a good approximation of the Charney or equilibrium climate sensitivity. Note that in this model there is only one slow feedback to correct for, namely the land-ice albedo feedback. When projecting onto the  $\Delta R_{[\text{CO}_2]}$  plane (Fig. 10a) the cold regime appears very diffuse and with very high (or sometimes negative) earth system sensitivities (ESS), suggesting that in the cold branch local climate dynamics are not entirely determined by  $\text{CO}_2$ . Similarly, when

**Fig. 10** Plots of  $T$  versus  $\Delta R$  for the Gildor–Tziperman model [17]; see [38]. **a** Projection onto the  $(\Delta R_{[CO_2]}, T)$ -plane, which should reflect the earth system sensitivity only considering  $CO_2$  as forcing; **b** projection onto the  $(\Delta R_{[LI]}, T)$ -plane, considering only the slow land-ice albedo feedback as forcing; **c** projection onto the  $(\Delta R_{[CO_2,LI]}, T)$ -plane, considering both  $CO_2$  and the slow feedbacks (in this model there is only the land-ice albedo feedback slow) as forcing. Following the formalism of [32] the slopes in this graph should reflect a good approximation of the Charney ECS



projecting only on land ice changes  $\Delta R_{LI}$  (Fig. 10b) the warm branch appears more diffuse, i.e.  $CO_2$  dynamics seem to be more important than land ice dynamics on this branch.

The left column of Fig. 11a–d shows the distribution of two-point climate sensitivities for  $R_{[CO_2,LI]}$  conditional on regime: this is comparable to Fig. 7a–d in that one can observe (i) clearly localised distributions of ECS in (a, c) conditional on remaining within the W or C regime, (ii) a broader distribution in (c): this seems to be associated with the curvature of the C regime branch in Fig. 10c), (iii) a clear association of tipping from W to C (b) or from C to W (d) being associated with extreme sensitivities. Note that for this model there is no energy balance model available and so it is not possible to compute the instantaneous sensitivity. For comparison we show in the right column of Fig. 11e–h the same distributions for the ESS, which are not compensated for the slow feedback (in this model the land-ice albedo feedback). Observe that for both regimes there is a much broader distribution for the ESS (e, g) than for the ECS in (a) and (c). In particular in the cold regime, earth system sensitivities (g) are much higher than equilibrium sensitivities (c) because the land-ice albedo feedback is very strong in the cold (land-ice covered) states.



**Fig. 11** Conditional sensitivities and probabilities of tipping for glacial cycles simulations with the Gildor–Tziperman model [17] (as in Fig. 7 for the energy balance model). **a–d** Conditional two-point sensitivities compensating for slow feedbacks (i.e. reflecting ECS) and probabilities of tipping (i.e. from the distribution shown in Fig. 10c). **e–h** Conditional earth system sensitivities ESS (not compensated for slow feedbacks) and probabilities of tipping (i.e. from the distribution shown in Fig. 10a)

## References

1. Ashwin, P., Wieczorek, S., Vitolo, R., Cox, P.: Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system. *Philos. Trans. R. Soc. A* **370**(1962), 1166–1184 (2012)
2. Ashwin, P., Perryman, C., Wieczorek, S.: Parameter shifts for nonautonomous systems in low dimension: bifurcation-and rate-induced tipping. *Nonlinearity* **30**(6), 2185 (2017)
3. Baatsen, M.L.J., von der Heydt, A.S., Huber, M., Kliphuis, M.A., Bijl, P.K., Sluijs, A., Dijkstra, H.A.: Equilibrium state and sensitivity of the simulated middle-to-late Eocene climate. *Clim. Past Discuss.* (2018). <https://doi.org/10.5194/cp-2018-43>.
4. Bloch-Johnson, J., Pierrehumbert, R.T., Abbot, D.S.: Feedback temperature dependence determines the risk of high warming. *Geophys. Res. Lett.* **42**, 4973–4980 (2015)
5. Budyko, M.I.: The effect of solar radiation variations on the climate of the earth. *Tellus* **21**(5), 611–619 (1969)
6. Caballero, R., Huber, M.: State-dependent climate sensitivity in past warm climates and its implications for future climate projections. *Proc. Natl. Acad. Sci.* **110**(35), 14162–14167 (2013)
7. Charney, J.: *Carbon Dioxide and Climate: A Scientific Assessment*. The National Academies Press, Washington, DC (1979)
8. Chekroun, M.D., Simonnet, E., Ghil, M.: Stochastic climate dynamics: random attractors and time-dependent invariant measures. *Physica D* **240**(21), 1685–1700 (2011)

9. Cox, P.M., Huntingford, C., Williamson, M.S.: Emergent constraint on equilibrium climate sensitivity from global temperature variability. *Nature* **553**(7688), 319 (2018)
10. Coxall, H.K., Wilson, P.A., Pälike, H., Lear, C.H., Backman, J.: Rapid stepwise onset of antarctic glaciation and deeper calcite compensation in the pacific ocean. *Nature* **433**, 53–57 (2005)
11. Dekker, M.M., von der Heydt, A.S., Dijkstra, H.A.: Cascading transitions in the climate system. *Earth Syst. Dyn.* **9**(4), 1243–1260 (2018)
12. Dijkstra, H.A., Viebahn, J.P.: Sensitivity and resilience of the climate system: a conditional nonlinear optimization approach. *Commun. Nonlinear Sci. Numer. Simul.* **22**(1–3), 13–22 (2015)
13. Ditlevsen, P.D., Kristensen, M.S., Andersen, K.K.: The recurrence time of Dansgaard-Oeschger events and limits on the possible periodic component. *J. Clim.* **18**(14), 2594–2603 (2005)
14. Drijfhout, S., Bathiany, S., Beaulieu, C., Brovkin, V., Claussen, M., Huntingford, C., Scheffer, M., Sgubin, G., Swingedouw, D.: Catalogue of abrupt shifts in intergovernmental panel on climate change climate models. *Proc. Natl. Acad. Sci. USA* **112**(43), E5777–86 (2015)
15. Friedrich, T., Timmermann, A., Tigchelaar, M., Timm, O.E., Ganopolski, A.: Nonlinear climate sensitivity and its implications for future greenhouse warming. *Sci. Adv.* **2**(11), e1501923–e1501923 (2016)
16. Ghil, M.: Climate stability for a sellers-type model. *J. Atmos. Sci.* **33**(1), 3–20 (1976)
17. Gildor, H., Tziperman, E.: A sea ice climate switch mechanism for the 100-kyr glacial cycles. *J. Geophys. Res. Oceans* **106**(C5), 9117–9133 (2001)
18. Gregory, J.M.: A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.* **31**(3), 147–4 (2004)
19. Herein, M., Márfy, J., Drótos, G., Tél, T.: Probabilistic concepts in intermediate-complexity climate models: a snapshot attractor picture. *J. Clim.* **29**(1), 259–272 (2016)
20. Hogg, A.Mc.C.: Glacial cycles and carbon dioxide: a conceptual model. *Geophys. Res. Lett.* **35**(1), L01701 (2008)
21. Huber, M., Caballero, R.: The early Eocene equable climate problem revisited. *Clim. Past* **7**(2), 603–633 (2011)
22. IPCC: Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge (2013)
23. Knutti, R., Hegerl, G.C.: The equilibrium sensitivity of the earth’s temperature to radiation changes. *Nat. Geosci.* **1**, 735–743 (2008)
24. Koehler, P., Stap, L.B., von der Heydt, A.S., de Boer, B., Van De Wal, R.S.W., Bloch-Johnson, J.: A state-dependent quantification of climate sensitivity based on paleodata of the last 2.1 million years. *Paleoceanography* **32**, 1102–1114 (2017)
25. Köhler, P., de Boer, B., von der Heydt, A.S., Stap, L.B., van de Wal, R.S.W.: On the state dependency of the equilibrium climate sensitivity during the last 5 million years. *Clim. Past* **11**(12), 1801–1823 (2015)
26. Lenton, T.M.: Early warning of climate tipping points. *Nat. Clim. Change* **1**(4), 201 (2011)
27. Lenton, T.M., Held, H., Kriegler, E., Hall, J.W., Lucht, W., Rahmstorf, S., Schellnhuber, H.J.: Tipping elements in the Earth’s climate system. *Proc. Natl. Acad. Sci. USA* **105**(6), 1786–1793 (2008)
28. Lucarini, V.: Revising and extending the linear response theory for statistical mechanical systems: evaluating observables as predictors and predictands. *J. Stat. Phys.* **173**, 1698–1721 (2018)
29. Lunt, D.J., Haywood, A.M., Schmidt, G.A., Salzmann, U., Valdes, P.J., Dowsett, H.J.: Earth system sensitivity inferred from Pliocene modelling and data. *Nat. Geosci.* **3**(1), 60–64 (2010)
30. Pfister, P.L., Stocker, T.F.: State-dependence of the climate sensitivity in earth system models of intermediate complexity. *Geophys. Res. Lett.* **44**(20), 10,643–10,653 (2017)
31. Roe, G.H., Baker, M.B.: Why is climate sensitivity so unpredictable? *Science* **318**(5850), 629–632 (2007)
32. Rohling, E.J., Sluijs, A., Dijkstra, H.A., Köhler, P., van de Wal, R.S.W., von der Heydt, A.S., Beerling, D.J., Berger, A., Bijl, P.K., Crucifix, M., et al.: Making sense of palaeoclimate sensitivity. *Nature* **491**(7426), 683 (2012)
33. Schwartz, S.E.: Determination of earth’s transient and equilibrium climate sensitivities from observations over the twentieth century: strong dependence on assumed forcing. *Surv. Geophys.* **33**(3–4), 745–777 (2012)
34. Sellers, W.D.: A global climatic model based on the energy balance of the earth-atmosphere system. *J. Appl. Meteorol.* **8**(3), 392–400 (1969)
35. Steffen, W., Rockström, J., Richardson, K., Lenton, T.M., Folke, C., Liverman, D., Summerhayes, C.P., Barnosky, A.D., Cornell, S.E., Crucifix, M., Donges, J.F., Fetzer, I., Lade, S.J., Scheffer, M., Winkelmann, R., Schellnhuber, H.J.: Trajectories of the earth system in the anthropocene. *Proc. Natl. Acad. Sci.* **115**(33), 8252–8259 (2018)
36. Tantet, A., van der Burgt, F.R., Dijkstra, H.A.: An early warning indicator for atmospheric blocking events using transfer operators. *Chaos* **25**(3), 036406 (2015)

37. Tantet, A., Lucarini, V., Lunkeit, F., Dijkstra, H.A.: Crisis of the chaotic attractor of a climate model: a transfer operator approach. *Nonlinearity* **31**(5), 2221–2251 (2018)
38. von der Heydt, A.S., Ashwin, P.: State dependence of climate sensitivity: attractor constraints and palaeo-climate regimes. *Dyn. Stat. Clim. Syst.* **1**(1), 1–21 (2017)
39. von der Heydt, A.S., Koehler, P., Van De Wal, R.S.W., Dijkstra, H.A.: On the state dependency of fast feedback processes in (paleo) climate sensitivity. *Geophys. Res. Lett.* **41**(18), 6484–6492 (2014)
40. von der Heydt, A.S., Dijkstra, H.A., Van De Wal, R.S.W., Caballero, R., Crucifix, M., Foster, G.L., Huber, M., Koehler, P., Rohling, E.J., Valdes, P.J., Ashwin, P., Bathiany, S., Berends, T., van Bree, L., Ditlevsen, P.D., Ghil, M., Haywood, A.M., Katzav, J., Lohmann, G., Lohmann, J., Lucarini, V., Marzocchi, A., Pälike, H., Baroni, I.R., Simon, D., Sluijs, A., Stap, L.B., Tantet, A., Viehmann, J.P., Ziegler, M.: Lessons on climate sensitivity from past climate changes. *Curr. Clim. Change Rep.* **2**(4), 148–158 (2016)
41. Zaliapin, I., Ghil, M.: Another look at climate sensitivity. *Nonlinear Process. Geophys.* **17**(2), 113–122 (2010)
42. Zhang, X., Kuehn, C., Hallerberg, S.: Predictability of critical transitions. *Phys. Rev. E* **92**, 052905 (2015)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.