Cascading transitions in the climate system

Mark M. Dekker[1,2,3], Anna S. von der Heydt[1,2], and Henk A. Dijkstra[1,2]

[1]Institute for Marine and Atmospheric research Utrecht, Department of Physics, Utrecht University, Utrecht, the Netherlands.
[2]Center for Complex Systems Studies, Utrecht University, Utrecht, the Netherlands.
[3]Department of Computer Science, Utrecht University, Utrecht, the Netherlands.

Correspondence to: Mark Dekker <m.m.dekker@uu.nl>

Abstract. We ~~provide a theory~~ introduce a framework of cascading tipping, i.e., a sequence of abrupt transitions occurring because a transition in one subsystem changes the background conditions for another subsystem. A mathematical framework of elementary deterministic cascading tipping points in autonomous dynamical systems is presented containing the double-fold, fold-Hopf, Hopf-fold and double-Hopf as most generic cases. Statistical indicators which can be used as early warning indicators of cascading tipping events in stochastic, non-stationary systems are suggested. The concept of cascading tipping is illustrated through a conceptual model of the coupled North Atlantic Ocean - El-Niño Southern Oscillation (ENSO) system, demonstrating the possibility of such cascading events in the climate system.

1 Introduction

Earth's climate system consists of several subsystems, e.g., the ocean, atmosphere, ice and land, which are coupled through fluxes of momentum, mass and heat. Each of these subsystems is characterised by specific processes, on very different time scales, determining the evolution of its observables. For example, processes in the atmosphere occur on much smaller time scales than in the ocean and hence in weather prediction, the upper ocean sets the background state for the evolution of the atmosphere. Similarly, in equatorial ocean-atmosphere dynamics associated with the El Niño - Southern Oscillation (ENSO) phenomenon, the global meridional overturning circulation can be considered a background state, as it evolves on a much larger time scale.

This notion that one subsystem provides a background state for the evolution of another subsystem is important when critical transitions are considered. In the climate system, a number of tipping elements have been identified ~~(Lenton et al., 2008)~~(Lenton et al. (2008 an overview of these), where changes in observables can occur relatively rapidly compared to the changes in their forcing near so-called tipping points. Examples of tipping elements are the Atlantic Meridional Overturning Circulation (AMOC) (Stommel, 1961), the Arctic sea ice (Bathiany et al., 2016), monsoon patterns, midlatitude atmospheric flow (Barriopedro et al., 2006), vegetation cover (Hirota et al., 2011) and more local systems like coral reefs and permafrost. When one subsystem undergoes a transition, which changes the background state of another subsystem, also a transition may be induced in that second subsystem. Such dynamical interactions leading to coupled transitions are examples of 'tipping cascades' or 'domino effects' (Kriegler et al., 2009; Lenton and Williams, 2013).

Many tipping points have been analysed in separate subsystems, both for phenomena of the present-day climate (Lenton, 2011; Bathiany et al., 2016), as well as in past climates (such as the abrupt cooling of the Younger Dryas (Livina and Lenton, 2007) and the desertification of the Sahel region (Kutzbach et al., 1996)). However, less attention has been given to the interaction between transitions in different subsystems. For example, when the AMOC collapses, precipitation patterns may change such that the equilibrium structure of the vegetation cover in the Amazon rainforest is shifted (Aleina et al., 2013). This may result in another transition, concerned with forest growth or dieback. Another example is the influence of the AMOC on the trade winds (through meridional sea surface temperature gradients), that in turn influence the amplitude of ENSO. In models, a collapse of the AMOC has been found to intensify ENSO (Lenton and Williams, 2013; Timmermann et al., 2007; Dong and Sutton, 2007), although there are also other effects that would weaken ENSO (Timmermann et al., 2005).

An example in past climates is the coupling between the ocean's overturning circulation and land ice. The rapid glaciation of the Antarctic continent around the Eocene-Oligocene boundary (34 Ma) is often explained in terms of a CO_2 threshold being reached that allows a major ice sheet to grow (DeConto and Pollard, 2003; Gasson et al., 2014). However, a two-step signal is found in the oxygen isotopic ratio, $\delta^{18}O$, which is attributed to a deep-sea temperature drop followed by the (slower) growth of the Antarctic Ice Sheet (AIS). One suggestion to explain the two-step transition is that the deep-sea temperature drop is related to a change in the pattern of the global MOC (Tigchelaar et al., 2011). The ice sheet formation is then argued to be driven by decreasing atmospheric CO_2 (Pearson et al., 2009). ~~This leads to the question whether a cascading tipping event did occur:~~ The switch in MOC (first tipping) has led to the changes in the atmospheric CO_2 (e.g. Elsworth et al. (2017)) which caused the growth of the AIS (second tipping). This leads to the question whether a cascading tipping event occurred.

In the last few years, much work has been done to formulate statistical indicators and early warning signals of tipping points. A system close to a critical transition shows features of a 'critical slowing down' (Dakos et al., 2008; Scheffer et al., 2009; Kuehn, 2011). In the vicinity of the tipping point, the system slowly loses its ability to recover from small perturbations. This results in increased variance, autocorrelation and potentially also increased skewness and flickering (Scheffer et al., 2009). Various methods providing a specific scalar together with a threshold when approaching the transition have been suggested, such as degenerate fingerprinting (Held and Kleinen, 2004; Thompson and Sieber, 2011) and detrended fluctuation analysis (DFA) (Peng et al., 1994; Livina and Lenton, 2007).

When considering cascading tipping points, the autocorrelation of two time series and their interaction needs to be ~~analyzed~~ analysed simultaneously. Podnobik and Stanley (2008) proposed an altered form of DFA to assess the cross correlation between two non-stationary time series and called this method detrended cross-correlation analysis (DCCA). In the computation of the fluctuation function, they used cross-covariance instead of auto-covariance and fit this to a power law. This concept is further extended by defining a coefficient ρ_{DCCA} that accounts for the auto-covariance of the individual time series (Zhou, 2008; Yuan et al., 2015). However, no statistical analysis and indicators have yet been formulated for cascading tipping events.

In this paper, we provide a quantitative approach to cascading tipping events. We start with a mathematical framework to formulate elementary cascading tipping points (section 2). Next, we ~~introduce statistical indicators and potential early warning indicators to analyse~~discuss statistical metrics to analyse, diagnose and potentially predict cascading transitions, and apply them to ensemble simulations of the elementary cascading tipping points (section 3). Finally, we apply the new concepts to an

example within the climate system: the potential cascading tipping mechanism between the AMOC and ENSO (section 4). We ~~summarise~~ summarize and discuss our findings in section 5.

2 Mathematical framework for cascading tipping

In the climate system, tipping points are usually related to rapid transitions, where an observable in the climate system may change abruptly in a relatively short time compared to changes in the forcing of the observable. Such rapid changes often involve transitions from one equilibrium state to another, which can often be explained with classical bifurcation theory for autonomous dynamical systems. ~~These concepts can to~~ To a certain extent, these concepts can also be applied to non-autonomous systems (so-called slow-fast systems) when the time variation of parameters can be viewed as a slow external forcing (Kuehn, 2011). They form ~~also~~ the basics to understand phenomena as noise-induced tipping (Thompson and Sieber, 2011) and rate-dependent tipping (Ashwin et al., 2012).

~~Here we~~ In this section, we present a mathematical framework for simple cascading transitions, that acts as a first step towards analysing the more complex transitions happening in reality. We focus on bifurcation-induced tipping points, and consider two types of bifurcations that are thought to be relevant to mechanisms of abrupt changes in the climate system; the back-to-back saddle-node bifurcation is often used to explain transitions between two co-existing equilibria (multi-stable systems), while the Hopf bifurcation can explain the appearance of oscillatory behaviour (Thompson and Stewart, 2002). In this view, abrupt change in the system appears as a consequence of a parameter crossing a specific critical value at the bifurcation point.

A back-to-back saddle-node bifurcation (two saddle-nodes connected by a common unstable branch) generically occurs in physical systems (having bounded states) when one parameter is varying and the simplest dynamical system having such a bifurcation is described by

$$\frac{dx}{dt} = a_1 x^3 + a_2 x + \phi \tag{1}$$

where the a_i $(i \in 1,2)$ are constants, ϕ is a parameter, x is the state variable and t is time. There are multiple equilibria in the system if and only if $a_1 < 0$, $a_2 > 0$ and within the parameter interval $|\phi| < ((-4a_1^3 a_2^3)/(27a_1^4))^{1/2}$. In this case, the back-to-back saddle-node bifurcations occur at $\phi_c = \pm((-4a_1^3 a_2^3)/(27a_1^4))^{1/2}$. In the sequel, we often use the terms 'fold bifurcation' and 'saddle-node bifurcation', although we in practise only use *back-to-back* saddle-node bifurcations.

A Hopf bifurcation also generically occurs in physical systems and the simplest dynamical system in which it occurs when one parameter is varied is described by

$$\begin{aligned} \frac{dx}{dt} &= a_1 y + a_2(\phi - (x^2 + y^2))x \\ \frac{dy}{dt} &= b_1 x + b_2(\phi - (x^2 + y^2))y \end{aligned} \tag{2}$$

where again the $a_i, b_i, i = 1,2$ are constants, ϕ is the parameter, (x, y) is the state vector and t is time. The state vector satisfying (2) reaches a stable periodic orbit if and only if $a_1 b_1 < 0$ and $\phi > 0$; the transition from steady to periodic occurs at $\phi = 0$.

There are two other bifurcations when one parameter is varied (the transcritical and pitchfork bifurcation) but they are non-generic because special conditions must hold (e.g. symmetry) and so these are not considered here. Using the ~~back-to-back~~

saddle-node and Hopf bifurcations, cascading tipping can be viewed as a combination of two coupled subsystems, where each subsystem undergoes one of these two types of bifurcations. ~~The coupling~~ Combining these bifurcations leads to four types of cascading tipping, discussed in this section.

Coupling two systems introduces a direction of the cascade and we take account of this by defining a *leading* system, which during its transition changes a parameter (that is, the coupling term) in the *following* system. The changing parameter in the following system then can ~~induce the~~ bring the following system closer to a bifurcation point, potentially even resulting in a second transition. In ~~the following, we discuss four types of cascading tippingin terms of combinations of saddle-node and Hopf bifurcations.~~this section, we only look at deterministic cases, which does not allow for noise to play a role in the tipping. Therefore, transitions in the leading system result in a changed coupling term that can lead to transitions in the following system. In bifurcation diagrams versus a forcing, the bifurcation points (for deterministic systems) therefore can overlap. However, the transitions are distinguishable in transients, because the following system always tips after the completion of the first transition. This is why we show the bifurcation diagrams of both systems versus the forcing (Fig. 1) and an example of a transient (Fig. 2) for each type. They will be discussed below.

2.1 Double-fold cascade

The most intuitive system that has the potential to undergo a cascading tipping event is a system where both the leading and the following system have saddle-node bifurcations~~('folds')~~. Analogous to ~~the system of (1)~~Eqn. 1, a dynamical system containing a double fold cascade is then:

$$\begin{cases} \dfrac{dx}{dt} = a_1 x^3 + a_2 x + \phi \\ \dfrac{dy}{dt} = b_1 y^3 + b_2 y + \gamma(x) \end{cases} \tag{3}$$

where x is the state vector of the leading system, y that of the following system, a_i, b_i $(i \in 1,2)$ are constants, and ϕ is a parameter in the leading system. The key is here that the function γ, which serves as a parameter in the following system, depends on the leading system. The most simple coupling between the two systems is represented by $\gamma(x) = \gamma_1 + \gamma_2 x$. Observe that a change in the forcing parameter ϕ can induce a transition in x, which may affect the coupling function γ such that also y undergoes a transition. We would like to emphasize that the forcing ϕ does not directly affect y but only through a change in x (which is effectively only significant when x tips).

Implementing this system with values shown in Table 1 gives insight in the system's equilibrium structure (Fig. 1a) and transient behavior (example in Fig. 2a). When ϕ is changed moving through the bistable regime of the leading system the coupling moves the following system through its own bistable regime (see Table 1). Figure 1a ~~shows~~ and b show the equilibria of the leading ~~system for different values of~~ (versus ϕ~~,~~) and following system (versus γ), respectively, showing the bistable regime in the ~~center~~ centre of the figure, embedded in the ~~back-to-back~~ saddle-node structure. ~~The equilibrium structure of the following system is displayed in Fig. 1e, as a function of ϕ.~~ Varying ϕ alters the state of the leading system, which through the coupling γ affects the state of the following system. This results in the existence of four stable states in the following system in the bistable regime of the leading system; two per state of the leading system~~.~~, as shown in Fig. 1i. The leading

system's state acts as a background condition modulating the position of the following system's equilibria and therefore, in case of transition, may drastically reposition the equilibria of the following system. This is intuitively visible in Fig. 2a, where a time series example of the dynamical system in (3) shows a cascading tipping event (parameters shown in Table 1). When the leading system (black) is forced (by changing ϕ) to move from a bistable to a monostable regime, it transits towards a new equilibrium. During this transition, the following system (red) is affected such that it leaves the regime in which it had four possible equilibria and also transits to a different state.

2.2 Fold-Hopf cascade

The second type of cascading tipping event involves a saddle-node bifurcation in the leading system and a subsequent Hopf bifurcation in the following system. Using analogous notation as in ~~(3)~~Eqn. 3, the simplest system that captures this so-called fold-Hopf cascade is

$$\begin{cases} \dfrac{dx}{dt} = a_1 x^3 + a_2 x + \phi \\ \dfrac{dy}{dt} = b_1 z + b_2(\gamma(x) - (y^2 + z^2))y \\ \dfrac{dz}{dt} = c_1 y + c_2(\gamma(x) - (y^2 + z^2))z \end{cases} \tag{4}$$

where x is again the state vector of the leading system, and (y, z) that of the following system. By slowly varying the parameter ϕ (e.g., linearly as $\phi(t)$) the leading system moves through its bistable regime (see ~~Tab.~~Table 1 for parameter values) and via the coupling $\gamma(x) = \gamma_1 + \gamma_2 x$ forces the following system across the Hopf bifurcation point.

The bifurcation structure of the leading system of (4), using parameters stated in Table 1, is displayed in Fig. 1b. ~~This (as~~ As in Fig. 1a~~) shows the back-to-back~~, this system's bifurcation diagram again shows a saddle-node structure~~,~~. The following system in Fig. 1f, now undergoes a Hopf bifurcation and ~~a subsequent oscillatory regime on part of the upper branch. The actual oscillation occurs in the following system, shown in Fig. 1f. On~~ subsequent oscillatory behavior. In Fig. 1j, it can be seen that increasing ϕ on the lower branch of the leading system ~~,~~ regime will still result in a stationary equilibrium for the following system~~does not oscillate, but on the upper branch, for many values of~~. When increasing ϕ ~~, the following system does. This~~ across ϕ_c, the leading system moves towards another state (seen in Fig. 1b) and the coupling γ increases (across γ_c in Fig. 1f) such that an oscillation occurs in the following system. This mechanism makes it possible for steady and oscillatory states to coexist on the right side of the Hopf bifurcation in Fig. 1~~f~~j. An example of a time series showing a fold-Hopf cascading event is shown in Fig. 2b. A transition in the leading system (black) brings the following system (red/orange) into an unstable equilibrium that eventually leads to an oscillatory state.

2.3 Hopf-fold cascade

A third type of cascading event involves a Hopf bifurcation in the leading system and a subsequent saddle-node bifurcation in the following system. Using a similar notation as in the previous subsection, the simplest system with a Hopf-fold cascade (see

Table 1 for parameter values) is given by

$$\begin{cases} \frac{dx}{dt} = a_1 y + a_2(\phi - (x^2+y^2))x \\ \frac{dy}{dt} = b_1 x + b_2(\phi - (x^2+y^2))y \\ \frac{dz}{dt} = c_1 z^3 + c_2 z + \gamma(x) \end{cases} \tag{5}$$

where (x, y) is the state vector of the leading system, and z that of the following system. Again, we can slowly increase ϕ such that the leading system (x, y) crosses a Hopf bifurcation; via the coupling $\gamma(x) = \gamma_1 + \gamma_2 x$ the following system is then moved through its bistable regime such that a fold is reached in z.

Fig. 1c contains the typical bifurcation structure of the leading system in (5), containing a Hopf bifurcation separating ~~stationary from oscillatory behavior~~a stationary from an oscillatory regime. The following system's equilibrium structure for varying ~~ϕ~~ γ is given by Fig. 1g. In this particular configuration, for any negative ϕ there are multiple stable equilibria~~.~~, seen in Fig. 1k. This makes sense, as ϕ only affects the following system via its impact on the leading system, and for negative ϕ the leading system remains constant. At $\phi = 0$, the Hopf bifurcation in the leading system is reached and (x, y) start oscillating. The following system oscillates a little along with the leading system due to the oscillatory changing value of γ.

When ϕ increases more, the amplitude of the leading system's oscillation grows, which ~~may make~~ eventually makes γ ~~to~~ cross the threshold such that the following system leaves its bistable regime (be it temporarily as γ will be reduced again due to the oscillation). This forces the following system into its upper branch, as can be seen in Fig. 1g~~by the red dashed lines in the lower branches ending at $\phi \approx 0.5$~~. The upper branch's ~~displayed~~ stable oscillation ends at $\phi \approx 0.8$ (in Fig. 1k), because the amplitude becomes large enough for the system to swap between multiple equilibria. An example of such a cascading transition event can be seen in Fig. 2c, where an oscillation starts in the leading system (black/grey), of which a particular phase makes the following system (red) transit into the second equilibrium.

2.4 Double-Hopf cascade

A fourth type of cascading tipping event discussed here involves a Hopf bifurcation in the leading system and a subsequent Hopf bifurcation in the following system. Using analogous notation as above, this double Hopf cascade is captured by the dynamical system

$$\begin{cases} \frac{dx}{dt} = a_1 y + a_2(\phi - (x^2+y^2))x \\ \frac{dy}{dt} = b_1 x + b_2(\phi - (x^2+y^2))y \\ \frac{du}{dt} = c_1 v + c_2(\gamma(x) - (u^2+v^2))u \\ \frac{dv}{dt} = d_1 u + d_2(\gamma(x) - (u^2+v^2))v \end{cases} \tag{6}$$

where (x, y) is the state vector of the leading system, and (u, v) that of the following system. If ϕ forces (x, y) such that it crosses the Hopf bifurcation point, the coupling $\gamma(x) = \gamma_1 + \gamma_2 x$ causes a crossing of the second Hopf bifurcation in (u, v).

Figure 1d ~~shows the bifurcation diagram~~ and h show the bifurcation diagrams of the leading ~~system, showing a typical system with a Hopf bifurcation~~and following systems, with supercritical Hopf bifurcations. The following system (in Fig. 1~~h~~l) is stationary for many values of ϕ, up to the point that the ~~term γ becomes high enough to start an oscillation~~leading system starts oscillating, which for high enough values of ϕ is large enough to make the following system cross the Hopf bifurcation. However, γ oscillates with the leading system (for $\phi > 0$). This means that only in a particular part of the leading system's oscillation period, oscillatory ~~behavior~~ behaviour can be expected in the following system. This interaction between the two oscillations result in torus bifurcations for particular values of ϕ. An example of a time series showing a Hopf-Hopf cascading transition is displayed in Fig. 2d. After a (slow) oscillation in the leading system (black/grey) has started, a (fast) oscillation in the following system (red/orange) arises in particular phases of the slow oscillation.

3 ~~Early warning signals~~ Statistics of cascading tipping points

In the previous section we have formulated elementary deterministic dynamical systems that can exhibit cascading tipping. In order to detect tipping events from~~e.g.~~, for example, observed time series in real systems, we need to detect whether a system is close to a critical transition. In general, ~~a system close to critical transition~~ such a system recovers more slowly from perturbations, which in turn increases memory in the time series. This leads to the phenomenon of 'critical slowing down' prior to bifurcation points. In this section, we ~~present statistical indicators which may be useful to detect~~ simulate cascading tipping events to (a) study the statistical character of such events, (b) diagnose (post-tipping) whether tipping events are causally related and (c) take a first step towards statistical indicators for the prediction of cascading tipping events.

3.1 Methods for single tipping points

Several methods have been suggested for the analysis of time series to detect the approach of a single tipping point. For saddle-node bifurcations, the key features of such a time series is a critical slowing down. This can be investigated as standard quantities such as increasing autocorrelation (e.g., the lag-1 autocorrelation), increasing variance and increasing skewness (Held and Kleinen, 2004; Scheffer et al., 2009; Kuehn, 2011). Although critical slowing down near critical transitions is a ~~more~~ general feature of (even chaotic) dynamical systems (Tantet et al., 2018), the standard ~~quantities may~~ metrics like autocorrelation at lag 1 and variance do not always provide an early warning ~~of a critical transition.~~signal (e.g. in Greenland ice core data in Livina and Lenton (2007)). Hence, more complicated indicators have been introduced, such as (i) the degenerate fingerprinting (DF) and (ii) the detrended fluctuation analysis (DFA) ~~,~~(Held and Kleinen, 2004; Thompson and Sieber, 2011; Peng et al., 1994; Livina and Lenton, 2007). DFA is argued to be a solution to the problem of the simple lag-1 autocorrelation that it does not capture the approach to a transition in highly non-stationary data in long time series (Peng et al., 1994; Livina and Lenton, 2007). The latter is characterized by high autocorrelation due to the gradual increase or decrease of the system itself, distorting the signal of the critical slowing down, a problem in standard metrics and DF.

As critical slowing down implies an increasing autoregressive ~~behavior~~ behaviour in the time series prior to a transition, the memory component is increased. In general, a first step in DF is the projection of the data fields onto the leading EOF,

which gives a time series x_n (Held and Kleinen, 2004). After time-equidistant interpolation and detrending of the data, in the DF method ~~,~~ one fits the following general autoregressive process to the time series x_n:

$$x_{n+1} = c \cdot x_n + \sigma \eta_n \tag{7}$$

where η_n is Gaussian white noise and $c = \exp(-\lambda \Delta t)$ the AR(1) coefficient. Here λ can be seen as the decay rate of perturbations in previous time steps. As the approaching of a bifurcation point involves an increase in memory, the value of c is presumed to increase towards one when approaching a saddle-node bifurcation point.

~~DFA copes well with non-stationarity in time series while searching for long-range correlations (Peng et al., 1994; Livina and Lenton, 20~~

In DFA, one first chooses an integer window size s and divides the (cumulative-summed) time series $X(n)$ in $N_s = N/s$ segments that do not overlap, where N is the length of the time series. In every window, the best polynomial fit of a chosen order is calculated. A quadratic polynomial is used here. The squared deviation from this quadratic polynomial for every window is summed, resulting in a measure of the auto-covariance fluctuating around the fit:

$$F^2(\nu, s) \quad = \quad \frac{1}{s} \sum_{i=1}^{s} [X((\nu - 1)s + i) - x_\nu(i)]^2 \tag{8}$$

with X the detrended time series and x_ν the best polynomial fit in segment ν. Then, an average is taken over all segments to obtain the fluctuation function $F(s)$:

$$F(s) \quad = \quad \sqrt{\frac{1}{N/s} \sum_{\nu=1}^{N/s} F^2(\nu, s)} \tag{9}$$

which depends solely on s. The long-range auto-correlations can now be recognized by fitting the fluctuation function to a power-law and looking at the resulting DFA-exponent α, according to

$$F(s) \quad \propto \quad s^\alpha \tag{10}$$

For $\alpha \leq 0.5$, there is no long-term correlation and the fluctuations are indistinguishable from white noise. However, when $\alpha > 0.5$, there are long-term correlations present and for $\alpha \geq 1.5$ the system has reached a bifurcation point. In the simulations analysed here the DFA scaling exponent is fitted explicitly for every (moving) window.

3.2 ~~Detrended cross-correlation analysis~~Methods for cascading tipping points

Cascading tipping involves two systems with their own bifurcation structure and their proximity towards bifurcation points. Although the leading system may be close to tipping, the following system might still be far away from its bifurcation point and needs the critical transition of the leading system to even come close to this point, which makes te behaviour of the following system more prone to noise, less dependent on the leading system and less auto-correlated prior to the first tipping. This is why the general measures for single tipping events cannot be used, nor can regular cross (Pearson) correlation. The reason is that the following and leading system do not have a one-to-one relationship ~~,~~ (that is, weakly Pearson correlated), but are rather coupled through specific parameters, ~~only seen~~ which is only visible in long-range correlations.

When approaching a cascading tipping point, the long-range cross-correlation between the ~~two state vectors (of the leading , say~~ the leading state x ~~, and following system, say~~ and the following state y ~~)~~ is expected to increase. The state ~~vector~~ x becomes more auto-correlated and is less susceptible to noise, and therefore through the coupling influences y in a more robust way. To find long-range *cross*-correlations, a method so-called detrended cross-correlation analysis (DCCA) was developed (Zebende, 2011; Podnobik and Stanley, 2008; Zhang et al., 2001; Zhou, 2008). Instead of taking the auto-covariance (8) to calculate the fluctuation function, one takes the cross-covariance,

$$F^2_{DCCA}(\nu, s) = \frac{1}{s}\sum_{i=1}^{s}[(X((\nu-1)s+i) - x_\nu(i)) \cdot (Y((\nu-1)s+i) - y_\nu(i))]^2 \tag{11}$$

with symbols similar to(8). With this function, one can calculate the fluctuation function and subsequent power-law scaling coefficient (Podnobik and Stanley, 2008; Zhang et al., 2001), similar to (9).

A variation on this method was proposed by Zebende (2011) and involves the ratio between F^2_{DCCA} and F_{DFA} of the two systems. Specifically, one chooses a certain segment size s and computes:

$$\rho_{DCCA} = \frac{F^2_{DCCA}}{F_{DFA\{x\}}F_{DFA\{y\}}} \tag{12}$$

which measures the level of the long-term cross-correlation between variable x and y; the quantity ρ_{DCCA} has values between -1 and 1.

There are, a priori, several limitations of using the power-law scaling coefficient and ρ_{DCCA}. First of all, the results are sensitive to choices in the maximum segment size s and the window size, induced by the noise in our simulations. Making the window size too large decreases the possibility to see any development prior to the tipping points, as windows including the tipping event itself are biased by strong autocorrelation and the (tipping-) trend in the data. However, making the window size too small increases the uncertainty in the exponential fit. Similarly, adding small segments co-determining the exponential fit makes the method prone to noise, while larger segments are limited by the window size and computation time. In our simulations, window sizes of 120 data points, and segments sizes between 10 and 60 within those windows were used to calculate the fluctuations per segment size and to calculate $F(s)$ (Eqn. 9). More research is needed to find optimal values of the window and segment sizes, to potentially reduce this limitation in the analysis of cascading transitions. Another limitation of using the power-law scaling coefficient and ρ_{DCCA} is that the way the two systems are coupled (e.g., linearity, with or without temporal lag) affects how an evolution in the leading system affects its cross-correlation with the following system in both magnitude and time. Finally, to observe trends in these metrics, the signal in the cross-correlation between the two systems has to overcome the (partly mutually independent) noise. However, close to critical transition, the recovery from noise actually decreases, making the often subtle *change* in detrended cross-correlation harder to distinguish from noise. These limitations may make the detrended cross-correlation metrics less useful in applications, but trends in the detrended cross-correlation metrics might still act as early warnings for cascading transitions, as shown in the next section.

3.3 ~~Analysis of cascading tipping systems~~Simulations

~~The previous section presented various quantities to analyse the occurrence of cascading tipping~~ In this section we discuss the earlier described metrics applied in ensemble simulations of cascading transition events. This ~~section is devoted to give insight into the accuracy and the usefulness of the indicators, by applying them to the~~ provides insight in the statistical characteristics of these events, the causal relation between tipping of both systems, and the potential prediction of these events. We focus on the double-fold and ~~fold-hopf~~ the fold-Hopf cascading tipping cases ~~.~~ for multiple reasons. First, these cases are most illustrative in terms of relation to physical systems. Second, in these cases the leading system starts with a fold bifurcation, which creates a clear threshold for the start of the event (convenient for analysis purposes).

3.3.1 Double-fold cascading tipping

To simulate these events and use statistical indicators, noise has to be included. The system of equations used here is:

$$\begin{cases} \dfrac{dx}{dt} = a_1 x^3 + a_2 x + \phi + \zeta_x \\ \dfrac{dy}{dt} = b_1 y^3 + b_2 y + \gamma(x) + \zeta_y \end{cases} \tag{13}$$

where now in addition to (3), ζ_x, ζ_y are Gaussian white noise terms. We simulate an ensemble of 100 members with the parameter settings and initial conditions as displayed in Table 2. The results of this ensemble are displayed in Fig. 3. Running windows containing the transition itself are shaded white because this data is misleading when one wants to know what happens before the bifurcation points. We make the distinction between the leading-transitional period (LTP), which is the time series before the tipping point in the leading system, and the following-transitional period (FTP), which is the time series between the first tipping point and the tipping point in the following system. The FTP can be seen as an in-between state that is stable for both systems, although the following system is close to bifurcation, meaning it is rather prone to noise. The duration of this state is therefore highly unpredictable. However, as we will see in this section, the FTP does provide information in the diagnosis of a potential second transition.

In the LTP, we can clearly see the gradual increasing leading system's variance, AR(1) coefficient and DFA scaling coefficient. These are all evidence of the leading system slowly approaching a bifurcation point, according to the change in the parameter. There is not much evidence of long-range auto-correlations in the time series of the following system, as its variance is low and the DFA scaling exponent remains below 0.5, pointing towards that the detrended fluctuations are statistically white noise. The AR(1) coefficient of the following system does increase just prior to the first tipping, but also stays small (compared to unity).

The detrended cross correlation scaling exponent (abbreviated here as DXA) does give > 0.5 values, but the range throughout the ensemble members is most of the time too large to see any structural development when approaching the bifurcation point. During the leading system's transition, a strong increase is visible, pointing towards the rather strong cross-correlated ~~behavior~~ behaviour during this period (as the following system also shifts its equilibrium a little).

The quantity ρ_{DCCA} seems to attain a small positive value (around 0.3) and stays relatively constant throughout the whole time interval. One important aspect of the calculation of ρ_{DCCA}, as we found by experimentation, is that the values is very sensitive to the segment size s and the moving window size. The moving window determines the amount of data that is

available to find long-range correlations, and the segment size has a strong impact on the accuracy of the fits and therefore on the segmented fluctuations. As in this type of problem, we need a temporal evolution of the statistical indicators, we need moving windows and thus encounter this problem. As these indicators (DXA and ρ_{DCCA}) have been applied successfully in simpler systems (Zebende, 2011; Podnobik and Stanley, 2008; Zhang et al., 2001; Zhou, 2008), more research on the sensitivity of the indicators with respect to the segment size and moving window size may lead to more robust results.

During the FTP, the variance, AR(1) and DFA of the leading system are strongly reduced. However, the gradual increasing of the following system's variance, AR(1) coefficient and DFA scaling coefficient are definitely visible, pointing towards the approaching of a bifurcation in the following system. Also notable is the contrast in the DFA of the following system between before and after the tipping of x. The DFA of y went from a white-noise regime (around 0.5) before the tipping of x towards a regime where the detrended fluctuations point towards long-range auto-correlations after the tipping of x (1-1.5). This illustrates the relation between the leading system's state and the following system's DFA scaling exponent. The DXA remains relatively high, but overall no structural development can be seen in this graph. The quantity ρ_{DCCA} exhibits the same ~~behavior~~ behaviour as in the LTP, probably for the reasons already mentioned.

To assess the effects of the cascading effect on the mentioned statistics, we compare the results with a case where the system (13) does not undergo a tipping in the following system (so only one transition remains). The resulting ensemble results are shown in Fig. 4. The most important differences between the regular cascading event and a single tipping event can be found when comparing the variance, AR(1) and DFA scaling coefficient changes between LTP and FTP (or period after the first transition). During the LTP, the leading system is close to transition and therefore has strong autoregressive ~~behavior~~behaviour, which is the opposite for the following system, being far from its bifurcation point. During the FTP, the following system generally gains memory because it is brought closer to its transition point, and the leading system loses this because it had just arrived at a new state. So we expect that from the LTP towards the FTP, the variance, AR(1) and DFA *decrease* in the leading system, and *increase* in the following system.

To quantify this effect, Table 3 shows the ratios of the different quantities, indicated by Q, during the FTP and LTP phases ($\bar{Q}_{FTP}/\bar{Q}_{LTP}$), for the cases *with* a second tipping (corresponding to runs shown in Fig. 3) and *without* second tipping (Fig. 4). All ensemble members are included in these numbers, accounting for a mean and standard deviation of these ratios. As expected, the leading system's autoregressive metrics decrease in both cases, visible in the mean values of the ratios of the leading system's autoregressive variables being lower than 1. Also as expected, the following system's autoregressive ~~behavior~~ behaviour increases (ratios > 1) in both cases, but striking is that in the case of a cascading tipping event (*with* second tipping), the following system's ratios are much higher than those in the case of a single tipping event (*without* second tipping). To investigate whether the difference in ratios between single or double tipping is indeed significant, a Student's t-test is done. The results are shown in ~~Tab.~~ Table 4. The high p-values for the leading system's ratios indicate no significant difference between single or double tipping, but the low p-values for the following system's ratios indicate a significant difference. This shows the potential of using the ratio of autoregressive variables before and after a transition to assess whether a cascading transition may follow. Further research is needed to quantify this expectation and to assess the sensitivity of these ratios to the system's parameters.

3.3.2 Fold-Hopf cascading tipping

Many statistical indicators have been applied on fold bifurcations specifically, because these transitions show a clear sign of critical slowing down and increased autocorrelation due to the irreversibility and process of going from one equilibrium towards another. A ~~supercritical Hopf bifurcation has a different nature with respect to the slowing down, as it is no critical transition~~state approaching a Hopf bifurcation reacts differently to perturbations, than a state approaching a saddle-node bifurcation. We will ~~now~~ therefore consider the fold-Hopf cascade in the light of the statistical indicators described before. For this, we use the following stochastic dynamical system:

$$\begin{cases} \dfrac{dx}{dt} = a_1 x^3 + a_2 x + \phi + \zeta_x \\ \dfrac{dy}{dt} = b_1 z + b_2(\kappa(x) - (y^2 + z^2))y + \zeta_y \\ \dfrac{dz}{dt} = c_1 y + c_2(\kappa(x) - (y^2 + z^2))z + \zeta_z \end{cases} \tag{14}$$

similar to (4) but now white noise is added through the terms ζ_x, ζ_y and ζ_z. We used an ensemble of 100 simulations with the parameter settings and initial conditions as displayed in Table 2. The results of the ensemble are shown in Fig. 5. Here, we do not make the distinction between the LTP and the FTP, because in contrast to the double-fold cascade, the following system undergoes a transition that is easily reversed and the system either is stationary or oscillating. ~~Noise directly starts the oscillation and completely removes the FTP~~We subtract a running average from the states and calculate the statistics from those series, to prevent the oscillation from dominating the signal in the autocorrelation. The following system (red) is quickly drawn towards the equilibrium state ~~$(x, y) = (0, 0)$~~$(y, z) = (0, 0)$, and the leading system (black) is in a steady state. During the time towards the bifurcation point, the variance, AR(1) coefficient and DFA of the leading system ~~z~~ x gradually increase, as is expected as we force this system towards its bifurcation point. It also seems that the AR(1) of the following system slightly increases during this period.

~~The DFA~~ After the transition of the leading system, the oscillation of the following system immediately starts due to noise. The variance and AR(1) of the following system after the bifurcation are ~~in strong contrast with~~ strongly increased with respect to before the bifurcation, ~~probably due to the autoregressive nature of the oscillation. The relation between the leading system's state and the following system's~~ despite the removal of the running average to lose the oscillation's signal itself. On average, the DFA scaling exponent ~~is also confirmed in this case~~also increases of the following system, which relates it to the leading system's state. The DXA sharply increased just prior to the critical transition but throughout the whole time series, retains relatively high values. The reason behind this might be found in the low level of noise that is taken, or other simulation-specific parameters. It could also be that it is because the following system on average has a high, weakly varying DFA scaling exponent~~on itself~~, which in turn might affect the height and variability in the cross-correlation. The ρ_{DCCA} coefficient remains positive and small, just like in the double-fold case. Again, this may have to do with the choice of window and segment sizes.

4 ~~The~~ Application: the coupled AMOC-ENSO system

In this section, ~~the theory~~ we find an application of the concept of cascading tipping ~~will be applied to investigate the coupling between~~ (the fold-Hopf case). This application reflects that cascading transitions are not only a purely mathematical concept, but do occur in idealized physical models. Here, we consider cascading tipping in a model that couples the Atlantic Meridional Overturning Circulation (AMOC) and the El-Niño -Southern Oscillation (ENSO).

4.1 ~~Coupling between AMOC and ENSO~~Background

~~First, to demonstrate~~ To demonstrate and quantify the coupling between AMOC and ENSO, we use output from global climate model simulations. In the ESSENCE project (Ensemble SimulationS of Extreme weather events under Nonlinear Climate changE) several simulations were performed with the ECHAM5/MPI-OM coupled climate model, including so-called hosing experiments (Sterl et al., 2008), where fresh water is added around Greenland to mimic ice sheet melting.

Of these climate model simulations we have used two ensembles; the first is the 'standard' experiment, where greenhouse gases evolve according to observations and from the year 2000 onwards following the SRES-A1b scenario (experiment name SRES-A1b). The second ensemble is the same as the standard one but has additional freshwater input (1 Sv = 10^6 m^3/s) around Greenland from the end of year 2000 onwards (experiment name HOSING-1). The HOSING-1 ensemble contains a hosing experiment in the classical way, following the procedure of Jungclaus et al. (2006). Five runs of each ensemble are taken, specifically runs 041-045 of the HOSING-1 and runs 021-025 of the SRES-A1b ensemble. The temporal resolution used is monthly data between 1950 and 2100. The spatial fields are on a curvilinear grid, with 40 vertical levels in the ocean. We use deseasonalised data because we are interested in interannual variability, not in seasonal variability, as El-Niño is associated with these timescales. As an AMOC index, we use the maximum of the Atlantic meridional overturning stream function at 35°N and as ENSO index, we use the NINO3.4 index, which is the average SST over the region 170°~~W-120~~W - 120°W ~~×~~ by 5°S - 5°N.

The results for the evolution of the AMOC and ENSO are shown in Fig. 6. It is clearly visible that the AMOC decreases strongly in the hosing experiments, by approximately 85%. Table 5 compares statistical properties for the time interval before and after 2001 (which is the year at which the hosing starts). We use the non-anomaly statistics, as this gives us information about the differences in the mean. We do note that we only use five runs per ensemble, which makes the uncertainty not statistically robust. We only state it in Table 5 to give an idea of the range of the variables among the different runs.

It is visible in Table 5 that the variability of NINO3.4 increases (bold numbers) if we compare the periods of 1950-2000 and 2001-2100. This increased variability is visible in both the standard and the HOSING-1 runs. However, the variability is increased much stronger in the HOSING-1 experiment, indicating that the decrease of the AMOC indeed has an amplifying effect on ENSO. The large difference between the standard and hosing runs suggests that the NINO3.4 index changed in the hosing experiment, potentially as a consequence of the decrease of the AMOC. As the first tipping is artificially induced (without any measurable critical slowing down), and the fact that the models used here are much more complex than the simple

dynamical systems we discussed in previous sections, the question of whether cascading tipping is actually happening in these runs is beyond the scope of this paper. We use this data only to justify the coupling of AMOC and ENSO.

Several mechanisms have been suggested in the literature on the coupling between the AMOC and ENSO. The first mechanism is concerned with oceanic waves. A colder North Atlantic creates density anomalies that induces oceanic Kelvin waves to propagate southward (along the American coast) across the equator. In West Africa, this energy radiates as Rossby waves towards the north and south, which induces Kelvin waves to move along the tip of south Africa into the Indian ocean, that eventually reach the Pacific. Consequently, the eastern equatorial Pacific thermocline deepens on a timescale of decades. This deepening has a weakening effect on the amplitude of ENSO (Timmermann et al., 2005).

~~The~~ A second mechanism is concerned with the trade winds. Cooling in the northern tropical Atlantic (due to AMOC weakening) induces anti-cyclonic atmospheric circulation (Xie et al., 2007) that intensifies the northerly trade winds over the northeastern tropical Pacific. This leads to a southward displacement of the Pacific ITCZ (Zhang and Delworth, 2005) and this generates a meridional SST anomaly due to anomalous heat transport and the wind-evaporation SST feedback in the Pacific. Also, Dong and Sutton (2007) found an atmospheric coupling through Rossby waves sent into the northeast tropical Pacific. This is in line with Dijkstra and Neelin (1995), who argue that part of the contribution to the zonal wind stress, τ_{ext}, arises from processes outside the tropical Pacific. The result of the wind stress as coupling between the two systems is an intensification of ENSO and this mechanism is argued to be stronger than the coupling through oceanic waves (Timmermann et al., 2005).

4.2 ~~A coupled AMOC - ENSO model~~Models and coupling

To study the possible cascading transition through the wind-stress coupling, we use a conceptual model. For the AMOC, the classical Stommel box model (Stommel, 1961) is used. It consists of a polar (subscript p) and an equatorial box (subscript e), both with a temperature T and salinity S and coupled by a density-driven flow rate. The state variables are then defined as $\Delta T = T_e - T_p$ and $\Delta S = S_e - S_p$. The time evolution of these variables is as follows (Cessi, 1994):

$$\begin{cases} \dfrac{d\Delta T}{dt} = & -\dfrac{1}{t_r}(\Delta T - \theta_0) - Q(\Delta\rho)\Delta T \\ \dfrac{d\Delta S}{dt} = & \dfrac{F_s}{H}S_0 - Q(\Delta\rho)\Delta S \end{cases} \tag{15}$$

where the first term in the temperature equation refers to relaxation towards a background temperature, and the second term refers to density-driven meridional transport. Specifically, t_r is the surface temperature restoring time scale and θ_0 is the equator-to-pole atmospheric temperature difference. $Q(\Delta\rho)$ is the transport function, which is calculated from a diffusion time scale and the meridional density gradient $\Delta\rho$. In the salinity equation, S_0 is a reference salinity, and H is the ocean depth. The parameter F_s is the freshwater flux, which can be used as a bifurcation parameter. The streamfunction

$$\Psi = \gamma_0 \Delta\rho/\rho_0 = \gamma_0(\alpha_T \Delta T - \alpha_s \Delta S)$$

represents the strength of the AMOC, with $\gamma_0 > 0$ a flow constant, ρ_0 a reference density and α_T, α_S the thermal and haline expansion/contraction coefficients.

For the El-Niño Southern Oscillation, we use the conceptual model as proposed in Timmermann et al. (2003). This model has a state vector consisting of the temperature of the western Pacific T_1, the temperature of the eastern Pacific T_2 and the thermocline depth of the western Pacific h_1. The model finds its basis in the Zebiak and Cane (1987) ENSO model, with a two-strip and two-box approximation, and a shallow-water model for the upper ocean with a fixed mixed layer depth:

$$\begin{cases} \frac{dT_1}{dt} = & -\alpha(T_1 - T_r) - \frac{u(T_2 - T_1)}{L/2} \\ \frac{dT_2}{dt} = & -\alpha(T_2 - T_r) - \frac{w(T_2 - T_{sub})}{H_m} \end{cases} \tag{16}$$

with $1/\alpha$ a typical thermal damping timescale, T_{sub} the temperature below the mixed layer, H_m and L the depths of the mixed layer and basin width, respectively, w upwelling velocity and u atmospheric zonal surface wind being linear to wind stress: $u/(L/2) = \epsilon\beta\tau$ and $w/H_m = -\zeta\beta\tau$. The parameters ϵ and ζ refer to the strength of zonal and vertical advection (bifurcation parameters).

The subsurface temperature T_{sub} is parametrized as

$$T_{sub} = T_r - \frac{T_r - T_{r0}}{2}\left[1 - \frac{\tanh(H + h_2 - z_0)}{h*}\right] \tag{17}$$

with h_2 the east equatorial Pacific thermocline depth (calculated as deviation from a reference depth H), z_0 the depth for which w becomes its characteristic value and $h*$ the sharpness of the thermocline. The thermocline depths are calculated as follows:

$$\begin{cases} h_2 = & h_1 + bL\tau \\ \frac{dh_1}{dt} = & r(-h_1 - \frac{bL\tau}{2}) \end{cases} \tag{18}$$

where b the efficiency of wind stress τ to drive the thermocline tilt. For further details~~and parameter values~~, we refer to Timmermann et al. (2003). In the Stommel and Timmermann models, we use the standard parameter settings, as given in the references, unless stated otherwise.

~~The coupling of~~ We couple the AMOC and ENSO ~~systems is mainly through influence on the wind stress. In the original~~ models through the relation between the Atlantic meridional temperature gradient (in the Stommel model) on the Pacific zonal wind stress (in the Timmermann model), the second (more important) mechanism found in literature, described in the previous section. Even in a simplified model, the relation between wind stress and meridional temperature gradient is physically justified: thermal wind balance indicates a direct connection between the adjustment of wind stress to changes in the meridional temperature gradient. In the Timmermann model, the zonal wind stress τ is expressed as:

$$\tau = \frac{\mu(T_2 - T_1)}{\beta} \tag{19}$$

with μ/β parameters that control the influence of the zonal temperature gradient on the wind stress, set to be 0.02 Pa·K^{-1}. ~~However, in Dijkstra and Neelin (1995) it was argued that part of the contribution to the zonal wind stress ,~~ We add an external wind stress term τ_{ext} ~~, arises from processes outside the tropical Pacific. Here, we model τ_{ext} to be~~ that is dependent on the meridional temperature gradient in the Atlantic ΔT, i.e.,

$$\tau = \tau_{ext}(\Delta T) + \frac{\mu}{\beta}(T_2 - T_r) \tag{20}$$

with a negative relation between τ_{ext} and Atlantic meridional SST gradient ΔT as we know from literature described above (stronger positive ΔT results in stronger easterlies, thus negative τ_{ext}). Note that both the ~~total~~ Pacific wind stress τ and specifically its external part τ_{ext} should always be negative. The total wind stress is negative because this area (at low altitude) is strictly dominated by easterly winds, and τ_{ext} is negative because through the meridional temperature gradient, it reflects the influence of the zonal mean Hadley cell on the equatorial Pacific. Physically, the Hadley cell only induces negative zonal wind stress in this region.

In the coupling (20), we fix β and vary μ as the coupling parameter. For τ_{ext} we take a linear relation:

$$\tau_{ext} = \alpha_\tau \Delta T + \gamma_\tau - \tau_0 \tag{21}$$

where all coefficients are constant over time. The parameters α_τ and γ_τ can be estimated from the ESSENCE data as discussed in section 4.1, and ~~the parameter~~ τ_0 ~~is there to remove a~~ reflects the constant part in the zonal mean wind stress, which we subtract because we are interested in the contribution of changes in the meridional overturning. Using five ESSENCE runs per ensemble for both the standard forcing and hosing-experiment, respectively, ΔT is computed as the absolute difference between the SST in the North Atlantic area ($50-60°$N $\times$ $50-20°$W) and the Equatorial Atlantic region ($0-20°$N $\times$ $45-20°$W). For the wind stress τ_{ext}, the zonally integrated wind stress averaged over the region $0-10°$N is taken. In Fig. 7, 5-year running means of annual averages are plotted for the hosing simulations (in red) and the standard simulations (in black). Clearly, τ_{ext} decreases with increasing ΔT, such that when the AMOC collapses (larger ΔT) the wind stress τ_{ext} becomes more negative and the external part of the trade winds increases. However, we also note that the spread in the simulation data is large, which in part can be attributed to internal variability present in the simulations. The coefficients α_τ and γ_τ were found to be (from a linear fit) -0.000376 Pa·°C^{-1} and -0.0119 Pa. By looking at the ΔT regime in Fig. 7, τ_0 is chosen to be the wind stress at 19 °C: $\tau_0 = \alpha_\tau \cdot 19 + \gamma_\tau \approx -0.0190$ Pa. This results in a final quantized expression for the coupling:

$$\tau_{ext} \approx -0.000376 \cdot \Delta T + 0.00715 \tag{22}$$

4.3 Results

The AMOC model's bifurcation diagrams are shown in Figs. 8a and b, clearly showing a ~~back-to-back~~ saddle-node structure. For an interval of values of the freshwater flux F_s, the system has multiple equilibria, and for other values, only one equilibrium remains. This means that when we are in the high-Ψ branch and F_s is large enough, the system can make a transition to the low-Ψ branch. This is depicted by the blue arrow in Fig. 8b.

The bifurcation diagram of the ENSO model with τ_{ext} as parameter is shown in Fig. 8c. First of all, the bifurcation diagrams become much simpler than in the original Timmermann et al. (2003) model, the reason for this being extensively discussed in Dijkstra and Neelin (1995). Fig. 8d shows the influence of μ for the position of the oscillatory regime: on each branch, two Hopf bifurcations can be found and the μ value of the first Hopf bifurcation decreases with more negative τ_{ext}. This indicates that ~~the~~ a Hopf bifurcation can be crossed if τ_{ext} is decreased, while μ is kept constant. In other words, for the right value of μ, the eastern Pacific SST starts oscillating (El-Niño ~~intensifies~~ 'intensifies') when the easterly external wind is increased. For the coupled model, we use $\mu = 0.00146$.

Using τ_{ext} to couple the AMOC and ENSO models, we performed simulations with $\Delta t = 0.1$ days and the Runge-Kutta fourth order integration method. To initiate the collapse of the overturning, a freshwater forcing F_s is applied in the form of a step function:

$$F_s = \begin{cases} 0.006 & \text{if } t \leq 500 \text{ y} \\ 0.01 & \text{if } t > 500 \text{ y} \end{cases} \tag{23}$$

Using the coupling of Eqn. 22, we attain the results shown in Fig. 9.The exact quantification of ~~this~~ the coupling partly modulates which effect the collapse of the AMOC has on ENSO. For the chosen coupling, the collapse of the overturning leads to the crossing of the first Hopf bifurcation point in the following system, and an oscillation starts growing. ~~Hence,~~ As is visible in Fig. 6, the relation between ΔT and τ has quite some spread, which implies a large uncertainty in the values of α_τ and γ_τ. We would like to stretch that the regime in which the Hopf bifurcation is crossed is dependent on multiple variables, among which are these coupling parameters. Running the forward integration of the coupled model for values between $\alpha_\tau = -0.00041$ and $\alpha_\tau = -0.00033$, learns that (*ceteris paribus*), for higher (lower) values of α_τ, the oscillation indeed becomes weaker (stronger), down to a disappearance of the oscillation at $\alpha_\tau \approx -0.000335$ at a time step of 0.25 days.

Despite the parameter sensitivity, this is a typical illustration of the fold-Hopf cascading ~~behavior discussed above.~~ behaviour discussed in earlier sections. This motivates the possibility that cascading transitions are possible in real physical systems.

5 Summary, Discussion and Conclusions

In this paper, we introduced the concept of cascading tipping, which can occur when a transition in a leading system alters background conditions for a following system such that it also undergoes a transition. We presented a mathematical framework around this concept, where we used generic bifurcations (~~back-to-back~~ saddle-node and Hopf) in both leading and following systems. Four types of deterministic dynamical systems with the possibility for cascading events were formulated, including the double-fold cascade, the fold-Hopf cascade, the Hopf-fold cascade and the double-Hopf cascade. In all cases we assumed a linear coupling between the following and leading system. The fold-fold coupled system has been considered before in another context (Brummitt et al., 2015), where it also has been noted that in systems with more than two coupled fold cascades not all subsystems undergo tipping ('hopping'). Moreover, stochastically coupled multi-stable systems have been considered in networks, where different types of domino effects can occur depending on the synchrony of the transition in the different network nodes (Ashwin et al., 2017; Creaser et al., 2018). Here we consider only two coupled systems, but allow different types of bifurcations, and the systems are physically coupled in a directional way.

We have discussed statistical indicators and analysis tools for cascading tipping points. Indicators for cascading tipping points are found in detrended cross correlation analysis (DCCA) and a special case of extrapolation using the DFA of the following system. These tools were applied in simulations involving both the double-fold and fold-Hopf cascades. The increased variance, AR(1) and DFA scaling exponent are clearly found in each case of single tipping. The cross-correlation indicators (DCCA and ρ_{DCCA}) did not evolve much throughout the time series, which indicates their insensitivity with respect to prox-

imity to single tipping points. Several limitations on the use of these variables have been mentioned. However, it seems that these metrics are highly sensitive to window and segment sizes, which leaves their potential as early warnings of cascading transition events inconclusive. The ratios of autoregressive metrics before and after the first transition seem to be a stronger warning of cascading transitions. More research is needed to exactly quantify these metrics.

The concept of cascading tipping was applied to study the ~~behavior~~ behaviour of a model describing a link between the Atlantic Meridional Overturning Circulation (MOC) and ENSO. We modelled this using a coupling between the Stommel (1961) model and the Timmermann et al. (2003) ENSO-model by a meridional temperature gradient-dependent term in the external wind stress of the ENSO model. Through analysis of the bifurcation diagrams and simulations, a cascading tipping event is indeed possible ~~in this case and our results are presented in the light~~ within this model in the form of the fold-Hopf cascade. Obviously, both models are highly idealized and more detailed models of both AMOC and ENSO are needed to demonstrate the occurrence of such a cascading transition in the climate system.

A potential example of a double-fold cascade, that was not further treated here, could be the impact of a bistable MOC on the (bistable) land ice formation on the Antarctic continent. In this case the coupling exists through the atmospheric CO_2 concentration, which depends on mixing and circulation in the ocean while strongly determining the existence of an ice sheet (DeConto and Pollard, 2003). During the Eocene-Oligocene transition, where a large ice sheet grew on Antarctica, a two-step ~~two-step~~ signal is observed in the deep-sea $\delta^{18}O$ ratio, suggesting two abrupt transitions. Using a box model by Gildor and Tziperman (2000), Tigchelaar et al. (2011) showed that a two-step signal can be produced by first a MOC transition which changes the CO_2 concentration such that a transition occurs in the land-ice model. Although from a physical perspective, this is a potential example of a cascading transition, we make no claim about whether such a transition likely occurred at the Eocene-Oligocene transition. Here, also more detailed models are needed and transition are expected to be more complicated (Tantet et al., 2018).

These two applications ~~indicate that there~~ reflect the relevance of this paper. There are likely many cases in which these cascading events occur in climate and therefore highlight the importance of the topic. Future research will point out whether these events are likely to happen in the future climate and whether these effects also occur in other fields than climate science. Of course, ~~the theory~~ this paper covers the very basics of deterministic cascading events. One can imagine a wide range of phenomena if more complicated transitions between attractors are considered and when noise is included. For example, when a leading chaotic system is coupled to a deterministic following system with a ~~back-to-back~~ saddle-node bifurcation structure, a slight change in the chaotic attractor may change the background conditions for the following system such that it undergoes a transition. An application here may be the effect of a midlatitude atmospheric jet on the Atlantic MOC. We hope that this paper will stimulate more research on the various types of cascading tipping and also on the development of well-suited indicators and early warnings of such events.

Acknowledgements. AvdH and HAD acknowledge support by the Netherlands Earth System Science Centre (NESSC), financially supported by the Ministry of Education, Culture and Science (OCW), Grant no. 024.002.001. AvdH thanks Peter Ashwin for discussions in relation

to this work and thanks the University of Exeter and the EPSRC funded Past Earth Network (Grant number EP/M008363/1) for funding an extended visit to the University of Exeter during the Summer of 2017.

References

Aleina, F. C., Baudena, M., D'Andrea, F., and Provenzale, A.: Multiple equilibria on planet Dune: Climate-vegetation dynamics on a sandy planet, Tellus B, 65, http://www.tellusb.net/index.php/tellusb/article/view/17662, 2013.

Ashwin, P., Wieczorek, S., Vitolo, R., and Cox, P.: Tipping points in open systems: Bifurcation, noise-induced and rate-dependent examples in the climate system, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 370, 1166–1184, doi:10.1098/rsta.2011.0306, http://rsta.royalsocietypublishing.org/cgi/doi/10.1098/rsta.2011.0306, 2012.

Ashwin, P., Creaser, J., and Tsaneva-Atanasova, K.: Fast and slow domino effects in transient network dynamics, Phys. Rev. E, 96, 459, http://arxiv.org/abs/1701.06148, 2017.

Barriopedro, D., García-Herrera, R., Lupo, A. R., and Hernández, E.: A Climatology of Northern Hemisphere Blocking, Journal of Climate, 19, 1042–1063, doi:10.1175/JCLI3678.1, http://dx.doi.org/10.1175/JCLI3678.1, 2006.

Bathiany, S., van der Bolt, B., Williamson, M., Lenton, T., Scheffer, M., van Nes, E., and Notz, D.: Statistical indicators of Arctic sea-ice stability - prospects and limitations, The Cryosphere, 10, 1631–1645, dOI: 10.5194/tc-10-1631-2016, 2016.

Brummitt, C. D., Barnett, G., and D'Souza, R. M.: Coupled catastrophes: sudden shifts cascade and hop among interdependent systems., Journal of The Royal Society Interface, 12, 20150 712–12, doi:10.1098/rsif.2015.0712, http://rsif.royalsocietypublishing.org/lookup/doi/10.1098/rsif.2015.0712, 2015.

Cessi, P.: A simple box model of stochastically forced thermohaline flow, J. Phys. Oceanogr., 24, 1911–1920, 1994.

Creaser, J., Tsaneva-Atanasova, K., and Ashwin, P.: Sequential noise-induced escapes for oscillatory network dynamics, SIAM Journal on Applied Dynamical Systemsiam Journal on Applied Dynamical Systems, 17, 500–525, doi:10.1137/17M1126412, http://arxiv.org/abs/1705.08462, 2018.

Dakos, V., Scheffer, M., van Nes, E. H., Brovkin, V., Petoukhov, V., and Held, H.: Slowing down as an early warning signal for abrupt climate change, Proceedings of the National Academy of Sciences, 105, 14 308–14 312, doi:10.1073/pnas.0802430105, http://www.pnas.org/content/105/38/14308.abstract, 2008.

DeConto, R. M. and Pollard, D.: Rapid Cenozoic glaciation of Antarctica induced by declining atmopsheric CO_2, Nature, 421, 245–249, 2003.

Dijkstra, H. A. and Neelin, D.: Ocean-Atmosphere Interaction and the Tropical Climatology. Part II: Why the Pacific Cold Tongue is in the East, J. Climate, 8, 1343–1359, printed, 1995.

Dong, B. and Sutton, R.: Enhancement of ENSO Variability by a Weakened Atlantic Thermohaline Circulation in a Coupled GCM, Journal of Climate, 20, 4920–4939, doi:10.1175/JCLI4284.1, http://dx.doi.org/10.1175/JCLI4284.1, 2007.

Elsworth, G., Galbraith, E., Halverson, G., and Yang, S.: Enhanced weathering and CO2 drawdown caused by latest Eocene strengthening of the Atlantic meridional overturning circulation, Nature Geoscience, doi:10.1038/ngeo2888, http://www.nature.com/doifinder/10.1038/ngeo2888, 2017.

Gasson, E., Lunt, D. J., DeConto, R. M., Goldner, A., Heinemann, M., Huber, M., LeGrande, A. N., Pollard, D., Sagoo, N., Siddall, M., Winguth, a., and Valdes, P. J.: Uncertainties in the modelled CO_2 threshold for Antarctic glaciation, Climate of the Past, 10, 451–466, doi:10.5194/cp-10-451-2014, http://www.clim-past.net/10/451/2014/, 2014.

Gildor, H. and Tziperman, E.: Sea ice as the glacial cycles' Climate switch: role of seasonal and orbital forcing, Paleoceanography, 15, 605–615, doi:10.1029/1999PA000461, http://dx.doi.org/10.1029/1999PA000461, 2000.

Held, H. and Kleinen, T.: Detection of climate system bifurcations by degenerate fingerprinting, Geophysical Research Letters, 31, n/a–n/a, doi:10.1029/2004GL020972, http://dx.doi.org/10.1029/2004GL020972, l23207, 2004.

Hirota, M., Holmgren, M., Van Nes, E. H., and Scheffer, M.: Global Resilience of Tropical Forest and Savanna to Critical Transitions, Science, 334, 232–235, doi:10.1126/science.1210657, http://science.sciencemag.org/content/334/6053/232, 2011.

Jungclaus, J. H., Haak, H., Esch, M., Roeckner, E., and Marotzke, J.: Will Greenland melting halt the thermohaline circulation?, Geophysical Research Letters, 33, n/a–n/a, doi:10.1029/2006GL026815, http://dx.doi.org/10.1029/2006GL026815, l17708, 2006.

Kriegler, E., W Hall, J., Held, H., Dawson, R., and Schellnhuber, H.: Imprecise Probability Assessment of Tipping Points in the Climate System, Proceedings of the National Academy of Sciences of the United States of America, 106, 5041–6, 2009.

Kuehn, C.: A mathematical framework for critical transitions: Bifurcations, fast-slow systems and stochastic dynamics, Physica D, 240, 1020–1035, doi:10.1016/j.physd.2011.02.012, http://dx.doi.org/10.1016/j.physd.2011.02.012, 2011.

Kutzbach, J., Bonan, G., Foley, J., and Harrison, S.: Vegetation and soil feedbacks on the response of the African monsoon to orbital forcing in the early to middle Holocene, Nature, 384, 623–626, doi:10.1038/384623a0, http://dx.doi.org/10.1038/384623a0, 1996.

Lenton, T. M.: Early Warning of Climate Tipping Points, Nature Climate Change, 1, 201–209, 2011.

Lenton, T. M. and Williams, H. T.: On the origin of planetary-scale tipping points, Trends in Ecology & Evolution, 28, 380 – 382, doi:http://dx.doi.org/10.1016/j.tree.2013.06.001, http://www.sciencedirect.com/science/article/pii/S0169534713001456, 2013.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J.: Tipping elements in the Earth's climate system, Proceedings of the National Academy of Sciences, 105, 1786–1793, doi:10.1073/pnas.0705414105, http://www.pnas.org/content/105/6/1786.abstract, 2008.

Livina, V. and Lenton, T.: A modified method for detecting incipient bifurcations in a dynamical system, Geophysical Research Letters, 34, n/a–n/a, doi:10.1029/2006GL028672, http://dx.doi.org/10.1029/2006GL028672, l03712, 2007.

Pearson, P., Foster, G., and Wade, B.: Atmospheric carbon dioxide through the Eocene–Oligocene climate transition, Nature, 461, 1110–3, 2009.

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., and Goldberger, A. L.: Mosaic organization of DNA nucleotides, Phys. Rev. E, 49, 1685–1689, doi:10.1103/PhysRevE.49.1685, https://link.aps.org/doi/10.1103/PhysRevE.49.1685, 1994.

Podnobik, B. and Stanley, H.: Detrended Cross-Correlation Analysis: A New Method for Analyzing Two Non-stationary Time Series, Physical Review Letters, 100(8), 084 102, 2008.

Scheffer, M., Bascompte, J., Brock, W., Brovkin, V., Carpenter, S., Dakos, V., Held, H., van Nes, E., Rietkerk, M., and Sugihara, G.: Early-warning signals for critical transitions, Nature, 461, 53–59, doi:10.1038/nature08227, 2009.

Sterl, A., Severijns, C., Dijkstra, H. A., Hazeleger, W., van Oldenborgh, G. J., van den Broeke, M., Burgers, G., van den Hurk, B., van Leeuwen, P. J., and van Velthoven, P.: When can we expect extremely high surface temperatures?, Geophys. Res. Lett., 35, L14 703, 2008.

Stommel, H.: Thermohaline Convection with Two Stable Regimes of Flow, Tellus A, 13, http://www.tellusa.net/index.php/tellusa/article/view/9491, 1961.

Tantet, A., Lucarini, V., and Dijkstra, H. A.: Resonances in a Chaotic Attractor Crisis of the Lorenz Flow, Journal of Statistical Physics, 170, 584–616, doi:10.1007/s10955-017-1938-0, https://doi.org/10.1007/s10955-017-1938-0, 2018.

Thompson, J. and Stewart, H.: Nonlinear Dynamics and Chaos, Wiley, https://books.google.nl/books?id=vQvu-5VPkmcC, 2002.

Thompson, J. M. T. and Sieber, J.: Predicting Climate Tipping as a noisy bifurcation: a review, International Journal of Bifurcation and Chaos, 21, 399–423, doi:10.1142/S0218127411028519, http://www.worldscientific.com/doi/abs/10.1142/S0218127411028519, 2011.

Tigchelaar, M., von der Heydt, A. S., and Dijkstra, H. A.: A new mechanism for the two-step δ^{18}O signal at the Eocene-Oligocene boundary, Climate of the Past, 7, 235–247, doi:10.5194/cp-7-235-2011, http://www.clim-past.net/7/235/2011/, 2011.

Timmermann, A., Jin, F., and Abshagen, J.: A Nonlinear Theory for El Niño Bursting, Journal of the Atmospheric Sciences, 60, 152–165, doi:10.1175/1520-0469(2003)060<0152:ANTFEN>2.0.CO;2, http://dx.doi.org/10.1175/1520-0469(2003)060<0152:ANTFEN>2.0.CO;2, 2003.

Timmermann, A., An, S.-I., Krebs, U., and Goosse, H.: ENSO Suppression due to Weakening of the North Atlantic Thermohaline Circulation, Journal of Climate, 18, 3122–3139, doi:10.1175/JCLI3495.1, http://dx.doi.org/10.1175/JCLI3495.1, 2005.

Timmermann, A., Okumura, Y., An, S.-I., Clement, A., Dong, B., Guilyardi, E., Hu, A., Jungclaus, J., Renold, M., Stocker, T., R.J., S., Sutton, R., Xie, S.-P., and Yin, J.: The Influence of a Weakening of the Atlantic Meridional Overturning Circulation on ENSO, Journal of Climate, 20, 4899–4919, doi:10.1175/JCLI4283.1, http://dx.doi.org/10.1175/JCLI4283.1, 2007.

Xie, S.-P., Miyama, T., Wang, Y., Xu, H., de Szoeke, S. P., Small, R. J. O., Richards, K. J., Mochizuki, T., and Awaji, T.: A Regional Ocean–Atmosphere Model for Eastern Pacific Climate: Toward Reducing Tropical Biases, Journal of Climate, 20, 1504–1522, doi:10.1175/JCLI4080.1, http://dx.doi.org/10.1175/JCLI4080.1, 2007.

Yuan, N., Zuntau, F., Zhang, H., Piao, L., Xoplaki, E., and Luterbacher, J.: Detrended Partial-Cross-Correlation Analysis: A New Method for Analyzing Correlations in Complex System, Scientific Reports, 5, doi:doi:10.1038/srep08143, 2015.

Zebende, G.: {DCCA} cross-correlation coefficient: Quantifying level of cross-correlation, Physica A: Statistical Mechanics and its Applications, 390, 614 – 618, doi:https://doi.org/10.1016/j.physa.2010.10.022, http://www.sciencedirect.com/science/article/pii/S0378437110008800, 2011.

Zebiak, S. E. and Cane, M. A.: A Model El Niño–Southern Oscillation, Monthly Weather Review, 115, 2262–2278, doi:10.1175/1520-0493(1987)115<2262:AMENO>2.0.CO;2, http://dx.doi.org/10.1175/1520-0493(1987)115<2262:AMENO>2.0.CO;2, 1987.

Zhang, D., Lee, T. N., Johns, W. E., Liu, C.-T., and Zantopp, R.: The Kuroshio East of Taiwan: Modes of Variability and Relationship to Interior Ocean Mesoscale Eddies, Journal of Physical Oceanography, 31, 1054–1074, doi:10.1175/1520-0485(2001)031<1054:TKEOTM>2.0.CO;2, http://dx.doi.org/10.1175/1520-0485(2001)031<1054:TKEOTM>2.0.CO;2, 2001.

Zhang, R. and Delworth, T. L.: Simulated Tropical Response to a Substantial Weakening of the Atlantic Thermohaline Circulation, Journal of Climate, 18, 1853–1860, doi:10.1175/JCLI3460.1, https://doi.org/10.1175/JCLI3460.1, 2005.

Zhou, W.: Multifractal detrended cross-correlation analysis for two nonstationary signals, Physical Review, 77, https://arxiv.org/abs/0803.2773v1, 2008.

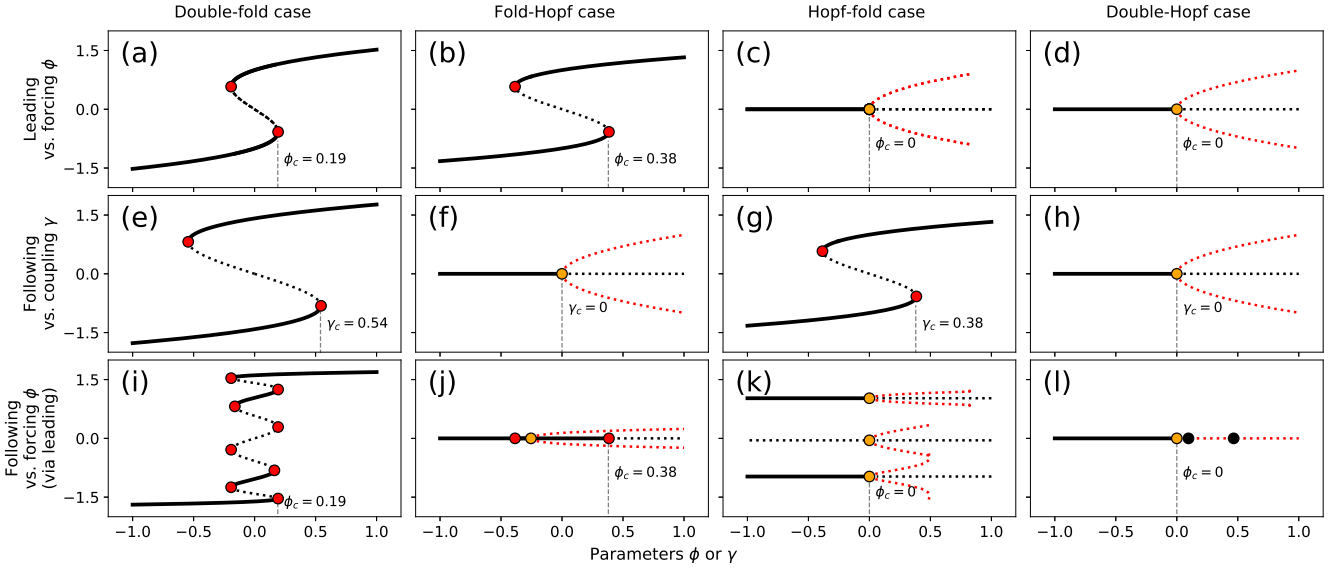

Figure 1. ~~Bifurcation diagrams of coupled dynamical systems that show various cascading tipping types. Forcing parameter ϕ of the leading system is used on the horizontal axis, the leading system's~~ Stable ~~(top) or following system's (bottom) equilibria are shown on the vertical axis. Lines indicate stable equilibria (black~~ solid), unstable ~~equilibria~~ (black ~~dashed~~dotted) and oscillatory ~~equilibria~~ (amplitude red ~~dashed lines indicate non-zero amplitudes in pairs~~dotted) regimes of various cascading tipping types (as depicted on top of the figure). ~~Dots~~ Black, orange and black dots indicate ~~important bifurcation points~~fold, Hopf and torus bifurcations, respectively. Top: ~~limit points (red/orange for the~~ leading ~~/~~system versus forcing ϕ. Middle: following system versus coupling γ. Bottom: (coupled) ~~, Hopf bifurcation points~~ following system versus forcing ϕ. Critical values of (~~blue/green for the~~ coming from lower branch) leading ~~/~~tipping (top), following ~~system~~tipping (middle) and ~~torus bifurcation points (black/grey for~~ the ~~leading/following system~~combined cascading tipping event (bottom) are marked by the grey vertical dashed line.

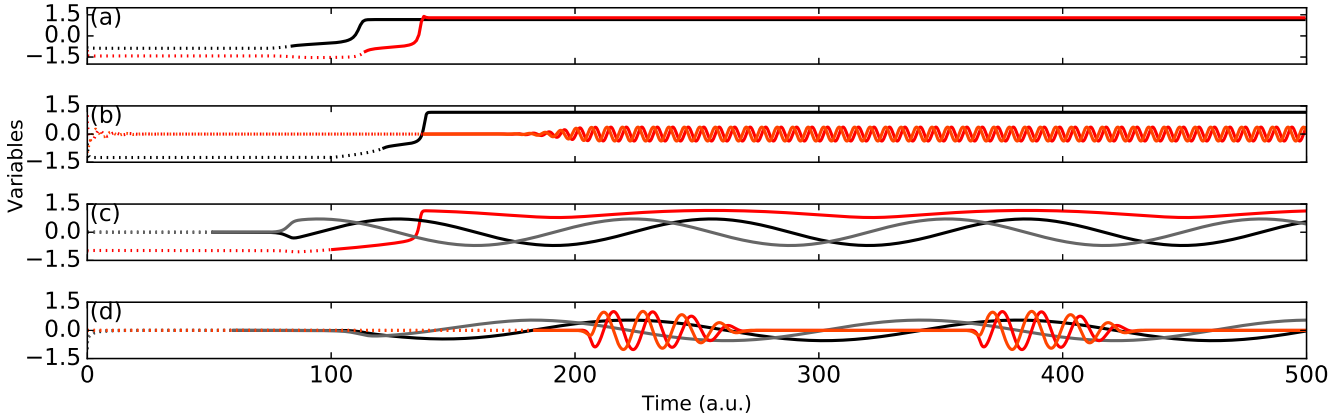

Figure 2. Example simulations for each cascading event type: the double-fold cascade (a), the fold-Hopf cascade (b), the Hopf-fold cascade (c) and the double Hopf cascade (d). Black and grey lines indicate the leading systems, red and orange lines indicate the following systems. Dotted lines indicate time before the critical threshold in the forcing $\phi(t)$ (black/grey) or coupling $\kappa(x)$ (red/orange) is reached, solid lines indicate the time after this. Parameter values for the modelled systems are given in Table 1.

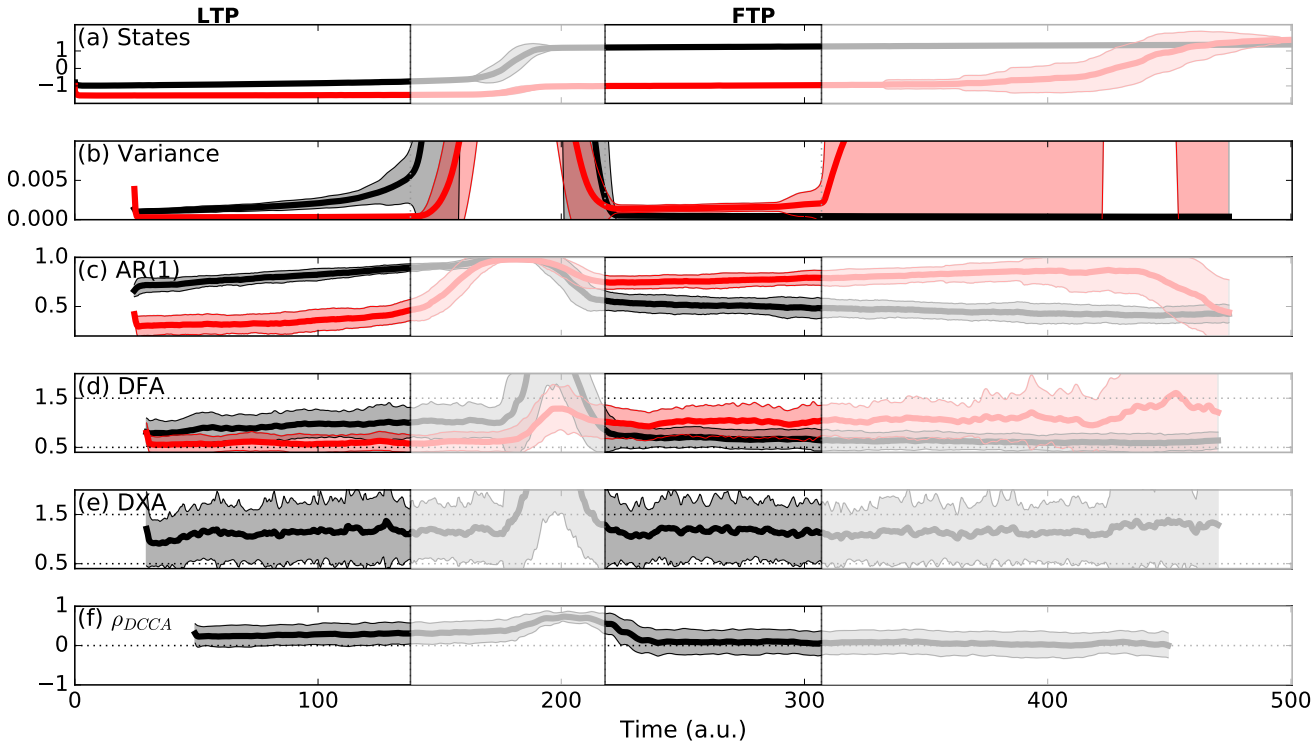

Figure 3. Ensemble (100-member) simulations of a dynamical system undergoing a double-fold cascade (Eqn. 13) where both systems undergo a transition, parameter values as in Table 2; (a) states of x (black) and y (red), (b) variance of x (black) and y (red), (c) autoregressive coefficient at lag 1 of x (black) and y (red), (d) detrended fluctuation analysis scaling exponent of x (black) and y (red), (e) detrended cross-correlation analysis scaling exponent and (f) detrended cross-correlation coefficient by Zebende (2011). White-shaded areas indicate windows containing the actual transitions. The increased variance, AR(1) and DFA scaling exponent prior to transition in the leading system and following system, respectively confirms the predicted increased memory through critical slowing down.

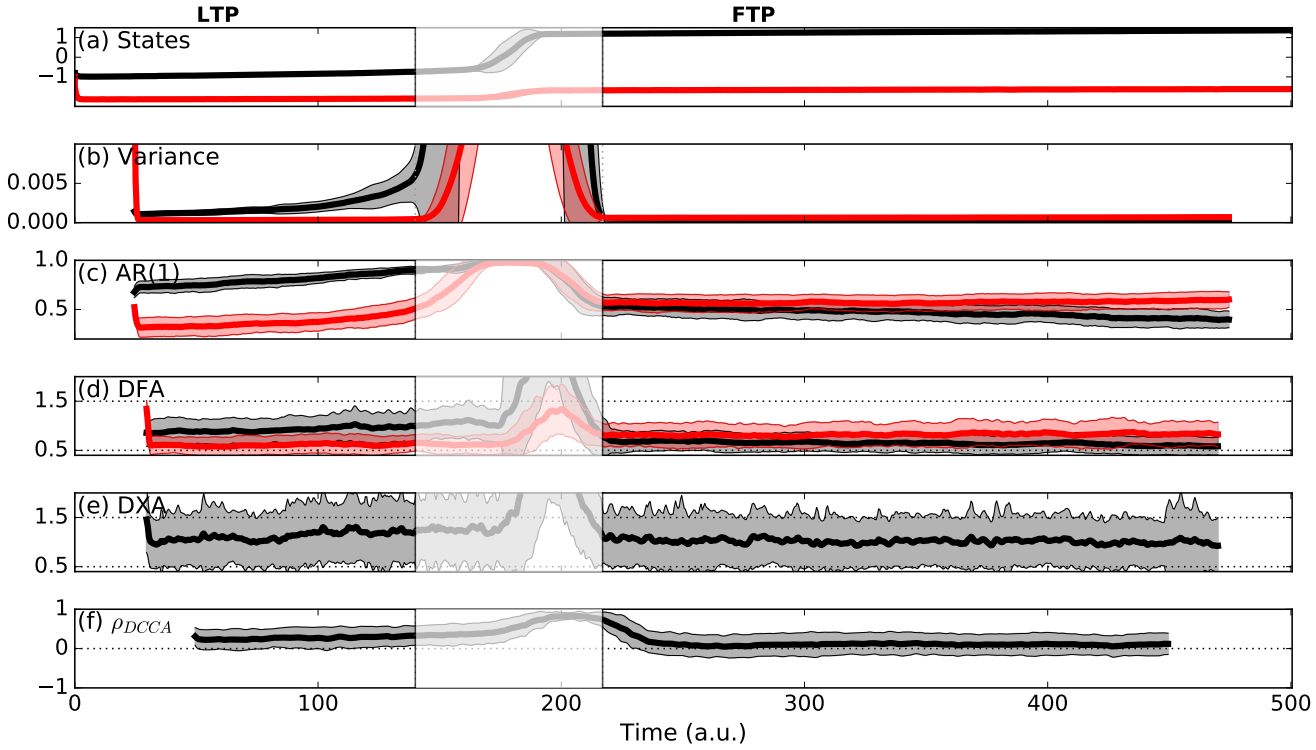

Figure 4. As in Fig. 3, but without any transition in the following system. In this case, only the leading system has a transition. Parameter values are given in ~~Tab.~~Table 2.

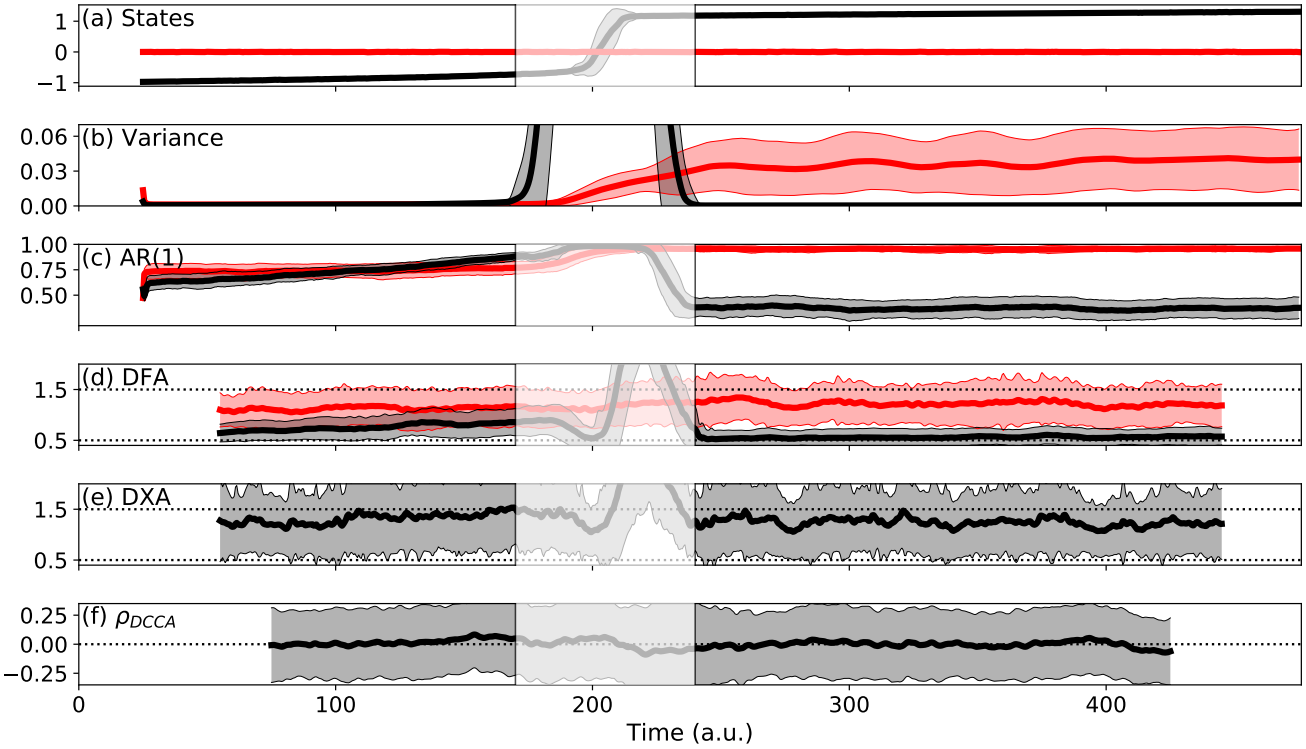

Figure 5. As in Fig. 3, but for the fold-Hopf cascade (Eqn. 14). Parameter values are given in Table 2.

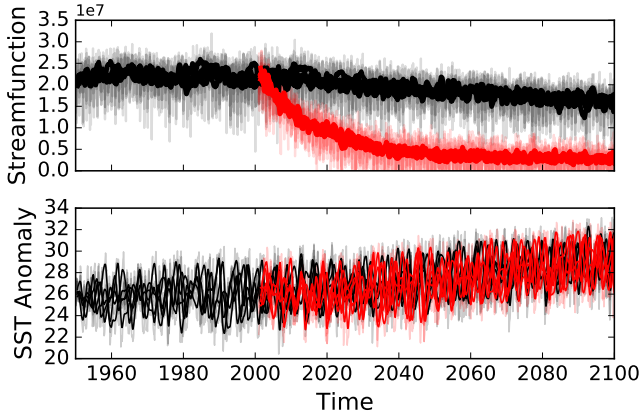

Figure 6. Top panel: Evolution of the five standard SRES-A1b runs (blue) and five HOSING-1 runs (red) in terms of the overturning streamfunction. Bottom panel: NINO3.4 of the standard SRES-A1b ensemble (blue) and the HOSING-1 ensemble (red). Shaded thin lines indicate monthly means, thick lines indicate the deseasonalised values.

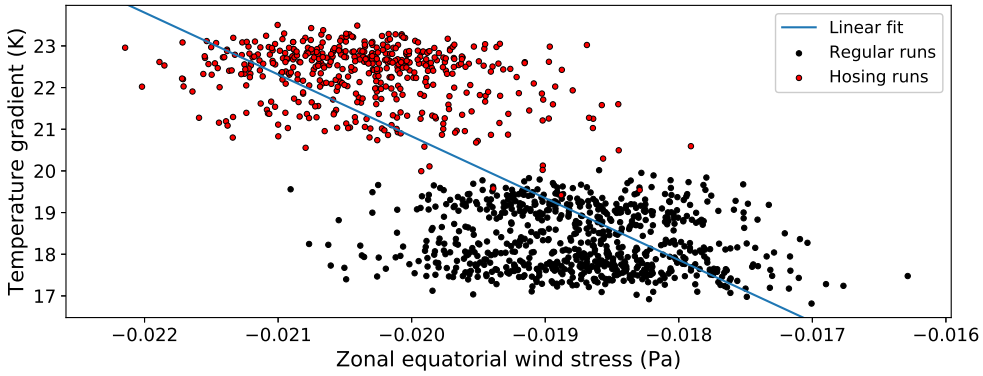

Figure 7. Zonal equatorial wind stress versus Atlantic temperature gradient. Data from the ESSENCE (Ensemble SimulationS of Extreme weather events under Nonlinear Climate changE) project is used, where the black dots refer to five members of the standard ensemble with SRES-A1b forcing (period 1950-2100), and the red dots refer to five members of the HOSING-1 ensemble where in 2000 a freshwater perturbation is applied (i.e., period 2000-2100). Five year running mean is applied, yearly averages are shown. The zonal equatorial wind stress here is defined as the average zonal wind stress over the latitudonal band 0-10°N. The Atlantic temperature gradient is defined as the difference between the SST in a northern box (50-60°N, 50-20°W) and a southern box (0-20°N, 45-20°W). The blue line indicates a linear fit.

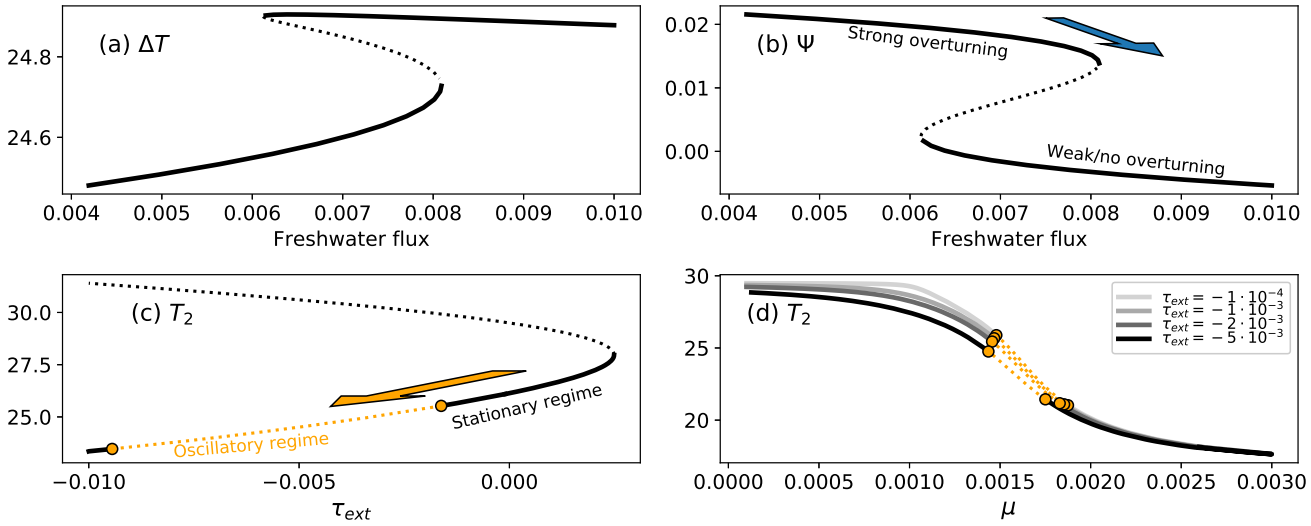

Figure 8. Bifurcation diagrams and forward runs of the Stommel (top panel) and of the Timmermann (bottom panel) models. Blue arrow indicates the collapse of the overturning circulation, which amplifies (negatively) the external zonal wind stress in the Pacific τ_{ext}, such that (orange arrow), the system enters an oscillatory state. Orange arrow indicates subsequent tipping in the following (ENSO) system. Top panels: (a) Meridional temperature gradient equilibria versus freshwater flux, (b) non-dimensional stream function versus freshwater flux. These figures show the multiple states of the overturning. Bottom panels: (c) eastern equatorial Pacific SST versus τ_{ext} (for $\mu = 0.00145$), showing a regime where the system is stationary and a regime where the system is oscillatory, (d) eastern equatorial Pacific SST versus μ for different values of τ_{ext}. Orange dots indicate Hopf bifurcation points, orange dotted lines indicate oscillatory regimes. Black and grey solid lines indicate stable equilibria, black dashed liens indicate unstable equilibria.

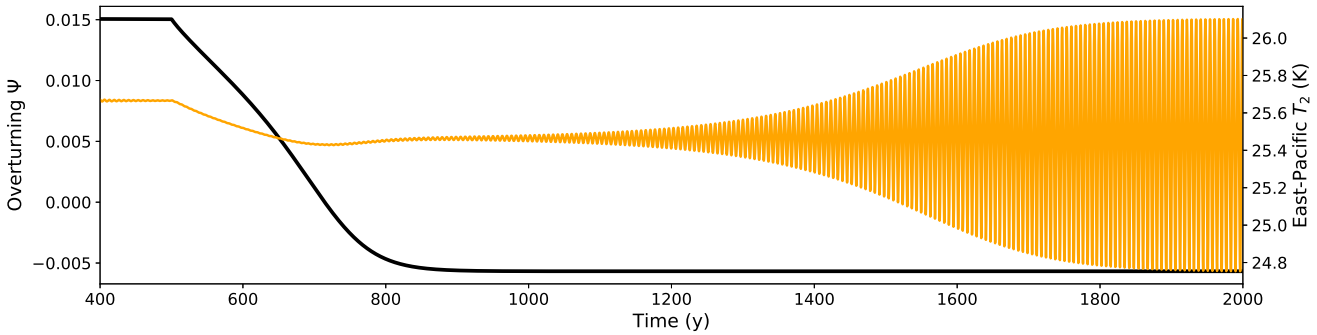

Figure 9. Simulation run of the coupled Stommel-Timmermann model for different model configurations, where the collapse of the overturning flow function (black) leads to the crossing of a Hopf bifurcation in the eastern equatorial-Pacific SST (orange). Parameter values as in Timmermann et al. (2003), with $\mu = 0.00146$.

Table 1. Parameter values and coupling for the four types of cascading tipping as shown in Figures 1 and 2.

Double fold (Eqn. 3)	**Fold-Hopf** (Eqn. 4)	**Hopf-fold** (Eqn. 5)	**Double Hopf** (Eqn. 6)
Leading system			
$\phi_c = \pm 0.19$ (Fold)	$\phi_c = \pm 0.38$ (Fold)	$\phi_c = 0$ (Hopf)	$\phi_c = 0$ (Hopf)
Bistable for	Bistable for	Oscillatory for	Oscillatory for
$\lvert\phi\rvert < \sqrt{\frac{-4a_1^3 a_2^3}{27a_1^4}}$	$\lvert\phi\rvert < \sqrt{\frac{-4a_1^3 a_2^3}{27a_1^4}}$	$\phi > 0$	$\phi > 0$
(if $a_1 < 0, a_2 > 0$)	(if $a_1 < 0, a_2 > 0$)	(if $a_1 b_1 < 0$)	(if $a_1 b_1 < 0$)
Coupling			
$\gamma = 0.48x$	$\gamma = -0.1 + 0.12x$	$\gamma = 0.05 + 0.5x$	$\gamma = -0.05 + 2x$
Following system			
$\gamma_c = \pm 0.54$ (Fold)	$\gamma_c = 0$ (Hopf)	$\gamma_c = \pm 0.38$ (Fold)	$\gamma_c = 0$ (Hopf)
Bistable for	Oscillatory for	Bistable for	Oscillatory for
$\lvert\gamma\rvert < \sqrt{\frac{-4b_1^3 b_2^3}{27b_1^4}}$	$\gamma > 0$	$\lvert\gamma\rvert < \sqrt{\frac{-4c_1^3 c_2^3}{27c_1^4}}$	$\gamma > 0$
(if $b_1 < 0, b_2 > 0$)	(if $b_1 c_1 < 0$)	(if $c_1 < 0, c_2 > 0$)	(if $c_1 d_1 < 0$)
Parameters			
$a_1 = -0.5$	$a_1 = -1$	$a_1 = 0.05$; $a_2 = 1$	$a_1 = 0.04$; $a_2 = 2$
$a_2 = 0.5$	$a_2 = 1$	$b_1 = -0.05$; $b_2 = 1$	$b_1 = -0.04$; $b_2 = 2$
$b_1 = -0.5$	$b_1 = b_2 = 1$	$c_1 = -1$	$c_1 = 0.4$; $c_2 = 1$
$b_2 = 1.0$	$c_1 = -1$; $c_2 = 1$	$c_2 = 1$	$d_1 = -0.4$; $d_2 = 1$

Table 2. Parameter values, coupling and initial conditions for the ensemble simulations of the Double Fold and Fold-Hopf systems as shown in Figures 3, 4 and 5.

Double fold (Eqn. 13) (following system tips)	**Double fold** (Eqn. 13) (following system does not tip)	**Fold-Hopf** (Eqn. 14)
Forcing and coupling		
$\phi(t) = 0.0012t$	$\phi(t) = 0.0012t$	$\phi(t) = 0.002t$
$\gamma(x) = 0.05 + 0.37x$	$\gamma(x) = 0.05 + 0.37x$	$\gamma(x) = -0.2 + 0.3x$
Parameters		
$a_1 = -0.5$	$a_1 = -0.5$	$a_1 = -1$
$a_2 = 0.5$	$a_2 = 0.5$	$a_2 = 1$
$b_1 = -0.5$	$b_1 = -0.25$	$b_1 = 0.1; b_2 = 1$
$b_2 = 1.0$	$b_2 = 1$	$c_1 = -0.5$; $c_2 = 1$
Integration time		
$t_{max} = 500$	$t_{max} = 500$	$t_{max} = 500$
$\Delta T = 0.5$	$\Delta T = 0.5$	$\Delta T = 0.5$
Noise		
Noise mean = 0	Noise mean = 0	Noise mean = 0
Noise variance = 0.1	Noise variance = 0.1	Noise variance = 0.1
Initial conditions		
$(x_0, y_0) = (-0.8, -1)$	$(x_0, y_0) = (-0.8, -1)$	$(x_0, y_0, z_0) = (-0.5, 1, -1)$

Table 3. Comparison of the ratios of autoregressive variables prior to and after the first transition, using the ensembles shown in Fig. 3 and Fig. 4.

Ratios	Mean	Standard deviation
With second tipping		
Leading variance	0.24	0.44
Leading AR(1)	0.62	0.09
Leading DFA	0.79	0.27
Following variance	3.95	1.53
Following AR(1)	1.92	0.39
Following DFA	1.70	0.49
Without second tipping		
Leading variance	0.15	0.06
Leading AR(1)	0.60	0.08
Leading DFA	0.74	0.23
Following variance	1.63	0.42
Following AR(1)	1.34	0.31
Following DFA	1.40	0.41

Table 4. Results of Student's t-test on the differences between the ratios (in ~~Tab.~~Table 3) of the cases *with* and *without* second tipping. *p-values calculated using the scipy.stats Python package.

t-test variable	T-statistic	dF	p-value*
Leading variance	1.62	101.67	0.11027
Leading AR	1.32	197.07	0.18904
Leading DFA	1.23	190.60	0.22264
Following variance	13.73	112.38	0.0
Following AR	9.93	184.12	0.0
Following DFA	4.13	190.17	0.00006

Table 5. NINO3.4 statistics (of deseasonalised data) for the different ensembles. The uncertainty stated is the standard deviation among the five runs within the ensemble. It is visible that in the case of a collapsed overturning, El-Niño intensifies more than without a collapsed overturning.

Time period	Ensemble	Variable	Value
1950-2000	Standard SRES-A1b	Mean	25.86 ± 0.046
	Standard SRES-A1b	Variance	$\mathbf{1.705} \pm 0.447$
2001-2100	Standard SRES-A1b	Mean	27.51 ± 0.032
	Standard SRES-A1b	Variance	$\mathbf{2.581} \pm 0.112$
2001-2100	HOSING-1	Mean	27.27 ± 0.053
	HOSING-1	Variance	$\mathbf{3.21} \pm 0.42$

Response to referee #1 (Alexis Tantet)

We thank the referee for the careful reading and the useful comments and will adapt the manuscript accordingly. Below is a point by point reply with the referees comments in bold font, our reply in italic font and the changes in manuscript in normal font.

1. Comment of the referee:

Section 2: The authors first describe possible scenarios of cascading tipping by combining the normal forms most relevant for applications and involving only one or a pair of stability exponents crossing the imaginary axis. As such, the framework is suited for coupled systems for which both the leading and the following systems are close to a saddle and/or a Hopf bifurcation, a situation relevant for the applications considered here. However, the climate system is a high-dimensional system with a large number of positive Lyapunov exponents, whereas the bifurcations considered here involve only one or two-dimensional attractors rather that chaotic sets. As such, while the mathematical framework considered here appears to be an important direction to explore for climate applications, I would consider it only as a first important step towards understanding more complex abrupt climate changes, such as the one studied in section 4. This point could be discussed more by the authors.

Authors reply:

We agree that abrupt climate changes in reality are connected to more complex chaotic sets and impossible to attribute to a single bifurcation or two bifurcations. As the referee also points to, the aim of this paper is to give a framework of cascading transitions with mathematical examples, analyses and applications to conceptual models. The step towards the real climate system should be taken with care. Especially in the beginning of the paper this can indeed be made more clear. In the discussion section, this was already mentioned (e.g. page 16, line 5-7).

Changes in text:

We will address the connection between the idealized cases of cascading tipping here and transitions in the real system in the revised introduction en discussion.

In the beginning of section 2 we will add: In this section, we present a mathematical framework for simple cascading transitions, that acts as a first step towards analysing the more complex transitions happening in reality.

2. Comment of the referee:

In bifurcations involving meta-stable states, such as the double saddle node bifurcation, or bifurcations involving strange attractors (e.g. (Tantet, Lucarini, Lunkeit, & Dijkstra, 2018)), a critical transition occurs through a saddle point, or a strange saddle. In this case, although the saddle set is globally unstable, its stable manifold may be responsible for a slowing down at the vicinity of the saddle, resulting in what also looks like a two step transition. Could you discuss why the cascading bifurcations may or may not be a better candidate to explain the two-steps transitions such as observed during the Eocene-

Oligocene transition?

Authors reply:
In DeConto and Pollard (2003), it is suggested that the atmospheric CO2 concentration influences the existence of an ice sheet on Antarctica, via its effect on the ice-albedo and height-mass balance feedbacks. As the box model by Gildor and Tziperman (2000) contains these feedbacks, and a boxed ocean in which Tigchelaar et al. (2011) found multiple steady states for the meridional overturning circulation, a cascading event (of two bistable systems) could be simulated here (as written on p. 15 line 29 to p. 16 line 2). Of course, the comparison with the Eocene-Oligocene transition as found in proxy records should be made with care, because of the simplicity of the model used in Tigchelaar et al. (2011). In the present manuscript we have added this mainly as an example, but clearly further work is necessary to substantiate the hypothesis of cascading tipping being relevant for the Eocene-Oligocene transition. In particular, the coupling of the two bistable systems via the carbon cycle (determining the atmospheric CO2) requires more attention. This goes beyond the scope of the present manuscript and will be elaborated on in a follow-up study.

Changes in text:
We will cite and shortly discuss the Tantet et al. (2018) paper. In the revised discussion, we will add: Although from a physical perspective, this is a potential example of a cascading transition, we make no claim about whether such a transition likely occurred at the Eocene-Oligocene transition.

3. Comment of the referee:
Section 4: In Fig. 7, there is indeed a strong correlation between the temperature gradient and the wind stress. However, as the author remark, there is also a strong spread, which should result in a strong variance in the estimate of the coefficients in Eq. 21. Could you use an ensemble method such as bootstrapping or a Bayesian model to test the probability that such a cascading tipping indeed occurs when sampling the different values of the coefficients of Eq. 21? This would allow to discuss the robustness of the results to the dependence of the wind stress on the temperature gradient.

Authors reply:
This is a good suggestion. Note that the results shown in Fig. 8 and 9 are dependent on multiple parameters and choices made (not only the ones that are derived from Fig. 7). To be precise, these are the definition of the North Atlantic and the Equatorial Atlantic regions, the zonal wind stress region, the reference wind stress parameter (τ_0), and it also turns out that the temporal resolution and running mean may dramatically change the values of α_τ and γ_τ in Eqn. (21).

Changes in text:
We will add such results (using bootstrapping with different values of α_τ and γ_τ in Eqn. (21)) to the revised paper.

4. Comment of the referee:

You explain well how the parameters of the wind stress equation are estimated from the model runs. However, it is not clear to me how the parameters of the Stommel and of the Timmermann are chosen. Are the parameter values used the same as in the references? Are they chosen so as to be as close as possible to historical data? So has to reproduce the mean state and variability found in observations? Or so as to favor the occurrence of the cascading tipping? In any case, I understand that estimating the parameter values of minimal models from observations or complex models is a difficult and not always relevant task. However, the sensitivity of the occurrence of the cascading tipping on the parameters of the coupled Stommel-Timmermann model should be discussed to better assess the likelihood of such tipping to occur.

Authors reply:
For the Stommel and Timmermann models, we have used the standard values as in the original references, except when stated otherwise (for example in the case of the freshwater forcing in Eqn. 23). The parameter mu (that partly determines the closeness of the ENSO system w.r.t. the Hopf bifurcation) has been chosen to be near critical for the Timmermann model.

Changes in text: We will explicitly mention how the value of mu (in caption of Fig. 9) was determined. On page 14 we will add: In the Stommel and Timmermann models, we use the standard parameter settings, as given in the references, unless stated otherwise.

5. Comment of the referee:
Discussion: Salinity biases, such as found in the GCM used in this study, have shown to have a strong impact on the bi-stability of the AMOC (Mecking, Drijfhout, Jackson, & Andrews, 2017). Considering that the strengthening of ENSO also occurs in the control run, could you discuss whether this is/is not an important factor to take into account when asking whether or not such a cascading tipping of the AMOC+ENSO system could occur in the future.

Authors reply:
Whether a cascading tipping event is what actually occurred in the HOSING-1 runs, is not known. Probably there is a more complicated reason behind the increased SST variance in HOSING-1 with respect to the standard runs, and likely a mix of different effects. The effect of salinity biases on the bimodality and hence on the AMOC-ENSO coupling is interesting but outside the scope of this paper. In the coupled Stommel-Timmermann model, we know in which parameter regime of the freshwater flux there is bimodality in the AMOC because that follows directly from the Stommel models design.

Changes in text:
In the revised discussion, we will shortly mention the effect of salinity biases on the bimodal behavior of the AMOC in GCMs, and its potential effect on the cascading behavior (and cite the relevant papers).

Response to referee #2

We thank the referee for the careful reading and the useful comments and will adapt the manuscript accordingly. Below is a point by point reply with the referees comments in bold font, our reply in italic font and the changes in the manuscript in normal font. We have taken the liberty to divide the comments into smaller pieces and discuss them separately.

1. Comment of the referee:

It is not very clear to me what the authors see as the main aim of the paper. In the abstract it is stated that they aim at providing a new theory / a mathematical framework. I am not convinced that these specific claims are supported by the contents of the paper. For example, isnt a theory something that provides an explanation for a certain number of facts? What is the new explanation here, and of what?

Authors reply:

The aim of the paper is introducing the concept of cascading transitions. The contents of the paper support this by giving mathematical examples of these events, describing their statistics, and giving an example from climate physics. We agree that it is only a framework, not a complete theory (which was never the aim). In that respect, the first sentence of the abstract and the first sentence of Section 4 need to be rephrased, as they indeed imply that we provide a new theory.

Changes in text:

Abstract, we will rewrite the first sentence to We introduce a framework of Section 4, in the first sentence we will change theory to concept.

2. Comment of the referee:

I like how Sect. 2 systematically explores conceptual models of two combinations of generic bifurcations. In this section, I have the impression that one tipping point immediately triggers the next (instead of the tipping point in the leading system only bringing the following system closer to its own tipping point, which is then again triggered by the changing control parameter)?

Authors reply:

We understand the confusion and will try to make it more clear in the text. First of all note that the bifurcation diagrams only show equilibrium states of purely deterministic systems. This means that if the bifurcation diagram does not show that the following system tips after the leading system is forced with phi, then the following system will never tip (there is no noise). So actually, one would expect that the bifurcations of the leading and following system overlap (when only looking in phi space), but not in time. The issue why the forward runs in Fig. 2 do show a gap between the two tipping points can be understood as follows. This is purely the transient as it takes a little time for the dynamical system to adapt to the new parameter setting (how long this takes may also depend on the specific form of the coupling). Fig. 2 nicely shows that the two tipping points are different events. The dashed lines show when the system is

still stable in their old equilibrium, and the solid lines show whenever they are drawn to the second equilibrium. For example, in Fig. 2a one can see that the following system (red) only reaches its critical threshold to be drawn to the next equilibrium exactly when the leading system tips (so not before due to the increased forcing, but it is caused by the first tipping).

Changes in text:
We will provide more explanation on the results of Fig. 2 in the text.

3. Comment of the referee:
If this is the case, how can early warning signals even be used to predict the second transition? I will elaborate on this point below. In general, the early-warnings analysis in Sect. 3 is less clear to me than Sect. 2. I have the impression that the authors present two analyses of a single tipping, and not one analysis of an induced tipping, which somewhat questions the novelty of the approach. At first, I thought that the authors aim to predict the second tipping before the first, or infer what kind of bifurcation to expect. However, after the first examples of cascading tipping it seemed like the approach was to use early warning signals to first predict the first transition and after that predict the second transition, but to do that the concept of cascading tipping is not necessary, since they basically predict two tipping points independent from each other.

Authors reply:
We agree with the referee that what is discussed in section 3 is broader than only the prediction of cascading transitions. The goal of section 3 is to give an overview of some important metrics during a cascading transition event, to see whether we can find early warnings of the complete (cascading) event, or to see whether we can find signals that a first (already happened) transition brings another system closer to second transition. So the question was indeed not narrowed to (only) predicting the second tipping before the first. Basically we tried to answer three questions: Can we predict the complete event (second tipping before the first)? For this we looked at DCCA and ρ_{DCCA}. Can we diagnostically see whether the second tipping is caused by the first? For this we looked at the ratios of AR1, variance and DFA before and after the first tipping. How do standard statistical metrics act during the event? For this we looked at AR1, variance and DFA.

Changes in text:
We will add additional text to the beginning of section 3 to explain better what the goal of the section is. We will also rephrase page 2, line 32. We will discuss the limits of the DCCA and ρ_{DCCA} usage.

4. Comment of the referee:
In this context, I was also wondering why the external shock that the second system receives must result from a bifurcation in the leading system. Could it not also result from other kinds of tipping points, or a sudden step or peak in forcing like a volcanic eruption or a pulse release of greenhouse gases? Why is the leading system needed at

all when the main aim is to detect if the first shock will trigger a transition in the following system?

Authors reply:
Of course, relating two transitions to each other is not trivial. It might indeed be that a second transition is completely unrelated to the first. If we would like to check the relation between a second transition to a first, we propose to look at the ratio in the autocorrelation, variance and DFA, because apparently (see Tab. 3-4) these are significantly different from events where there is no second transition or when they are unrelated.

Changes in text:
We will clarify this together with the changes according to the comment 3 above.

5. Comment of the referee:
The model example in the end (ENSO-AMOC model) is interesting, but its purpose is not clear enough to me. Maybe the authors can clarify what it is that they want to demonstrate exactly and state this clearly in the introduction and draw conclusions using the results they show.

Authors reply:
Section 4 is an example of a cascading transition in a physical model. We do not claim that this experiment reproduces reality, but it does reflect that cascading transitions are not a pure mathematical construct, but that they present inside idealized climate models.

Changes in text:
We will elaborate on the goal of section 4 at the beginning of the section and discuss its implications.

6. Comment of the referee:
The conclusion section should be extended by a discussion about what questions are answered and what the implications of the results are. What can we do or understand with the approach in this paper that we were not able to do or understand before? What should be done next?

Authors reply:
Agreed, the implications can be more explicit. The main new notion is that tipping in one system can lead to tipping in another system through modification of the stability of the state of the latter system, even if both systems are only weakly coupled. An important implication of this is that several climate subsystems may actually be highly vulnerable through their coupling with other subsystems.

Changes in text:
We will add a paragraph in the revised discussion section to reflect on the main implications of the results in the paper.

7. Comment of the referee:

The reasons for the choice of methods should be explained better. This is often linked to the problem mentioned above, i.e. the lack of clarity about the aim of the study. Once this aim becomes clearer, it should also be easier to explain why certain methods are applied. Specifically, the choice of statistical indicators needs better justification. I currently do not see what the early warnings approach can add to previous studies. For example, why is DFA used as a warning signal instead of just the autocorrelation? Since autocorrelation is simpler to calculate and more intuitive, I would like to see an argument for the added value of DFA.

Authors reply:
The referee makes a fair point that DFA does not add much in the simple systems we look at right now. As we can see in Fig. 3-5 and Tab. 3-4, AR1 and variance mark the slowing down pretty well, even better than DFA in terms of robustness and statistical significance. DFA is argued to be needed when one needs to filter long-range correlations/non-stationarity in data that has a relatively short size with respect to short-range noise, e.g. in Greenland ice core data in Livina & Lenton (2007). For completeness, we added DFA because when one applies these ideas to actual data, DFA might be necessary.

Changes in text:
We will motivate better why we use DFA, in line with the changes from comment 3.

8. Comment of the referee:
The statement that standard quantities not always provide an early warning signal (page 6, line 26/27) should be backed up with an argument and references, and then it should be explained why DFA can cope with this.

Authors reply:
Agreed. Standard quantities is ill-defined in this sentence. As discussed in comment 3 and 7, the DFA argumentation is also elaborated on more.

Changes in text:
The statement will be changed into standard metrics like autocorrelation at lag 1 and variance do not always provide an early warning signal (e.g. in Greenland ice core data in Livina & Lenton 2007).

9. Comment of the referee:
I would actually expect DFA to fail whenever autocorrelation fails, which happens when the system is more complicated than the typical Langevin equation / AR1-process with one fixed time scale. One argument the authors give is that DFA copes well with non-stationarity. First: What is the explanation for this statement? What is the tradeoff when using DFA (more data needed?). Second: Couldnt one just remove non-stationarity with a high-pass filter (which is what the authors seem to do already) and then use traditional early warning signals? The authors use relatively simple models here, where

the parameter can be varied as slowly as necessary to remove non-stationarity (or they could even make long stationary time series for different fixed parameter values). Another argument the authors provide is that DFA captures long-range correlations. But why should one expect such long-range correlations in the simple models the authors use? Can they even exist? So, in a nutshell, why is DFA needed in this paper?

Authors reply:
The argumentation for the usage of DFA is found in (among others) in Livina & Lenton 2007 and Peng 1994. Livina & Lenton (2007) argue that DFA filters long-range correlations or non-stationarity better. They apply it on a dataset of Greenland ice cores and argue why degenerate fingerprinting is less applicable (due to the short length of the dataset with respect to the time scale of the non-stationarity).

Changes in text:
We already refer to these articles in the paper, but in line with comments 3, 7 and 8, we will explain this better.

10. Comment of the referee:
Then the authors generalise DFA to capture the involvement of several state variables (using DCCA). This could make sense if they were trying to detect something about the coupled system, for instance, which variable is leading, what will happen after tipping 1 and 2. However, the main results seem to consist in predicting tipping 1, and then detecting that the following system has moved closer to a tipping point (by the way, how do we know that there is a second tipping point? The fluctuations could just have changed for another reason.). As far as I can see, DCCA is not needed to do so, an AR1-analysis of each single variable may have sufficed.

Authors reply:
As stated in comment 3, we look at multiple statistical aspects of cascading transitions. DCCA is used to see whether the detrended cross-correlation gives any signal of the complete system prior to a cascading transition, which lets it act as an early warning signal. We agree that this can be made more clear, and that DCCA is not adding much in terms of results. However, we think that it does add to the completeness of the statistical description and invites the scientific community to do further research on this.

Changes in text:
We will add some text in 3.2 to explain why DCCA is included.

11. Comment of the referee:
The explanation on page 7, lines 26-30 is unclear to me. In what way and to what purpose and why can Pearsons correlation not be used? And what is meant with a one-to-one-relationship (line 30)?

Authors reply:

Not much is predictable about the behaviour of the following system prior to the first transition, as it might be still far away from its bifurcation point. The pure (Pearsons) cross-correlation might therefore be very noisy or even decreasing prior to the first tipping. A one-to-one relationship refers to that the systems correlate well, but this is clearly not necessarily the case. However, we suggest that by looking at long-range correlations (e.g. using DCCA), one might filter out an increase in the long-range cross-correlation prior to the first transition, acting like an early warning signal for cascading transitions. The results are too noisy to interpret, but we invite other researchers to look at this more in-depth.

Changes in text:
The first part of 3.2 will be adapted as in the reply to comment 10.

12. Comment of the referee:
Sect. 3.3.1: The authors state several times that DXA and DCCA are sensitive to the segment size and moving window size, but have different values been tried? It would be nice to show how sensitive they are, and what this means for the results. It could help already to just show more runs with different parameter settings.

Authors reply:
Indeed, we have tried different values. The results are indeed quite sensitive to segment size (for detrending in DFA/DCCA) and moving window (for running averages).

Changes in text:
We will a short discussion on the effect of segment size on the results.

13. Comment of the referee:
In section 3.1 The essential part of degenerate fingerprinting is the projection on the leading EOF in a multivariate system. However, the manuscript skips this part of the method, and therefore, right now, just explains the lag-1 autocorrelation and not degenerate fingerprinting. Could one learn something about a system with cascading tipping points by using degenerate fingerprinting on the multivariate signal?

Authors reply:
We agree with the referee. We only focus on the proposed c-propagator as in Livina & Lenton 2007, without explaining the background of degenerate fingerprinting. The c-propagator is reflected by the AR1 coefficient and therefore is an estimator of the degenerate fingerprinting technique.

Changes in text:
In section 3.1, we will add a few lines on the decay rate of κ and the projection on the leading mode.

14. Comment of the referee:
Sect. 3.3: Why are only the double-fold and fold-Hopf systems tested for the early-warning approach, and not the two systems with a Hopf bifurcation in the leading system? This choice should be explained

or the other two examples should be included as well.

Authors reply:
We agree that this seems arbitrary. We focus on those two systems because adding more would be repetitive and not adding much, and the choice of these two is because they both start with a clear first transition, making them more illustrative.

Changes in text:
At the beginning of 3.3, we will add additional text to explain why we focus on the double-fold and fold-Hopf cases.

15. Comment of the referee:
Sect. 3.3.2, page 10, last paragraph: The oscillation seems to affect the measurement of autocorrelation. I think that one should here measure the auto-correlation of the residuals around a mean oscillation, either by subtracting this mean cycle somehow, or by defining a period and working with Poincare sections (snapshots after each period). Otherwise the result would probably be meaningless.

Authors reply:
This is a good suggestion. We will subtract the mean oscillation.

Changes in text:
We will change Fig. 4 and its discussion by removing the mean oscillation.

16. Comment of the referee:
In both figure 1 and figure 2, the choice for the coupling of the two subsystems seems to be arbitrary. These choices could be explained better to make it more understandable for the reader. For example, one can shift the two systems versus each other (by varying parameter gamma1), such that the two tippings are well separated, or that they are really intertwined (one tipping inducing the other immediately). How would the stability landscape then look like, and what would we see in early-warning signals? I was also wondering why the values of gamma1 have been chosen in a way that gamma1 is 0 for the double-fold, ¡0 for the Fold-Hopf, and double-Hopf, and ¿0 for the Hopf-fold. Conceptually it would make a difference if the second tipping is triggered by the changing parameter or a direct consequence of the first tipping. It seems that the latter is always the case here for all parameter choices? This should be made clear from the beginning (as I mentioned above).

Authors reply:
For the discussion on the separation of the transitions, we refer to our answer to comment 2. The parameter settings are chosen such that a cascading event is produced. As specified in Tab. 1-2, there are boundaries within which one can vary the parameters (for some values, there will not be a second tipping). Within these boundaries, one could vary γ_1 to make the transient of the following system progress slower or faster towards the new equilibrium, but the stability

landscape would remain the same as in Fig. 1 (note that this landscape has phi on the horizontal axis). And again; the second tipping is triggered by the first tipping, and therefore indirectly by the forcing parameter ϕ.

Changes in text:
None.

17. Comment of the referee:
Probably related to this point: In figure 2 it seems to be the case that the leading system tips before the following system, whereas the bifurcation plots seem to indicate that this happens at the same time. Where does this time delay come from?

Authors reply:
See our reply on comment 2 and 16.

Changes in text:
None additional.

18. Comment of the referee:
Similarly, a time delay can be seen in the Fold-Hopf system (Fig. 2b), while Fig. 1b would make me expect a discontinuous jump from a stationary solution to a cycle with some non-zero amplitude.

Authors reply:
It takes some time before the cycle arises. Strictly speaking, the following system in 2b would actually remain near the unstable equilibrium for a long time (if there would be no numerical imperfections), as the system is deterministic. The latter can, however, be stated more clearly. In any case, there would be no jump to a non-zero amplitude cycle in the transient. This would always be gradual.

Changes in text:
We will add a sentence to say that in a pure deterministic case, the following system in the Fold-Hopf case would remain near the unstable equilibrium for a long time.

19. Comment of the referee:
Also, according to Table 2, the control parameter Phi increases linearly with time, but I do not see any change in state (or the amplitude of its oscillations) in Fig. 2, and on page 5, last paragraph, it is mentioned that at some point the amplitude would jump to a large value when both equilibria are accessed, but this is not seen in the Figure.

Authors reply:
Assuming the referee is talking about the Hopf-Fold, a change in state in the leading system is visible in 2c in the fact that the black/grey lines start oscillating as soon as ϕ is large enough (i.e., when the lines go from dashed to solid, reaching the critical value). In the following system, the change in state is also visible in 2c, by observing the sharp increase in the red curve around $t = 130$.

Changes in text:
None.

20. Comment of the referee:
In this context, Fig. 1 and 2 appear contradictory to me. This point is actually a crucial one because the period between the two tippings is used to detect early warning signals for the second tipping. How can it even be that there is enough time to detect them, when the system is already in the process of tipping? Here it looks like the second tipping is actually not caused directly by the first, but by the changing control parameter (in contrast to the impression I got in the previous section).

Authors reply:
The second tipping is caused by the first, and the first is caused by the changing control parameter. You can directly see this in the equations, where the leading systems state variables act directly as bifurcation parameters for the following system. An example is the double fold case (Eqn. 13), where gamma modulates the (bi-)stability of the following system, and is dependent on X, not (directly) on ϕ. However, the equilibrium of X is determined by ϕ, which is why for a varying ϕ, X might transition towards a completely other state, which affects $\gamma(X)$ so dramatically that it might affect the stability of Y. To summarise: ϕ indirectly affects the transition in the following system, but always through the transition in the leading system. Concerning the detection time: Fig. 1 and 2 are for deterministic systems and therefore not really about detection, as for detection, a lot of information is gained from its behaviour around its equilibrium (in terms of noise and recovery from perturbations). Also note that in deterministic systems, cascading transition per definition does not have a state between the two tipping events, where the following system is stable (as in that case, the following system would remain stable; there is no noise to change that). Early warnings are analysed in section 3 (Fig. 3-5), where we look at stochastic systems. In contrast with the systems in Fig. 1-2, in section 3, there is time between the first and second tipping (called the following transition period in this paper) where the following system is stable, but close to its bifurcation point (e.g., resulting in high AR1).

Changes in text:
We will add additional text to the beginning of section 2 to emphasise the fact that the systems presented there are deterministic and explain what we replied to this comment.

21. Comment of the referee:
If it is a real cascade (tipping 2 directly induced by tipping 1), wouldnt the systems state suddenly be very far from equilibrium after tipping 1. Can early warnings even be expected in this situation (mind they sample the equilibrium when the state fluctuates around it)?

Authors reply:
The idea here is that there are two different systems; a leading system and a

following system. The first tipping alters background conditions such that the following system comes closer to its bifurcation point. It might be that the referee means that the cascade might not be a consequence of crossing bifurcation points, but rather a cascade as a consequence of a strong perturbation, which is unrelated to changing equilibria or stability in the leading system. That would indeed bring the leading systems state far away from its equilibrium, and depending on its recovery rate, might bring it back to the pre-existing stable equilibrium. During this phase, it might also affect the following system, if that system is coupled in the right way to the leading system. Perturbations are (without any pre-knowledge on the source of the perturbations) unpredictable, but the time between the first and second tipping might give warning indicators about a second tipping.

Changes in text:
None.

22. Comment of the referee:
Moreover, I imagine that the relative time scale between the systems matters (controlled by the different coefficients in the equations). For example, in case of the Hopf-fold system, it would matter how fast and how large the oscillation in the leading system is compared to the following systems response time.

Authors reply:
The relative time scale between the systems indeed is a factor to take into account. For now, we used relatively equal time scales. Of course, if the following system has a very short time scale, and the leading system a rather long time scale, the following system might still transit to a new state as a consequence of the leadings transition, but is hard to find any time in between the transitions, and in real data it might be hard to distinguish the sequence.

Changes in text:
None.

23. Comment of the referee:
So why has this particular coupling been chosen for the paper, and how representative is that for the climate system?

Authors reply:
The linear coupling is based on the thermal wind balance, where the wind stress adjusts to the changes in meridional temperature gradient. So the coupling formulation has a physical basis.

Changes in text:
We will motivate the coupling better in the revised section 4.

24. Comment of the referee:
The climate model (coupled ENSO and AMOC) seems very interesting. However, it is not completely clear to me what point exactly the authors want to make by showing it. Sect. 4.3 is very short and I

dont really understand its purpose. In the conclusion section and in the abstract it seems to be argued that it illustrates that cascading tipping can occur in climate models, but as this is already known according to the introduction, and given that the model has been designed like this on purpose, what new information does this model provide?

Authors reply:
Agreed; this is indeed less clear from the text.

Changes in text:
We will add a short motivation in the beginning of section 4 why this model is chosen.

25. Comment of the referee:
Also, it should be more clearly explained how the two existing models have been coupled. I found it difficult at first to identify the common variables in the models that were linked. More precise wording might help (e.g. through influence of the wind stress - influence on what?, in the original model - which model?). It seems that the authors introduce an equation for the wind stress τ which links τ from the ENSO model to the temperatures from Stommels model? Then one could say so from the start, followed by the details.

Authors reply:
The sentence containing through influence of the wind stress is indeed unclear. Section 4.2 gives a technical overview of the complete coupled model. The coupling itself is part of that and the introduction and explanation of the coupling appears sufficient.

Changes in text:
We will correct the mentioned sentence.

26. Comment of the referee:
The model seems to be a representation of the Fold-Hopf case above? This should be explicitly stated from the beginning.

Authors reply:
We agree.

Changes in text:
At the beginning of section 4, we will add that this is an example of the Fold-Hopf case.

27. Comment of the referee:
- Why have the authors not done an early-warning analysis with this AMOC-ENSO model? This would be a natural step to do after the generic models above.

Authors reply:

This is indeed something we have thought about. The largest problem is that the model is deterministic. Some stochasticity would be needed to be build into the model, which would require a thorough sensitivity analysis and needs to be done with care as the pre-existing models were not designed for that. This is beyond the scope of this paper, if even possible. Moreover, it should be noted that this is not the aim of section 4, which is illustrating an application of cascading transitions in physics, implying the possibility of these events in physical systems.

Changes in text:
None.

28. Comment of the referee:
The authors use data from a complex model to tune their conceptual model. What can we learn from that data directly about predicting each tipping, or the coupling (or whatever the authors aim to do)? Could one apply a statistical analysis and infer something about that model from the data?

Authors reply:
We only use the complex model data to back up the ideas we have for (a) the coupling (relation between τ_{ext} and meridional temperature) and (b) what happens to El-Niño when the AMOC collapses. Trying to infer predictions/early warning elements in this data has a number of problems. A first problem is that the first transition is caused by a large perturbation, and no critical slowing down has happened prior to that. So the period before the first transition is useless in the scope of cascading tipping prediction. The second problem is that we are unsure (and we also do not claim otherwise in the paper), whether a (stochastic) Hopf bifurcation can be found or is crossed in Pacific Equatorial SSTs in the hosing runs, like in our simplified model. More research is needed for that.

Changes in text:
We will add a few lines on the limitations of the data from the complex model in the revised section 4.

29. Comment of the referee:
Minor comments What I find most interesting is the analysis of the coupled deterministic systems, e.g. in Fig. 1. A very interesting aspect is the occurrence of intermediate (in terms of the state variable) stable states which are inaccessible when varying the control parameter. It seems that only noise can bring the system on these branches. This aspect is however not discussed in the paper. It is of course up to the authors if they want to go into this, but I would recommend them to at least comment on these hidden states, which I personally find more novel and exciting than the early-warning part of the paper. Could there be such hidden stable states in the climate system and how can they be found?

Authors reply:
Thank you for mentioning this. The occurrence of these in-between states (called

following transition period, FTP in section 3) in stochastic systems is indeed interesting.

Changes in text:
We will add a few sentences about the FTP, explaining it more elaborately, in section 3.3.1.

30. Comment of the referee:
Section 2.1 + Figure 1: It took me quite some time to understand what is going on. It could be helpful to create an X-Y bifurcation plot in addition to the phi-X plots and phi-Y plots that are shown already (to see how each system behaves in isolation). More emphasis can be put on explaining this figure, because this in itself is already an interesting result. The authors might even think of making an animation as extra material, to show how the subplots relate to each other. Also, it could be nice to show how figure 2 relates to figure 1.

Authors reply:
This is indeed confusing, see also the reply on comment 2.

Changes in text:
We will add a sentence in the beginning of section 2, explaining Fig. 1 in more detail. We will also add extra panels in Fig. 1 to see the bifurcation diagram of the following system with respect to the coupling and to emphasize the separation between the systems .

31. Comment of the referee:
- Several different names are sometimes used for the same thing, at least for the fold bifurcation (fold / back-to-back / back-to-back saddle-node). I had never come across the term back-to-back before. Is one term a subset of another? The authors should clarify this and unify the language.

Authors reply:
We indeed used all those terms interchangeably. The back-to-back saddle-node consists of two saddle-nodes at ends of two stable equilibrium branches, connected by a common unstable equilibrium branch.

Changes in text:
We will explain the prefix back-to-back in section 2 and use only the terms saddle-node bifurcation and fold in the revised paper.

32. Comment of the referee:
- Some of the references are a bit outdated (e.g. Kutzbach 1996 on page 2; a lot has happened since then), or could be a bit more specific. Page 2: Scheffer 2009 is a review of some of the earlier papers like Held 2004, some of which are cited later; Peng 1994 is not about predicting tipping points. Also, note that there are papers from the 80ies dealing with statistical precursors already, e.g. by Wiesenfeld, 1984.

Authors reply:
Although perhaps outdated, Kutzbach 1996 is illustrative for what is said in our paper about the desertification of the Sahel region. We agree that Scheffer 2009 is a review paper, but it combines various simple metrics together. Held 2004 is cited multiple times when we specifically mention degenerate fingerprinting, about which Scheffer 2009 does not go into detail. Peng 1994 lays the foundation of DFA to detect long-range signals, and therefore should (only) be cited when we talk about DFA, which we do. Wiesenfeld 1984 is specifically about period-doublings.

Changes in text:
None.

33. Comment of the referee:
page 1 (lines 17ff): Lenton et al. 2008 do not show evidence that there are tipping points in the climate system (though the paper is often cited in that way), so this paragraph should be formulated more cautiously.

Authors reply:
We agree.

Changes in text:
Part of the sentence will be changed to Lenton et al. (2008) give an overview of these.

34. Comment of the referee:
Also, the vegetation states found by Hirota et al. are purely ecological phenomena, and do not imply any tipping points in the climate.

Authors reply:
Hirota et al. 2011 show various equilibrium states of tree cover as modulated by precipitation, and also discusses transitions between these states, and related hysteresis effects and bifurcations. Although the paper focuses on the interaction between vegetation and precipitation, the results are illustrative of tipping elements in (a subsystem of) the climate system.

Changes in text:
None.

35. Comment of the referee:
- page 6, line 18: close to critical transition (2x), should be close to a critical transition.

Authors reply:
Agreed.

Changes in text:
This will be changed accordingly.

36. Comment of the referee:
- page 10, line 17/18: as it is no critical transition: why not? And what is a critical transition?

Authors reply:
Agreed that this needs to be rephrased.

Changes in text:
The sentence will be rewritten.

37. Comment of the referee:
- In Fig. 8, I found it confusing that the labels are not next to the vertical axes but inside the figure. I do understand that this is consistent with the previous figures, so I dont have strong feelings about this.

Authors reply:
Understandable, and a matter of choice. As we have had no other comments about this, we will leave it as it is.

Changes in text:
None.