Early Warning Signals of Environmental Tipping Points

Submitted by Christopher Andrew Boulton, to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy in Geography, June 2015.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: ……

Abstract

This thesis examines how early warning signals perform when tested on climate systems thought to exhibit future tipping point behaviour. A tipping point in a dynamical system is a large and sudden change to the state of the system, usually caused by changes in external forcing. This is due to the state the system occupies becoming unstable, causing the system to settle to a new stable state. In many cases, there is a degree of irreversibility once the tipping point has been passed, preventing the system from reverting back to its original state without a large reversal in forcing. Passing tipping points in climate systems, such as the Amazon rainforest or the Atlantic Meridional Overturning Circulation, is particularly dangerous as the effects of this will be globally felt. Fortunately there is potential for early warning signals, designed to warn that the system is approaching a tipping point. Generally, these early warning signals are based on analysis of the time series of the system, such as searching for 'critical slowing down', usually estimated by an increasing lag-1 autocorrelation (AR(1)). The idea here is that as a system's state becomes less stable, it will start to react more sluggishly to short term perturbations. While early warning signals have been tested extensively in simple models and on palaeoclimate data, there has been very little research into how these behave in complex models and observed data. Here, early warning signals are tested on climate systems that show tipping point behaviour in general circulation models. Furthermore, it examines why early warning signals might fail in certain cases and provides prospect for more 'system specific indicators' based on properties of individual tipping elements. The thesis also examines how slowing down in a system might affect ecosystems that are being driven by it.

Acknowledgements

Firstly I want to thank Tim for his patience and understanding when he realised I am a true ‘pure mathematician’ (who needs units?) and for his unlimited supply of ‘pixie dust’ to sprinkle over our publications. Hopefully I’ve come on at least a bit since I started.

Thanks to Ben, David and Glen at the Met Office for giving me the summer placement right before I started my PhD. It gave me a huge leg up in the type of research I would be doing for the next few years. Also thanks for letting me ‘steal’ data! Thanks also to Peter Good for the help he’s given me on certain chapters/publications.

University-wise, thanks to Steve, Tom and Alastair (the Hatherly office bunch) for good times, even if our office did end up being covered in soil at one time, and cheese and broccoli soup another time. I liked all the R challenges I was set! Thanks to Loocy who although isn’t even at university kept me company on Facebook (when I was supposed to be working!) Thanks to Hywel for giving me loads of teaching hours which partially supported me during my continuation year. Also thanks to Peter Cox for his ideas and suggestions and Lesley for her help.

Finally thanks to Mum and Dad for their support. To quote David Beckham. ‘My parents have always been there for me, ever since I was about 7.’

Table of Contents

List of Figures

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Tables

Chapter 4

Authors Declaration

All analysis in this thesis has been carried out by Christopher Andrew Boulton (CAB).

The FAMOUS GCM run of AMOC collapse was run by Hawkins et al. (see Chapter 2 for reference) and obtained by Lesley Allison. Streamfunction data from other GCMs (Fig. 2.1, 2.2) was also obtained by Lesley Allison. Chapter 2 was co-written with Tim Lenton and Lesley Allison.

The HadCM3-ESE used in Chapters 3 and 4 was devised and run by members of the Hadley Centre. CAB ran the RCP 8.5 members of the ensemble (used in Chapter 4) under the guidance of Hadley Centre colleagues. Chapter 3 was co-written with Tim Lenton, with input from Peter Good of the Met Office. Chapter 4 was written with input from Peter Good and Ben Booth of the Met Office.

Chapter 5 was co-written with Tim Lenton.

Publications

Chapters in this thesis are created from the publications listed below. They have been adapted from their published form to allow them to form chapters.

Chapter 2: Boulton, C. A., Allison, L. C., and Lenton, T. M., 'Early warning signals of Atlantic Meridional Overturning Circulation in a fully coupled climate model', 2014, Nature Communications, 5, 5752

Chapter 3: Boulton, C. A., Good, P., and Lenton, T. M., 'Early warning signals of simulated Amazon rainforest dieback', 2013, Theoretical Ecology, 6, 3, 373-384

Chapter 4: Boulton, C. A., Booth, B. B. B., Good, P., 'Exploring uncertainty of Amazon dieback in a perturbed parameter Earth system ensemble', in preparation, previously submitted and reviewed at Global Change Biology

Chapter 5: Boulton, C. A., Lenton, T. M., 'Slowing down of North Pacific climate variability and its implications for abrupt ecosystem change', 2015, PNAS, 112, 37, 11496-11501

Chapter 1

Introduction

Although a relatively new concept in climate and environmental systems, tipping points in dynamical systems have a long research history. These occur when a system is forced such that the stable equilibrium it occupies becomes unstable and the system settles to equilibrium in a different stable state. This can cause drastic changes that occur seemingly randomly and without warning, and so are of huge interest to society, especially from a climate perspective where their impacts could be globally felt. Early warning of these transitions could be useful in prevention or at the very least adaptation to the changes that will occur.

Methods for predicting the approach towards tipping points in dynamical systems are based on analysis of their time series. Systems will begin to respond more sluggishly to perturbations as the state it occupies loses stability and will also sample more of the surrounding state space, quantified by an increase in autocorrelation (Dakos *et al.*, 2008, Held & Kleinen, 2004) and variance (Carpenter & Brock, 2006) respectively over time. These are examples of ‘generic early warning signals’, others of which are detailed later.

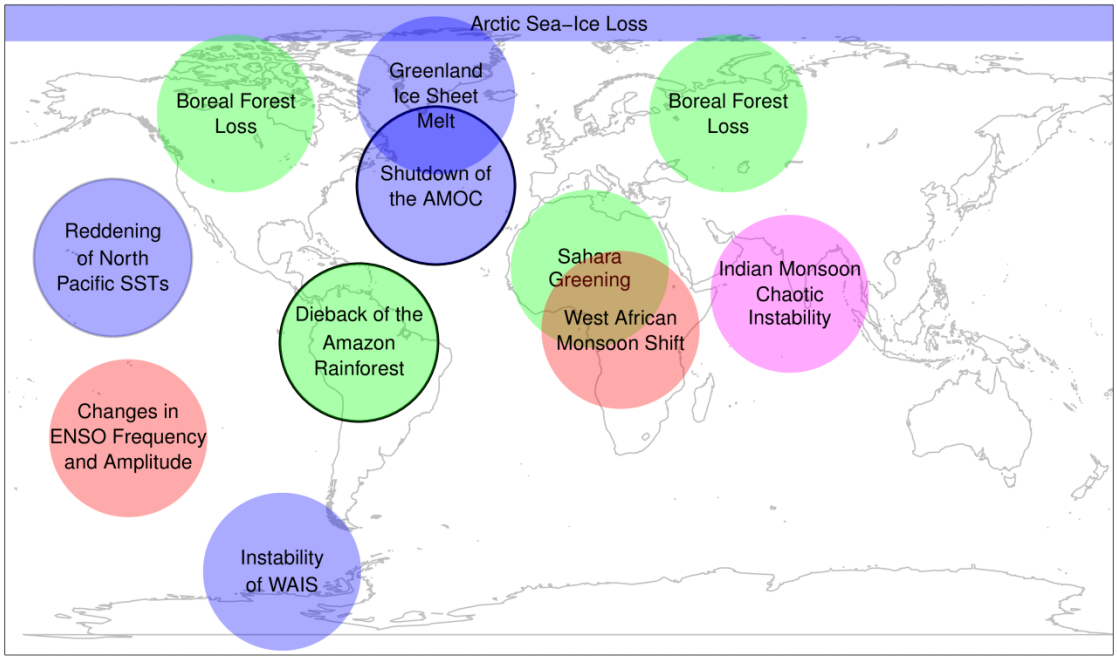

Figure 1.1: A map of climate systems that may potentially exhibit tipping points (adapted from Lenton *et al.* (2008)). Systems coloured green are concerned with changes to vegetation, blue, changes in ocean circulations and the cryosphere and red, changes to atmosphere. The two systems circled in black are explored with the thesis and the system circled in grey ('Reddening of North Pacific SSTs'), although not considered a tipping point, exhibits signals similar to a system approaching a tipping point and is also explored in this thesis.

Tipping points are thought to occur in a number of climate systems (Fig. 1.1), all of which will have impacts which would be at least nationally felt, with the majority having global implications should the tipping points be passed (Lenton *et al.*, 2008). Candidate tipping points include changes in ocean circulation, such as Atlantic Meridional Ocean Circulation (AMOC) collapse, to loss of large-scale vegetation like the Amazon rainforest, among others. The potential for these systems to pass their tipping points depends on how they are forced in

the future, and there is a large uncertainty in future anthropogenic forcing of the climate system.

The central aim of this thesis is to determine the performance of generic early warning indicators of tipping point behaviour in climate systems that could potentially tip in the future, with the use of general circulation models (GCMs). This provides an increase in the complexity of systems these indicators have been tested on, and involves plausible forcing scenarios that could be seen over the 21st century. Furthermore the statistical significance of any indication of an approach towards a tipping point is tested throughout. A method for testing how soon before a tipping point a significant warning (usually an increase in an indicator) and therefore a prediction of an approach can be made is also suggested. In cases where generic early warning signals fail, alternative methods of indication are provided and the reasons for the indicator failing are determined. Lastly, there is an exploration into how the approach towards a tipping point in a climate system might affect ecosystems that are influenced by it. This provides a basis to begin to determine if this change in ecosystem behaviour could be an indicator of tipping point approach itself.

A Generic Tipping Point

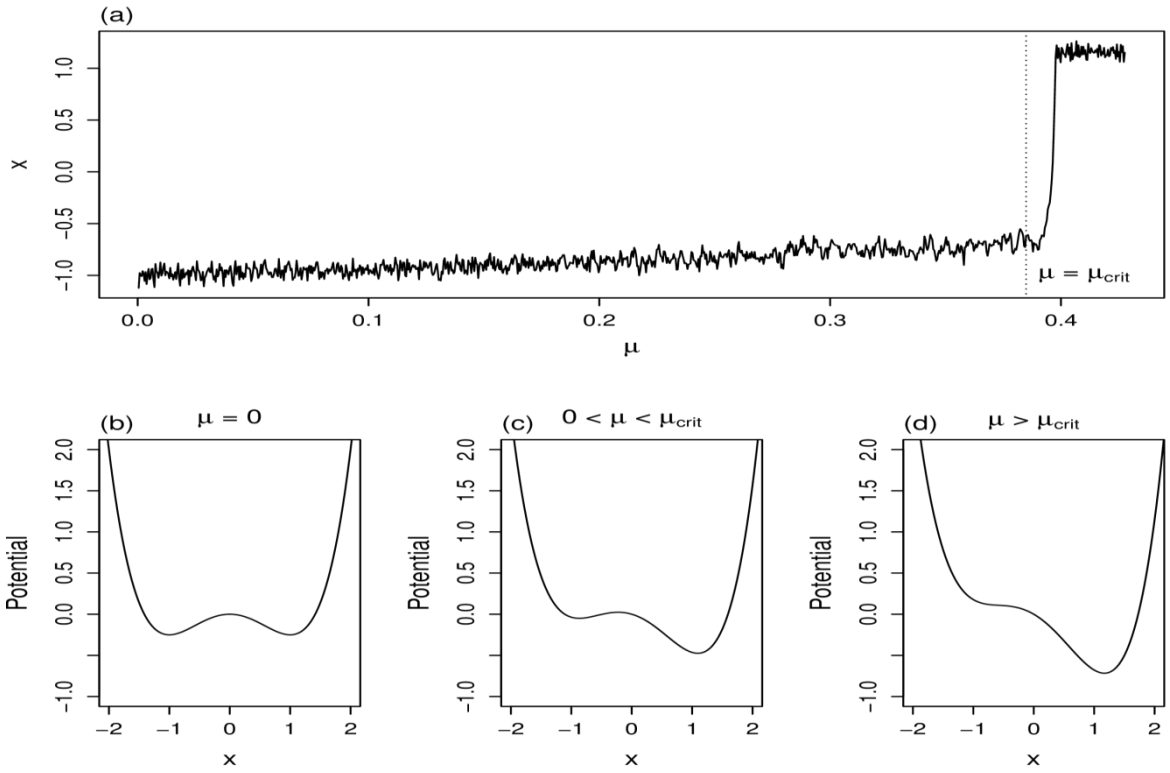

Figure 1.2: An example of a tipping point being passed. Forcing of the system governed by Eqn. 1.1 is provided by increasing parameter μ to a critical value, μ_{crit} (dotted line in (a)), causing time series x to abruptly shift (a). The potential of the system for (b) μ=0, (c) 0<μ<μ_{crit} and (d) μ>μ_{crit} shows how the left hand stable equilibrium loses stability as μ is increased, causing the system to shift to the right hand equilibrium.

To begin to understand what causes tipping (Fig. 1.2a), we imagine a typical system with two stable states. These two states are separated by an unstable steady state. It is easy to visualise this by plotting the potential of the system (Fig. 1.2b-d). The two 'wells' of the system represent the stable states whereas the 'hump' in between them is the unstable steady state. The longer time scale

dynamics of the system are deterministic, whereas the system will be moved around by shorter time scale stochastic noise, which from a climate perspective could be viewed as ‘weather’. The steepness of the sides of the wells are related to the strength of their stability and how strong the restoring feedback, which keeps the system in that state, is. The system in Fig. 1.2 is represented by Eqn. 1.1, with the bifurcation parameter μ determining the distance away from the bifurcation or tipping point.

$$\dot{x} = -x^3 + x + \mu \tag{1.1}$$

Tipping generally occurs in one of two ways. The first of these, via a bifurcation, is when the state the system occupies loses its stability, causing the system to move to the other stable attractor. This bifurcation is approached due to a change in one (or more) of the parameters in the system known generally or collectively as the bifurcation parameter (μ in Eqn 1.1). In this case there is a sense of irreversibility as the system would have to be forced in the opposite direction to obtain the stability of the original state once again. The system could also move to a different state due to being forced purely by the short term stochastic noise/weather without the stability of the states changing. The most likely way for tipping to occur is a combination of these two. As a system is approaching a bifurcation, it becomes easier for noise to induce a shift to another state as the basin of attraction of the current state gets smaller and thus perturbations to the system are more likely to push the system outside of this basin (Fig. 1.2b-d). Tipping points can also occur when rate of change in a parameter is too fast for the system to keep up and so moves away from stable

attractors. This ‘rate dependent’ tipping is not discussed here but information can be found in Ashwin *et al.* (2012).

As the system approaches a bifurcation, the leading eigenvalue of the Jacobian matrix of the system approaches zero from below (Wissel, 1984). As this happens, the basin of attraction of the state decreases and the potential well widens (Fig 1.2b-d). Because of this, the system will begin to respond more sluggishly to noise. This phenomenon is known as ‘critical slowing down’ (CSD) and invokes the idea that in a time series of the system, points will begin to become similar to the previous point in the series. For daily data, this could be simply thought of as ‘today is becoming more like yesterday’.

The changes in the potential well and behaviour of the system can be seen in Figure 1.2, where three examples of the potential well of a system at different distances from the tipping point are shown. The potential well of the state occupied by the system is steep when it is far away from bifurcating (Fig. 1.2b) and shallower near the threshold (Fig. 1.2c). Once the critical value of μ, μ_{crit} has been passed, the potential well the system originally occupied has lost all stability and only the well/state on the right hand side remains (Fig 1.2d). The time series in Fig. 1.2a is created by slowly increasing the bifurcation parameter as well as perturbing the system by white noise of constant variance. The time series of the system when μ is near 0 acts similarly to white noise (Fig. 1.2a) seemingly random much like the perturbations. When the system is closer to tipping, the time series exhibits increased memory with successive points more similar to the previous point. This critical slowing down causes the time series to meander away from the equilibrium for longer times (observed in Fig 1.2a), rather than appearing like white noise.

As well as this slowing down, the basin of attraction of the equilibrium reduces (due to the leading eigenvalue approaching 0) causing the system to sample more of the surrounding area. Close to the tipping point, the variance of the time series is larger than far away from it (Fig. 1.2a). In the typical ‘fold bifurcation’ example, where there are two states with one losing stability, the time series also becomes more skewed closer to the tipping point due to the landscape of the potential well not being symmetric. In other examples such as a ‘pitchfork bifurcation’ where there is one state that symmetrically separates into two, skewness is not expected to change as the basin of attraction of the state will remain symmetric. These characteristics will be discussed in more detail below.

Tipping Points in the Climate System

Potential tipping points in the climate system are detailed by Lenton *et al.* (2008) and some examples can be seen in Fig 1.1. Expert elicitation carried out by Kriegler *et al.* (2009) asks for the opinions of experts on various systems' likelihood of passing a tipping point under 3 different warning scenarios. Across all the experts, there is a range of likelihoods found for the various tipping points with increases in the probability of the tipping point occurring under the higher warming corridors. This range in responses suggests that there is high uncertainty in the future of these systems. The expert's opinions also suggest that a tipping point passing in one system will change the potential for another system to tip with the majority of these interactions promoting further tipping. However in some cases, the experts were unsure on the sign of the influence of tipping in one system on another and in other cases, collectively disagreed on the sign of the influence. This again shows the uncertainty associated with these tipping points. Lenton (2011) suggests that overall, there is a range of likelihoods and impacts of these tipping points occurring. For example, the loss of summer sea ice has a relatively high likelihood with low impact, whereas abrupt changes in ENSO are relatively unlikely to happen, but will have a large impact if they occur. There are three climate systems explored within this thesis, collapse of the Atlantic Meridional Overturning Circulation (AMOC), dieback of the Amazon rainforest and shifts in the Pacific Decadal Oscillation (PDO). The first two are generally considered systems that could exhibit tipping point behaviour, whereas there is debate about whether shifts in the PDO occurred due to some nonlinear dynamics, or a red noise process that takes larger

excursions from a mean state (see below). While the chapters concerning these systems contains detailed background information, an overview, specifically from a tipping point prospect is provided below:

Atlantic Meridional Overturning Circulation (AMOC) collapse

The idea of AMOC exhibiting two stable states, an ‘on’ and an ‘off’ mode, is supported by theory (Stommel, 1961) and model studies (Hawkins *et al.*, 2011, Lenton *et al.*, 2009, Manabe & Stouffer, 1988, Rahmstorf *et al.*, 2005). Original theoretical arguments came from Stommel’s box model (Stommel, 1961), which consisted of 2 well-mixed boxes, representing the high and low latitudes, connected by pipes representing surface flow and deep water flow. Each box has a temperature and a salinity and the flow strength between boxes is regulated by the gradient of temperature and salinity between the boxes. The ‘on’ state (the current state of AMOC in the real world) is regulated by the temperature gradient, where surface flow is poleward, from the low to high latitudes due to the high latitude box being denser (having a higher salinity). However a salinity gradient driven ‘off’ state exists, where the flow is reversed, which exists under different initial conditions of temperature and salinity.

Both equilibria can be found in coupled atmosphere-ocean models (Manabe & Stouffer, 1988), as well as hysteresis exhibited in a number of models (Hawkins *et al.*, 2011, Rahmstorf *et al.*, 2005), with the tipping point between states induced by increasing a freshwater forcing in the high latitudes to alter the salinity in the region.

In its current state, the AMOC is responsible for a large poleward heat transport (Johns *et al.*, 2011). Without this, the North Atlantic region could cool by 1-3°C with local cooling due to large sea-ice changes of up to 8°C (Jacob *et al.*, 2005, Vellinga & Wood, 2002). A southward shift in inter-tropical convergence zone would also be expected, causing drought in the Sahel (Brayshaw *et al.*, 2009, Vellinga & Wood, 2002), and sea level changes of up to 80cm along the coasts of North America and Europe (Levermann *et al.*, 2005, Vellinga & Wood, 2008).

There is large uncertainty associated with the possibility of future AMOC collapse, largely due to differences in the ways it is modelled in GCMs which produce different streamfunctions (Boulton *et al.*, 2014). Freshwater forcing similar to rates observed currently (Lenton *et al.*, 2008) was applied to a GCM (Hawkins *et al.*, 2011) and the AMOC collapsed after 800 years, suggesting that it is far away from tipping in that model, but any early warning would be useful due to the significance consequences if this were to occur. However it is important to note that the initial state used in the model could be very different from the real world and in reality, the tipping point could be closer or further away. (Hawkins *et al.*, 2011) suggest that the FAMOUS run is closer to the tipping point by assessing the 'F_{OV}' indicator in their system, a measure of the net freshwater import into the Atlantic. This determines how strong the negative feedback restoring the AMOC is and is detailed below.

Dieback of the Amazon rainforest

Sudden dieback of the Amazon rainforest was first simulated in an offline vegetation model forced by changes in climate from the GCM HadCM2 (White *et al.*, 1999). Since then it has also been found in coupled GCMs such as

HadCM3LC (Cox *et al.*, 2000), with an even stronger non-linear response to temperature found when the same model is run to equilibrium (Jones *et al.*, 2009). However there is high uncertainty concerning the future of the rainforest, with other GCMs showing less dieback and a transition to a seasonally dry forest in the East Amazon appears a more realistic prospect (Malhi *et al.*, 2009). Experts during an elicitation gave a 20% chance of dieback (in this case over half of area of the Amazon loses year-round forest) by 2200 if global temperature increases between 2-4°C and a 70% of dieback if it increases by more than 4°C.

For tipping point behaviour to be observed, a positive feedback inducing forest loss needs to exist in the system. The rainforest is partially self-sustaining, in that water is recycled to the atmosphere via transpiration and becomes precipitation (Betts, 1999, Salati & Vose, 1984). Winds inland ensure rainforest far from the Atlantic receives this recycled precipitation. If forest were to die back, a positive feedback would be caused by the reduced rainfall and thus further dieback (Betts *et al.*, 2004).

The importance of the Amazon rainforest in future climate change is due to it being a large carbon store and generally a carbon sink, except in recent drought years when it has become a temporarily carbon source (Lewis *et al.*, 2011, Phillips *et al.*, 2009). Dieback of the Amazon rainforest on a regional scale could cause the area to become a permanent carbon source, having a large effect on the global carbon cycle and thus future climate change (an estimated increase of 280ppmv of atmospheric CO_2 when including climate effects on the carbon cycle, resulting in a further contribution of 1.5°C to global warming in HadCM3) (Cox *et al.*, 2004, Cox *et al.*, 2000). Loss of the Amazon rainforest also results

in a loss of habitat for the biodiversity native to the Amazon region (Dirzo & Raven, 2003).

Analyses carried out within this thesis concerns changes to the Amazon rainforest due to changes in climate only, rather than through deforestation or fire effects as these are not implemented in the GCMs used.

Changes in the Pacific Decadal Oscillation (PDO)

The Pacific Decadal Oscillation (PDO) is the time series of the first empirical orthogonal function (EOF1) of North Pacific (20-60°N) sea surface temperatures once the annual cycle and warming trend have been removed (Mantua *et al.*, 1997). It represents the polarity between surface temperatures in the eastern and northern sections of the domain, compared to the western and central sections. An example of the PDO (although using a different dataset to Mantua *et al.* (1997)) can be seen in Figure 1.3.

The PDO index switches sign on a multidecadal time scale and abrupt shifts between the positive and negative phases (Hare & Mantua, 2000) have been linked to changes in salmon productivity (Mantua *et al.*, 1997), drought regimes in the US (McCabe *et al.*, 2004), and Indian summer monsoon rainfall (Krishnan & Sugi, 2003) amongst other biological and ecological time series (Hare & Mantua, 2000).

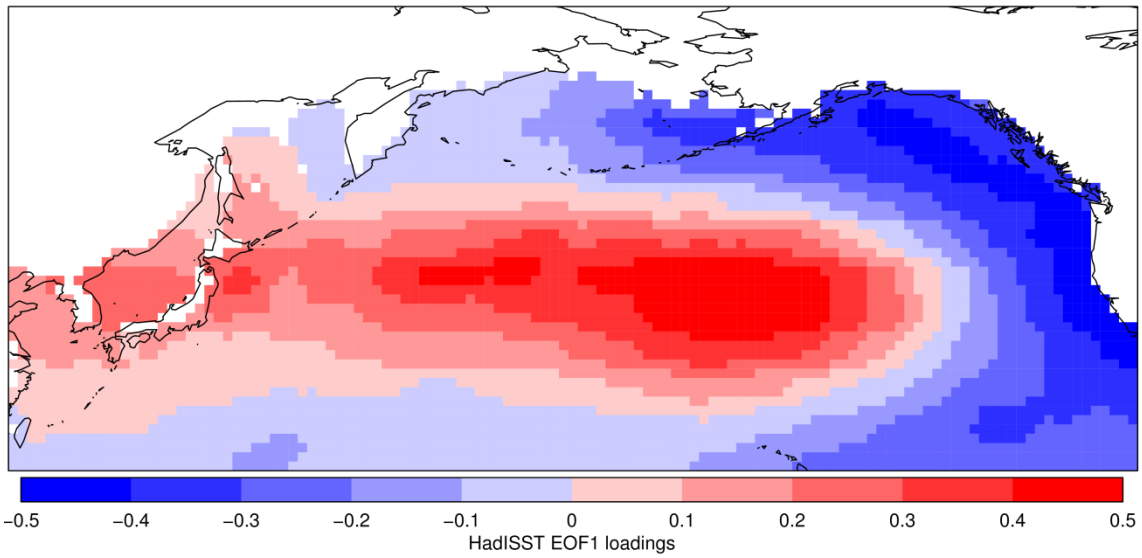

Figure 1.3: A map of the first empirical orthogonal function (EOF1) of HadISST North Pacific SSTs. SSTs of grid points have had their annual cycle and quadratic warming trend removed. Principal component analysis is conducted on the residual time series.

Sea surface temperatures are influenced by the white noise like forcing from the atmosphere which is integrated over time, resulting in red noise type behaviour (Hasselmann, 1976). This has led to debate about whether or not the PDO exhibits true tipping point behaviour with separate attractors, or if it is a red noise process which over time, will meander above or below a mean value (Rudnick & Davis, 2003).

The marine ecosystems driven by indices such as the PDO are thought to integrate them over time, 'double- integrating' the original white noise atmospheric forcing (Di Lorenzo & Ohman, 2013). Whether these systems have more than one stable attractor and could exhibit nonlinear, tipping point behaviour (Hsieh *et al.*, 2005), or will respond linearly to the index (Hsieh & Ohman, 2006), a redder forcing will be more easily tracked by the marine

ecosystems (Steele *et al.*, 1994). While the PDO index is not considered a tipping point in this thesis, it appears to be getting redder over time (producing signals similar to early warning signals) and the potential effect of this on other ecosystems is explored.

Early Warning

The methods used within this thesis generally use the idea of testing early warning indicators on time series of a system, considering a section or 'window' of it at a time. This window then moves along the time series one point at a time to build up a time series of the indicator. A change (usually an increase) in the indicator over time is considered an early warning of the approach towards a tipping point. The default choice of window length in existing literature (e.g. Dakos *et al.* (2008)) appears to be half the length of the time series, noting that only data up to but not including the tipping point should be used. A window length needs to be long enough to provide a reliable estimate of the indicator within that window. It also needs to be short enough to provide a time series of the indicator sufficiently long to determine if there is any trend that would give rise to early warning. In cases where drift occurs (the equilibrium of the system moves), the time series should first be detrended, usually with a smoothing function. Kernal smoothing functions are used for this within this thesis.

The time series of these indicators are tested using Kendall's τ rank correlation coefficient to determine how strong they are.

$$\tau = \frac{(number\ of\ concordant\ pairs) - (number\ of\ discordant\ pairs)}{\frac{1}{2}n(n+1)}$$

Where n is the length of the time series and a pair (x_i,y_i) and (x_{i+1},y_{i+1}) is concordant if $x_i > x_j$ and $y_i > y_j$ or $x_i < x_j$ and $y_i < y_j$ and discordant otherwise. With one variable as time which will always be increasing, Kendall's

τ gives the indicator's tendency with a value of 1 being always increasing and -1 always decreasing.

Figure 1.4 shows an example of calculating an indicator from the system shown previously (Eqn. 1.1). In this case, the bifurcation parameter starts at -0.235 and goes to 0.5. Indicator values are plotted at the end of the window length used to calculate them as shown by the different colours in Fig. 1.4. The indicator time series, when detrending with a bandwidth of 100 and using a window length of 200 shown in Fig. 1.4 has a τ value of 0.9, suggesting it is a strong indication of approaching a tipping point. However this is only one example using a chosen window length and detrending bandwidth. The robustness of this signal can be tested by measuring the strength of the indicator whilst varying detrending bandwidth and window length. The τ values from each calculated indicator can be viewed in a contour plot as is shown in Fig. 1.4. In this instance, the AR(1) signal tested is robust to varying detrending bandwidth and window length. The AR(1) indicator is described in full detail later in this chapter.

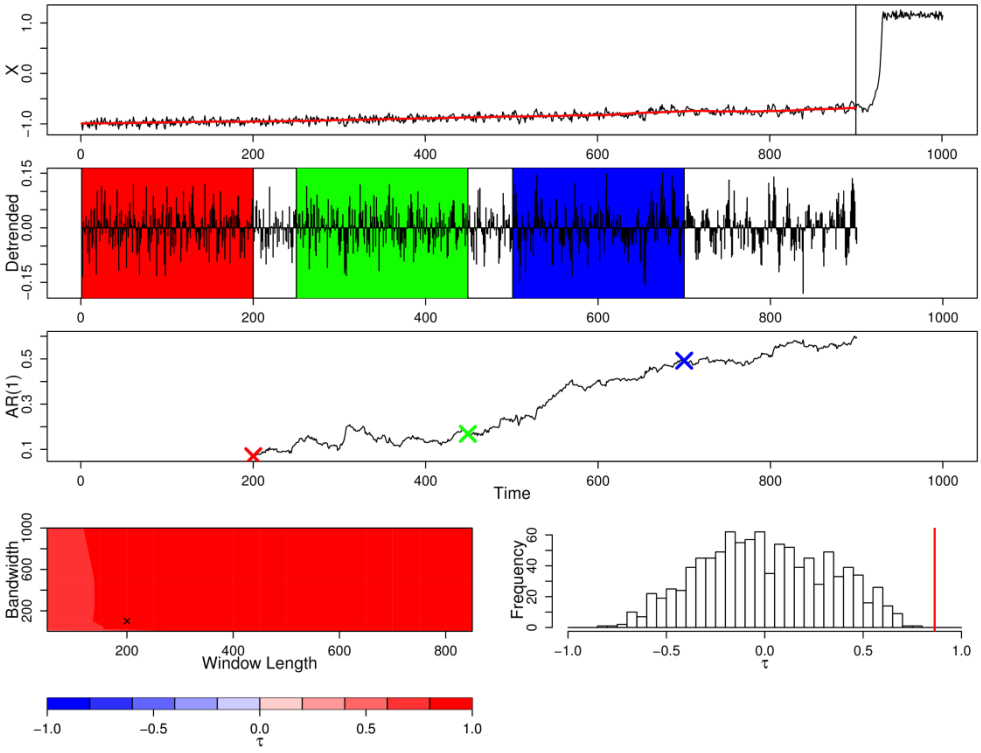

Figure 1.4: An example of an early warning indicator (AR(1); see below) and its tests of robustness and significance. The time series (a) is cut prior to the tipping point (vertical line) and using a Kernal smoother (red line), detrended (b). The early warning indicator is calculated on a moving window and plotted at the end of the window used (c). Examples of the windows and the resulting indicator are shown as different colours in (b) and (c). To test robustness, Kendall's τ values are calculated for the indicator when varying window length and detrending bandwidth and are plotted as a contour plot (d), with a cross denoting the window length (200) and bandwidth (100) chosen as an example in (c). A null model ensemble (created by bootstrapping, see main text) is used to determine the significance of the result by calculating a P-value, the probability an early warning signal or equal or greater τ value is found. In this case all 1000 τ values are lower than the τ found in the indicator (0.9), thus P<0.001.

The significance of the strength of these indicators is measured using a null model. This null model can be created a number of ways. In a created example, the indicators could be tested on the system when there is no approach to the tipping point (i.e. keeping the bifurcation parameter at a constant value). In reality, not enough will be known about the system in question to be able to run a simulated null model. In this case, the time series created for the null model can be created by sampling values from the original time series, creating a number of time series of the same length (Dakos *et al.*, 2008). This is known as bootstrapping. The time series created from this will have similar properties such as mean and variance whilst the memory in the time series will be destroyed. With this memory lost, the signals are not expected to be present and the probability of strong signals being found by chance can be explored. From null models, a P-value can be calculated. This is the probability of finding a stronger signal by chance. An example of a null model ensemble calculated for the example system (using the bootstrapping technique) is shown in Fig. 1.4, showing that the signal found is highly significant with P-value < 0.001.

Examples of Generic Early Warning Signals

Critical slowing down

As mentioned in the introduction, on the approach to a tipping point, critical slowing down (CSD) occurs in the system. The increase in memory as this occurs can be seen in the time series in Fig. 1.2a and in the detrended time series of Fig. 1.4. As previously stated, CSD is caused by the leading eigenvalue of the Jacobian matrix of the system approaching 0 (Wissel, 1984). However it is impossible to calculate the exact value of this eigenvalue without knowing the governing equations of the system. In reality, only a time series is likely to be available. There are many ways to estimate that the eigenvalue is changing however, and it could be possible to directly estimate the eigenvalue itself with more information than just a time series (Lade & Gross, 2012).

Lag-1 autocorrelation (AR(1) coefficient estimation)

Critical slowing down can be estimated by an increasing lag-1 autocorrelation:

$$x_{t+1} = \alpha x_t + \varepsilon \qquad (1.2)$$

Where x is the time series, α is the lag-1 autocorrelation (AR(1)) coefficient and ε is random white noise that perturbs the system. As the system approaches the tipping point, α is expected to approach 1. Clearly the increase in α causes each x_{t+1} to become more like x_t. Usually α is estimated by fitting a model of the form of Eqn. 1.2 to the data.

The indication of an increasing AR(1) coefficient on approach to a tipping point was first used in climate data by Held and Kleinen (2004). They used a climate model of intermediate complexity (CLIMBER2) to simulate a shutdown of the Atlantic Thermohaline Circulation (THC). To do this they aggregate their data into fast, medium and slow time scales, the fast representing the noise, the medium the response to those perturbations, and the slow in which the approach to the bifurcation is happening. This time scale separation is also described by Thompson and Sieber (2011).

Dakos *et al.* (2008) test this same method, with the use of different detrending bandwidths and window lengths for robustness (in their supplementary information), on 8 climate proxy time series from palaeo-records. Positive results were found by both Held and Kleinen (2004) and Dakos *et al.* (2008) although some results by Dakos *et al.* (2008) were not as convincing as others, mainly due to lack of data points. These results showed only general increases in AR(1) coefficient estimation under a lot of noise and so had low but positive Kendall's τ, compared to time series containing more data points which had higher τ values. Dakos *et al.* (2008) also simulate 3 tipping points occurring and these suggest an increasing AR(1) coefficient. In the supplementary information, Dakos *et al.* (2008) use null models as previously described to determine how significant the Kendall's τ of the AR(1) coefficient time series are.

An approach of the AR(1) estimate to 1 suggests the tipping point has been reached. Held and Kleinen (2004) find that they are able to predict when the tipping point would occur by fitting a linear regression through their AR(1) time series, although their 95% confidence interval of this reaching 1 is around

10000 years wide (about 1/5 the length of the time series). Generally only increases in the value of the AR(1) estimate are considered early warning signals as frequency of data points and detrending bandwidth can affect its value.

Increasing DFA exponent

Another method for observing critical slowing down comes in the form of an increasing detrended fluctuation analysis (DFA) exponent. This tests for the decay in correlations over time and is related to estimating the AR(1) coefficient, although the latter only compares correlations to the previous time point rather than over longer time distances. The DFA exponent is calculated by integrating the time series (creating the profile), dividing it into windows of a chosen size, fitting a polynomial of chosen order to each window and assessing the squared difference between this polynomial fit and the profile.

$$F^2(v,s) = \frac{1}{s}\sum_{i=1}^{s}[Y\big((v-1)s+i\big) - y_v(i)]^2$$

Y is the integrated time series (profile function), $y_v(i)$ is the best fit polynomial of chosen order for segment *v* and *s* denotes the size of the windows. These differences are averaged over the all the windows and repeated for different window sizes. The fluctuation function increases by a power law as $F \propto s^{\alpha}$ with α being the DFA exponent.

The DFA exponent theoretically approaches 1.5 as the tipping point is reached (but is usually rescaled so it approaches 1), although as with the AR(1) estimate, its value can be altered by frequency of data points and detrending

method so any increase is deemed an early warning signal. DFA calculations could be considered more robust than AR(1) estimates as only a choice of window length has to be made since detrending is part of the method. However this does depend on the type of detrending as different orders of polynomials can be fitted through the data to detrend it (similar to using different Kernal smoothing bandwidths). Generally polynomials of order 1 (straight lines) are fitted through the data.

The first application of this method on climate data was on a model shutdown of the THC and palaeoclimate data over the last deglaciation by Livina and Lenton (2007), which was successful in that increases in the DFA exponent were observed, despite the poor resolution in the palaeo record. It was also used in a test of robustness against other methods by Lenton *et al.* (2012) for 6 tipping systems, including model data and palaeo records. In some cases stronger trends were found compared to the AR(1) estimate and in others the trend was not as strong but was always positive except in 1 instance where it failed, although in all of these instances, it is not certain that there was an approach towards a tipping point and the abrupt shifts may have been noise induced (see beginning of Chapter 1). The AR(1) estimates which could give either sign for the same series depending on the choice of bandwidth and window length sometimes showed values of τ closer to 1 than the DFA analysis. However the 2 methods complement each other as testing both can provide some measure of robustness (if both suggest a tipping point is approaching then it is less likely to be a false alarm).

Spectral 'Reddening'

Spectral reddening although more difficult to estimate numerically, can potentially show graphically that there is an approach to a tipping point with higher power found at lower frequencies nearer the tipping point (Fig 1.5a). A shift in power from high to low frequencies occurs as perturbations tend to have a longer lasting effect on the system. This is analogous to the AR(1) coefficient estimation as a loss of power at higher frequencies (contains less high frequency noise) suggests that the system is becoming more sluggish and slowing down. Kleinen *et al.* (2003) show this for a simple 2-box model of the THC by plotting the spectrum of the time series at different distances away from the tipping point.

There is potential for spectral reddening to be used in a time series analysis method (using a sliding window). The power law exponent can be estimated by fitting a regression line through the log power spectrum of a segment of the time series (Fig 1.5b). An exponent of approximately 0 implies white noise whilst lower (negative) values equate to a reddening. The redder a time series is, the more negative the exponent becomes so recording its negative value may make it easier to understand (so increases suggest reddening). The power law exponent is related to both the AR(1) coefficient and the DFA exponent by the equations $\beta = 2\alpha - 1$ and $\alpha = 1 - \frac{\gamma}{2}$, where β is the power exponent, α is the DFA exponent and γ is the AR(1) coefficient (Livina & Lenton, 2007).

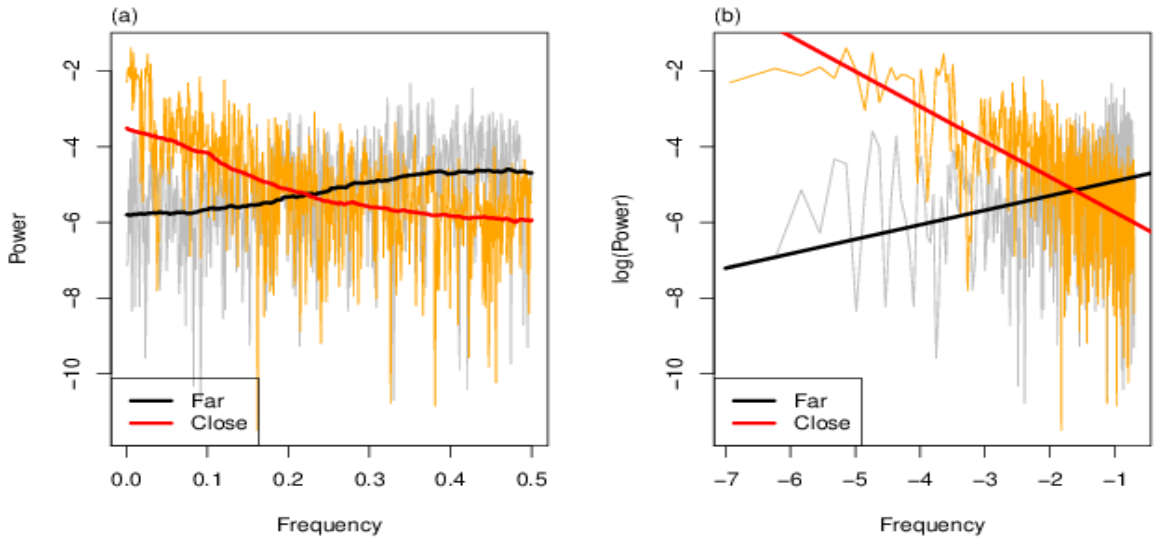

Figure 1.5: Power spectrum of two time series (created using Eqn. 1.1) where system is far (grey/black) and close (orange/red) to the tipping point showing (a) power and (b) log(power) as a function of frequency. A Kernal smoothing function has been applied to the spectrum for power (a) and a linear regression line fitted to the log(power) (b).

Increasing Return Time

In an idealised situation, the resilience of the system could be measured by artificially perturbing and recording how long the recovery or return time is to settle to equilibrium again. Wissel (1984) noted that the return time after a perturbation to the system will be longer if closer to the bifurcation. Using this as an early warning signal in the real world would be difficult, especially on the large scales of climate systems where large scale perturbations would be difficult to perform and may also have adverse effects. Also the recovery time might be difficult to determine under short term noise that appears in real

systems. However it has shown potential in small ecological systems where perturbations can be more easily implemented, as well as in modelled systems.

The method was used by van Nes and Scheffer (2007) when looking at 6 different ecological models with the increase in return time being observed a long time before the transition. Measuring the return time has been adopted more recently in a laboratory experiment by Veraart *et al.* (2012) where 2 populations of cyanobacteria are shown increasing levels of light until they pass a threshold (due to being unable to shade one another) and the tipping point is passed.

Eigenvalue estimation and higher dimension systems

Early warning signals which use extra information could potentially diagnose the eigenvalue and its approach to 0 directly rather than trying to estimate this from the AR(1) coefficient. One example of this uses information such as the functional form of the equations governing the system. From this, the system is linearised i.e. the Jacobian matrix is found and calculated at each time step based on multiple time series of known information to find an estimate for the eigenvalue (or eigenvalues for a system with two or more dimensions). In dynamical systems analysis, eigenvalues with negative real parts imply the state is stable and an increase towards 0 means the system is approaching a bifurcation or tipping point. The AR(1) estimate acts as a proxy for the eigenvalue whereas this method attempts to directly calculate it. The eigenvalue's absolute value is unaffected by any data frequency or user choices and so can potentially be more useful, especially in cases where the sampling of a system can be a sparse or inconsistent.

Lade and Gross (2012) first use this method on models of increasing complexity, the simplest of these being an Allee effect, 1-dimensional model similar to that of Takimoto (2009) (who looks for increases in variance, see later). The method successfully predicts the tipping point with only 15 time points prior to tipping, knowing only the population and the birth rate of the species at each time point. They do however assume the death rate is linearly increasing with the population.

Lade & Gross also test their method on the fisheries model from Biggs *et al.* (2009) and a tri-trophic food chain model both of which are higher dimension systems than the Allee effect model. More information is needed but the method still shows promise for a climate system. There is also the ability to track more than one eigenvalue (for systems of greater dimension than one) rather than just one in case there is a sudden switch of which one is closest to 0. For example, one eigenvalue could stay constant but be close to 0, and another could be largely negative but quickly increasing towards 0. An approach to the tipping point would be diagnosed earlier, compared to a method such as AR(1) coefficient estimation which would only register the constant eigenvalue until the switch occurred. Furthermore, the type of bifurcation being approached can be determined. By observing both the real and imaginary parts of the eigenvalues in the fisheries model, it was determined that a Hopf bifurcation was crossed (Lade & Gross, 2012).

Although this method shows promise, it is unclear if this type of method could be transferred to the real world or even to more complex, higher dimension climate models due to the extra information on the function forms that would be needed. The estimation of the eigenvalue can also be heavily influenced by

strong noise (Lade & Gross, 2012). There is however prospect for it to continue to work in simple systems.

The concept of extending early warning signals to systems of higher dimensions has been explored by Williamson and Lenton (2015), who show that other phenomena such as a decrease in the frequency of oscillations could provide early warning in more complex systems.

Increasing Variance

Another indication of the approach to a tipping point is an increase in the variance of the system (also measured as an increasing standard deviation). As previously mentioned, when the state of the system loses stability, the basin of attraction increases and the system is able to sample more of the surrounding space. Ditlevsen and Johnsen (2010) show that variance should increase as AR(1) increases and an increase in one but not the other should not be viewed as an early warning signal. The fluctuation-dissipation theorem shows that $\langle x^2 \rangle / T = \sigma^2/2$, where x is the time series, $T = 1/\alpha_\mu$ with the AR(1) coefficient $c = e^{-\alpha_\mu |t|}$, such that α_μ is the dominant eigenvalue for bifurcation parameter μ. α_μ tends to 0 as μ approached the critical value for tipping μ_{crit}. Hence variance will also increase according to the ratio above.

Visually an increase in variance near the tipping point can be seen in Figure 1.6. The example system is shown at two distances away from the tipping point (far away and near, as in Fig. 1.5) and the histograms show the changes in variance of the time series because of the differences in the bifurcation

parameter. The difference in variance is also noticeable in the time series themselves.

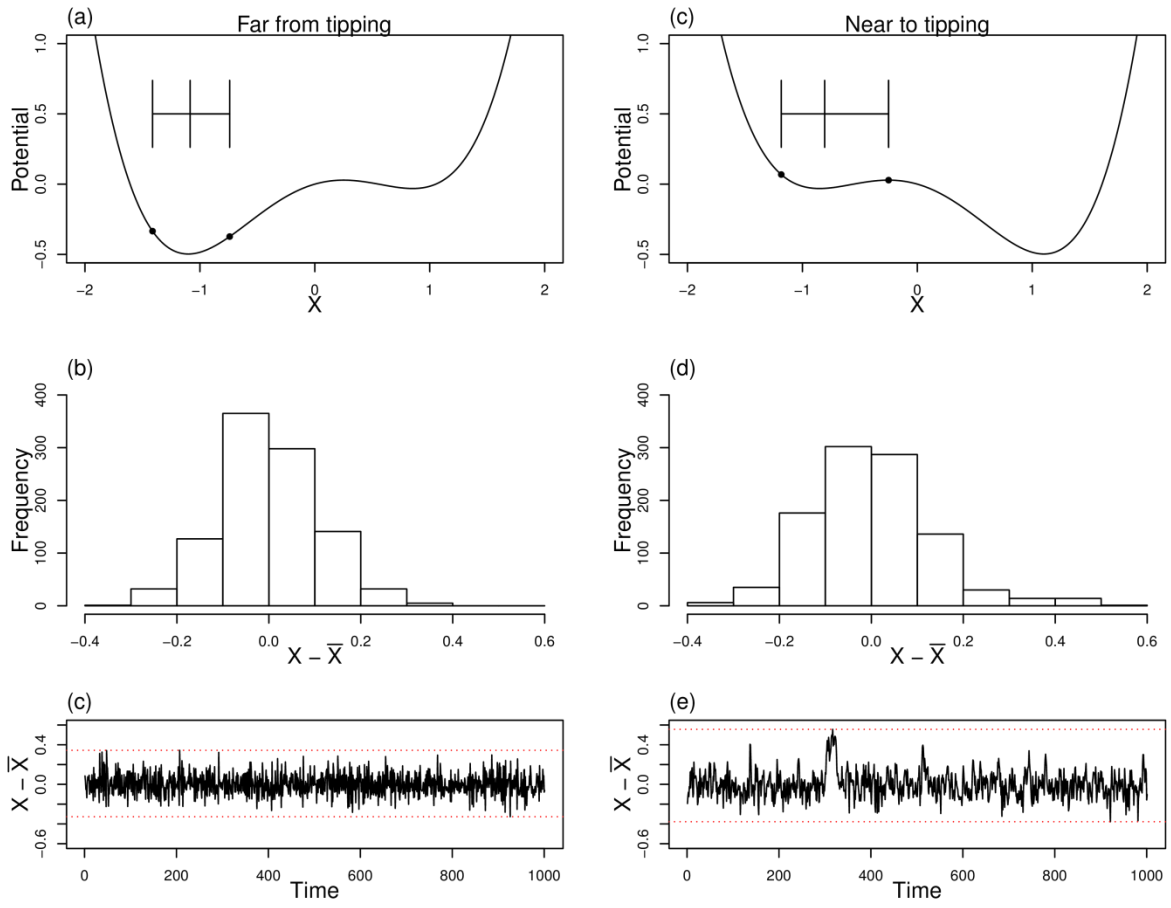

Figure 1.6: Differences in the variance of the system when it is far from (a-c, left column) and close to (d-f, right column) a tipping point. Black points in the potential well plots (a,c) correspond to the ranges in the time series (c,e) with this spread including the mean shown above. Histograms show the distribution of the system when it is far (b) and close (d) to the tipping point. The mean of the distribution/time series has been removed in b-c and e-f.

Variance is calculated temporally in the same way the AR(1) coefficient estimation, by first detrending the time series and calculating it on a moving window of a specified length. The trend of the resulting time series of variance is again calculated using Kendall's τ correlation coefficient. Unlike the AR(1) coefficient which approaches 1 towards a tipping point, variance could take any positive value and thus just an increase in variance can be viewed as an early warning signal. Variance can also be used as an indicator on the time series without detrending and the results will be less affected by inconsistent sampling of the time series compared to AR(1) coefficient estimation.

Measuring the variance was used as part of the robustness test carried out by Lenton *et al.* (2012) previously described. For the 6 time series, 3 were mainly negative, 1 of these robustly negative. The other 3 were robustly positive suggesting that in climate test so far results are mixed, although being real world systems, it is not certain that the time series are approaching a bifurcation. For a true test of variance as an indicator for early warning, a system where it is known for sure that a tipping point is being approached would have to be tested.

The method has been previously used in ecological systems, which are better understood in terms of knowing they are approaching tipping points, all with increases of the variance. The first of these by Carpenter and Brock (2006), uses a lake eutrophication model. Biggs *et al.* (2009) use a fisheries model and Takimoto (2009) measures variance as well as other methods to test for an early warning in an Allee effect model.

In some cases variance could decrease on the approach to a bifurcation. Dakos *et al.* (2012) show that in some cases, factors in the environment can fluctuate

stochastically and the system becomes less responsive to these fluctuations near a tipping point. However it will always increase at the tipping point. Variance can also be seen to decrease if the Kernal smoothing function leaves only high frequency data due to a low bandwidth. An approach to a tipping point would result in less high frequency movement (and so a slowing down as seen already) meaning the variance at high frequencies would decrease.

Temporal Skewness

Changes in the skewness of the distribution of a system can also be an indicator of approaching a tipping point of the kind shown in Fig. 1.2, commonly known as a fold or 'saddle-node' bifurcation. Skewness is not expected to change for a pitchfork bifurcation, where a stable state splits symmetrically into two stable states separated by an unstable steady state (although other indicators are expected to work for these as shown by Kuehn (2011)). However these bifurcations do not exist in the systems analysed in this thesis. As the stability of the state decreases and the potential well shallows, it will also 'lean' towards another stable state, becoming asymmetric. Also in certain cases, when the amplitude of the noise driving the system is high, the system will begin to sample the other state before coming back to the state losing stability (known as 'flickering', see below). These two characteristics cause the system to become skewed towards the new state. The direction of the skewness is dependent on the position of the other state, thus a change in skewness regardless of sign is considered early warning rather than indicators based on critical slowing down or variance where strictly an increase is searched for.

Skewness, the third moment of the distribution, is defined as:

$$\gamma = \frac{\mu_3}{\mu_2^{\frac{3}{2}}}$$

With:

$$\mu_k = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^k$$

Where *N* is the total number of points and $\bar{x}$ the mean of the data. As with variance and AR(1) coefficient estimation, skewness is measured again on a moving time window (which would be equal to *N*). Note that detrending can affect the result of this method and can be tested with the residuals or the original data. As the calculation of skewness involves residuals from the mean on a moving window, any drift in the original time series does not interfere with a signal. An example of skewness changes in the simple model used previously can be seen in Fig. 1.7, where it can also be seen graphically as the window is moved along the time series as an evolving histogram. To ensure skewness can be seen more clearly in the histograms (Fig. 1.7 c,d) μ is increased more slower than previously. Skewness is first seen in use by Guttal and Jayaprakash (2008). They test the indicator on model data from a vegetation collapse simulation which shows a clear increase in indicator from the beginning of the time series and a lake euthrophication model which gives an early warning signal which could be observed around 5 years before the tipping point occurs (from 45 years of data). However when skewness is used as an indicator on palaeo records of sediment data showing the desertification of the Sahara, there is no trend suggesting a tipping point is being reached, although the small length of time series prior to desertification makes it difficult to test the

robustness of this result. It is also unclear if this system approached a tipping point, or if it was a noise induced transition (see beginning of Chapter 1).

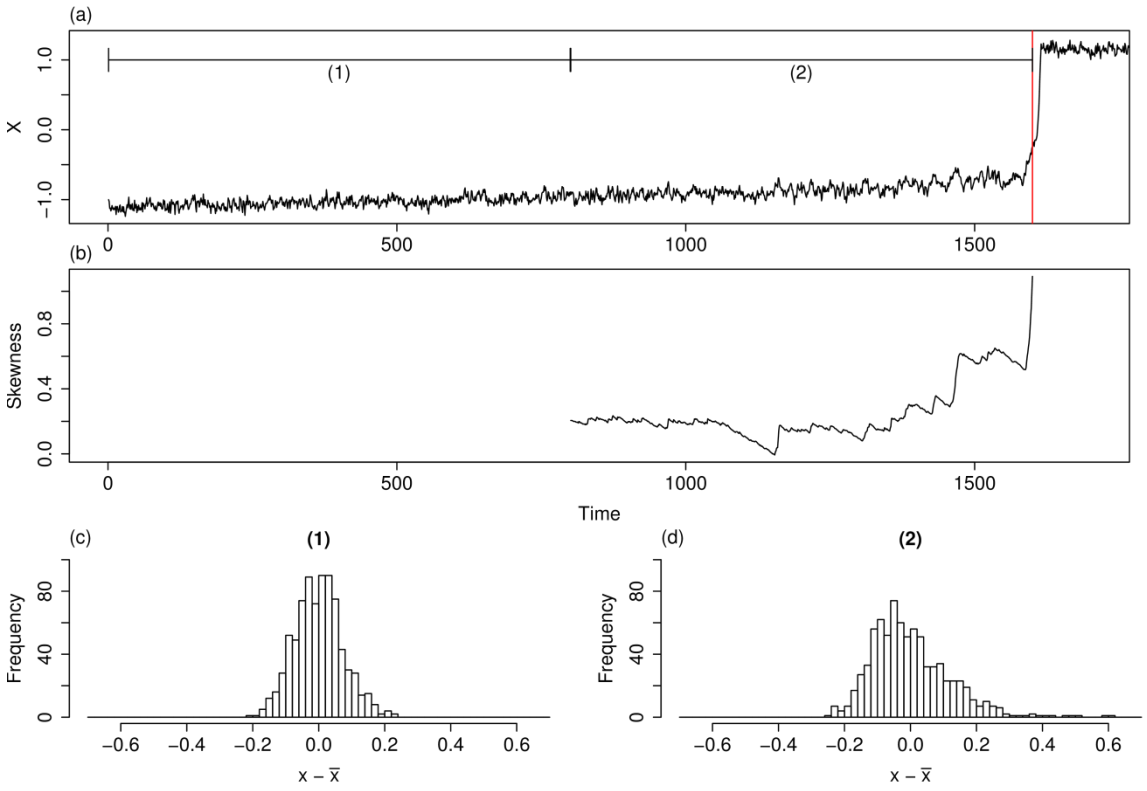

Figure 1.7: Changes in skewness as the system approaches the tipping point. Time series (a) is cut prior to tipping point and skewness (b) is measured on a moving window of 800 points. Histograms (c,d) of the sections of time series with the mean subtracted shown in (a) show an increasing positive skewness.

Flickering

Connected to changes in skewness, flickering of a system occurs in time series which start to sample another state for a short amount of time before returning back to state that is losing stability. It requires a large enough amplitude of noise to move between states and for the state the system is eventually moving to not be too stable such that hysteresis occurs (i.e. the noise can move the

system away from that state and back to the other). While it is difficult to be used as a quantitative early warning signal, the phenomenon can provide important information about the other state such as how far away it is. Dakos *et al.* (2013) show flickering in a number of modelled systems which is accompanied with increases in skewness and variance when the flickering occurs.

Other early warning indicators

The indicators mentioned above are generally deemed ‘time series early warning signals’ and are tested on a single time series representing the system. However other indicators exist in the literature. These include spatial indicators which use different locationally connected time series of a system (such as time series of vegetation changes) and more ‘system-specific’ indicators that are specialised to suit the system they are tested on.

Spatial early warning indicators

The indicators described above have also been tested in a spatial context. While the temporal methods need a relatively long time series to be able to use a window length of the data to create a time series of an indicator, spatial indicators can be calculated in one time step (although multiple time steps are needed to create a time series as before), thus the extra spatial data can supplement the lack of temporal knowledge. Indicators include spatial correlation (akin to CSD), variance and skewness.

Dakos *et al.* (2010) use spatial correlation on a lake eutrophication model, a resource-overharvesting and a vegetation-turbidity model. They use Moran's coefficient to express the spatial correlation which needs a binary matrix describing which grid points are neighbouring and test the models under different levels of heterogeneity and connectivity (how much exchange occurs between neighbouring cells). Then measuring Kendall's τ for each instance, they find that generally the spatial correlation is a better indicator than the AR(1) estimate for the systems which have high connectivity.

Spatial correlation as suggested above can be thought of as 'spatial critical slowing down'. As a tipping point is approached, diffusion will start to dominate the reaction-diffusion dynamics of the system which in turn suggests that neighbouring grid boxes will start to become more like each other and correlated.

Rietkerk *et al.* (2004) show that self-organised patchiness suggests the system is becoming more vulnerable and these early warning signals stem from this idea. Bailey (2010) suggests that single ‘snapshots’ of a system may be unreliable in determining how far away from a threshold the system is and that observing changes in the patterns over time could be more beneficial.

Spatial variance measures the variance using the spatial data points from one time point. In theory, increases should be seen due to the shallowing of the potential well as is the case with temporal variance. Spatial variance is yet to appear in the literature as an indicator within climate science, possibly due to the difficulty in accounting for different weather occurring in different regions of a climate system due to its size, which will alter the variance signals observed. This could also be true of spatial correlation indicators.

Litzow *et al.* (2008) use spatial variance on data from fish stock levels around the North Pacific and Atlantic. Although no quantification is made (such as calculating Kendall's τ), they argue that an increase in variance can be seen prior to the regime shifts of fish population.

Guttal and Jayaprakash (2009) use the method on a model of vegetation collapse. Increases are seen but not until near the tipping point (within 10 years) when compared to the whole length of the data set prior to the transition (around 50 years).

Bathiany *et al.* (2013) use a spatial method to detect 'hot spots' in spatial systems that maximise an early warning signal of vegetation collapse in a general circulation model (GCM). These subsets of the system are thought of as governing the tipping on a larger scale.

Network analysis can also be used as an early warning signal, which assesses the connectivity of spatial points (van der Mheen *et al.*, 2013). Points are connected if, over time, their time series are correlated (above a threshold correlation coefficient). An increase in connectivity over time indicates an approach to a tipping point. This has been tested on AMOC collapse by Feng *et al.* (2014) on the same data used in this thesis (Chapter 2).

Spatial skewness is measured the same way as temporal skewness, except rather than measure skewness on a moving window, a single temporal point can be used by collapsing the spatial information to 1-dimension, supplementing a potential lack of temporal data with spatial. Like temporal skewness, increases or decreases in spatial skewness can be considered indicators.

Guttal and Jayaprakash (2009) use this method along with the spatial variance (described above) on a vegetation collapse model, similar to their 2008 paper when testing temporal skewness as an indicator. They find, as with the spatial variance, increases in spatial skewness occur near the tipping point (around 10 years before) with 50 years of data before tipping occurs. Whether this is enough time to provide a useful early warning signal is dependent on the system as the system has to have time to respond the prevention measures before the tipping point is reached.

System specific stability indicators

There is also the prospect that early warning signals that are more specific to a subset of systems could exist. The methods above work generally for time series. However for some systems it can be hard to obtain a time series where these methods work well, for example, most readily available data on the Amazon rainforest is its area coverage which remains fairly constant and so indicators may not change. The methods described below show prospect for conversion to early warning signals but at this time are closer to stability indicators or tests of vulnerability rather than tracking a movement towards a tipping point.

Good *et al.* (2011) have created an index from model data (HadGEM2-ES) which expresses the vulnerability of an area of the tropical forest based on the temperature and the dry-season length (number of months with precipitation below a given threshold, DSL). With information from all tropical grid points combined including the forest coverage of each point, the combinations of temperature and DSL suitable for sustainable forest growth can be determined.

From this, areas of forest that are vulnerable (i.e. near temperature/DSLs unsuitable for forest) can be found and with future projections of changes to temperature and DSL, predictions can be made on the future of the forest in these areas. Converting this to the real world makes assumptions on unknown real world parameters which are set to values in the model, along with assumptions that the world behaves the same way. However potentially, changes in external parameters such as temperature and DSL could be monitored as an indicator of approaching a tipping point.

Also concerning the Amazon rainforest, Hirota *et al.* (2011) use real world observations of forest coverage and annual precipitation to reconstruct the potential wells of the system (as in Livina *et al.* (2010)) using the precipitation as an axis rather than time. From this, forest vulnerability can be determined by observing how close it is to crossing the threshold in the potential well based on how much precipitation there is. Although this shows promise and appears to work well, it uses only annual precipitation as a proxy for the climate of the area ignoring important factors such as seasonality and temperature. Adaptations however could be made to account for this to create an early warning signal which would include monitoring changes in precipitation to see if areas are becoming too dry or even if the potential well itself is changing over time.

To measure the stability of the AMOC when it is in its 'on' state, the net freshwater import into the Atlantic (F_{ov}) has been proposed (de Vries & Weber, 2005, Hawkins *et al.*, 2011, Rahmstorf, 1996).

$$F_{ov}(\Phi) = -\frac{1}{S_0}\int_{-D}^{0} \overline{v*}\,(z,\Phi)\langle S(z,\Phi)\rangle dz$$

Where S_0 is a reference salinity (35 psu), z is depth,Φ is the latitude, $\overline{v*}$ is the zonal integral of northward baroclinic ocean velocity, $\langle S(z,\Phi)\rangle$ is the zonal mean salinity and D is the total depth of the ocean. The F_{ov} is measured at 34°S with a positive value meaning the AMOC is exporting salt out of the Atlantic, promoting a negative feedback and thus a strengthening of the AMOC should it weaken. A negative value means salt is being imported and suggests that any weakening of the AMOC would cause a positive feedback and so the AMOC is unstable. The F_{ov} indicator has been measured in the model used in this thesis (Chapter 2) where it becomes negative when bistability is present (Hawkins *et al.*, 2011).

Potential problems with early warning signals

Although the methods described above have sound theory explaining why they should work, in practice this is not always the case, particularly if the system does not fit the requirements of the theory. These requirements include the ability to separate the time scales of the system (as done by Held and Kleinen (2004)) and that the noise driving the system on small time scales is white and of a constant variance across the time series. Caution needs to be taken if any signal is found. Both 'false alarms' (an indication of a tipping point when there is no change to the system, such as those found in null models by Dakos *et al.* (2008)) and 'missed alarms' (no indication of a tipping point when there is one approaching, for example the desertification of the Sahara tested by Guttal and Jayaprakash (2008), assuming a bifurcation occurs in this system) have been observed in the literature listed above. Generally only a few time series exist for a given system each of which can be viewed as a single realisation meaning it is difficult to determine if any results found are a false alarm or missed alarm. More research is needed to show how effective these methods are, using both model systems (not climate specific) and climate records. The use of models allows more realisations to be made to see how an indicator behaves generally, but at the loss of simplifying the system.

Work by Ditlevsen and Johnsen (2010) involves the use of a simple double potential well system which is linearly transformed towards a bifurcation and a 1 state system, thus creating a tipping point. Alongside this, a control simulation is run where there is no change in the bifurcation parameter which would cause the approach to the tipping point. With these runs, both the AR(1) coefficient

and the variance are estimated using the methods above for each of the time series. Once these indicators are calculated, they are plotted on top of each other and the error bars (a 2 σ confidence band) on the control run and the bifurcating run AR(1) time series overlap up to the tipping point, although the Kendall's τ rank statistic is not tested here with error bars attached to it, something which may provide extra information. The error bars on the variance of the bifurcating run and the control separate before the transition suggesting that in this instance it is a better indication for using as an early warning signal in this system.

Ditlevsen and Johnsen (2010) also ran the same system but create a noise induced transition whilst keeping the bifurcation parameter constant. Using the same methods as the bifurcation and control runs, they find that neither of the indicators increase as expected. Then using high resolution NGRIP data, they overlay 25 Dansgaard-Oeschger (DO) events so the transition takes place at the same point and test the same indicators as above and find no robust increase overall, suggesting that these are noise induced transitions. However (Cimatoribus *et al.*, 2013) suggest from their results that the switches in states are due to changes in some external forcing when they test indicators on the ensemble mean of the DO events (similar to the method of Kuehn (2011) detailed below).

The method in Ditlevsen and Johnsen (2010) can be considered as testing the DO events against a null hypothesis. They are aware of how the indicators behave when a tipping point is not being approached, and compare the results from time series of which they are unsure if there is an approach and test how significantly different these indicators are.

Kuehn (2011) tests variance and AR(1) indicators using Monte Carlo ensembles. With a sample size of 1000 mathematical models of approaches to different bifurcations (fold, transcritical and pitchfork) each ensemble member is different due to the added noise. From each ensemble, the overall trend in the indicator can be calculated. Indeed all indicators show an increase as expected. However taking randomly selected individual members implies that sometimes the indicator can show negative tendency, again reiterating that real world time series are just one realisation and that missed and false alarms could easily be found within them.

Boettiger and Hastings (2012) test indicators observed in an ensemble of runs of a system simulating the Allee effect where the bifurcation parameter, and so the distance to the tipping point, is not changing. However in some cases the tipping occurs by chance. The indicators for these ensemble members give stronger signals than those which remain in the same state. They conclude that the stochastic process driving the system, increase the autocorrelation prior to the system tipping and that these chance transitions create a false positive for early warning signals, which are supposed to indicate the bifurcation parameter is changing, and so potentially a false positive for past tipping points which exhibit early warning. This suggests that there is a bias towards finding a positive early warning signal when looking for one in time series that have experienced tipping points and care should be taken when assuming a positive result means a tipping point was approached, rather than noise induced.

Unanswered questions

Although these methods have been tested on abrupt changes in palaeoclimate (Dakos *et al.*, 2008, Livina & Lenton, 2007) and modelled experiments (Held & Kleinen, 2004), little effort has been placed in researching if these indicators could help determine if climate tipping points are being approached now, or could be in the future. One of the main problems preventing this is that monitoring of these climate systems has yet to produce time series that are long enough to observe significant signals from. One counterexample of this is by Livina and Lenton (2013), where the signals are tested on daily sea ice extents to determine if the ice free summers are being approached in the Arctic.

This lack of data naturally leads to testing these methods in climate models. As well as giving longer time series for systems in the present and towards the future, the indicators can also be tested under different levels or types of forcing and with repeats, allowing a test of robustness of the signals that could not be obtained from one realisation. Signals can also be tested in many spatial locations compared to in the real world where monitoring of a system could be sparse. This could even inform potential locations to monitor real world systems. Furthermore, general circulation models (GCMs) are more complex than simple systems that theory states these early warning signals will work on. Hence testing generic early warning signals on data from these would begin to determine how useful they are in systems that might not necessarily be reducible to 1-dimenstion and so begin to determine how useful they would be when testing for early warning in even more complex, real world data.

Another aspect of generic early warning signals that needs more research is determining the significance of an observed signal. There have been some attempts at using null models to determine significance (Dakos *et al.*, 2008) but these were not emphasised as a main result in the paper. It is vital to understand how significant any early warning signal observed is, preventing signals which appear insignificant from being dismissed. It is also important to determine how long a time series would be needed until significant signals are observed. Of course in the real world, the distance from the tipping point is unknown but an estimate of how close to the tipping point significant signals could be found could be tested for in GCMs for example.

In the instances when these generic indicators fail, understanding why they do not work is important. One of the main causes for this is that the forcing could be too fast for the systems and it lags behind. The methods could be tested on the system when it looks to be near equilibrium but actually the stability of the state is lost and it is yet to tip. This fast forcing problem is likely to occur when testing for early warning under anthropogenic forcing over the next century. Another example of why early warning signals may fail is if the noise level driving the system on the short timescales is too high. This prevents the separation of timescales needed for generic early warning signals from working. If the strength of this noise also changes over the time series, it could also cause problems. For example, an increasing noise level will increase the variance over time, regardless of whether or not the tipping point is being approached.

With these problems with generic early warning signals, ‘system specific indicators’ could be a better prospect. Generic early warning signals work for

tipping points that are approached slowly and linearly. However with these conditions being unlikely for real world climate systems, it could be that early warning systems specifically created for the system in question could give better warning.

Finally, with slowing down occurring in systems approaching tipping points, little research has been conducted into the effect this will have on other systems. For example, Di Lorenzo and Ohman (2013) show that indices created by atmospheric white noise forcing integrated over time (Hasselmann, 1976) drive systems that themselves integrate this now red noise forcing over time. If this forcing system were to get redder over time, then this will have an effect on the systems it is driving.

Outline

In the next four chapters of this thesis there are various attempts to answer the problems detailed above, specifically determining if early warning signals could be used on climate systems approaching tipping points over the next century and beyond. This is done through the use of GCMs which are forced with plausible 21st century emission scenarios. The significance of any signals are also tested with the use of null models, and the length of data needed to produce a significant signal is also tested in one case study. In another case study, where the generic early warning signals fail, a system specific indicator is suggested. Finally the effect of slowing down in a system, on ecosystems that are driven by it is explored. While the explored driving system in this thesis is not approaching a tipping point itself, the results could be extended to a system that is, showing that even the approach towards a tipping point could have adverse effects on other systems.

In Chapter 2 ‘Early Warning Signals of Atlantic Meridional Overturning Circulation Collapse in a Fully Coupled Climate Model’, these generic early warning signals are tested on output from a GCM of a climate system that could potentially collapse in the future. While this model was slowly and linearly forced, the model itself is complex. However significant generic early warning signals are still found. Furthermore, a novel technique is developed to calculate how far in advance the tipping point can be predicted and how many years of data are needed to do that.

In Chapter 3 ‘Early Warning Signals of Simulated Amazon Rainforest Dieback’, the generic early warning signals are further tested when an ensemble of GCMs

are run with realistic emissions scenarios. An attempt to predict an approach to simulated Amazon rainforest dieback fails. However there is potential in using the increasing sensitivity of the forest to temperature as it approaches dieback. This has formed the basis of a system specific indicator as mentioned previously.

Chapter 4 ‘Exploring the Uncertainty of Modelled Amazon Dieback’ continues to look at the Amazon rainforest by observing measures other than generic early warning signals. In Chapter 3, it is revealed that part of the reason generic early warning signals fail is because the system lags behind the forcing. This suggests that potentially, the tipping point may have already been passed without realising it. Exploration in Chapter 4 concerns how much the system lags the forcing and how much loss the forest could be committed to, based on forcing up to the end of the 21st century. These methods also have potential to become a system specific early warning signal as they examine the vulnerability of the system and its predicted distance from the tipping point.

In Chapter 5 ‘Slowing Down of North Pacific Climate Variability and Implications for Marine Ecosystems’, explores the idea that slowing down of systems (both critical and non critical) could affect other systems. Finding that increasing AR(1) and variance are found in the Pacific Decadeal Oscillation (PDO) and associated sea surface temperatures, the impact of this on marine ecosystems is explored. In this instance, it is not believed that the PDO is approaching a tipping point, the indicator rises can be at least partially explained by other processes. Simple example systems (representing the marine ecosystems) are shown to react differently when the forcing upon it starts to ‘redden’, i.e. show signs of slowing down, by tracking the forcing system more closely and

increasing itself in variance. This change in behaviour of the dependent system could potentially be used as an early warning signal, especially if it is monitored better than the system driving it.

Chapter 6 is a discussion of the main messages of the previous four chapters and how the results within them collectively contribute to the field of research. Challenges for the future of early warning signals are also discussed. Finally the findings of this thesis are concluded.

Chapter 2

Early Warning Signals of Atlantic Meridional Overturning Circulation Collapse in a Fully Coupled Climate Model

This chapter is based on Boulton, C. A., Allison, L., C., and Lenton, T. M., 'Early warning signals of Atlantic Meridional Overturning Circulation collapse in a fully coupled climate model', 2014, Nature Communications, 5, 5752

Abstract

The Atlantic Meridional Overturning Circulation (AMOC) exhibits two stable states in models of varying complexity. Shifts between alternative AMOC states are thought to have played a role in past abrupt climate changes, but the proximity of the climate system to a threshold for future AMOC collapse is unknown. Generic early warning signals of critical slowing down before AMOC collapse have been found in climate models of low and intermediate complexity. Here we show that early warning signals of AMOC collapse are present in a fully coupled atmosphere-ocean general circulation model, subject to a freshwater hosing experiment. The statistical significance of signals of increasing lag-1 autocorrelation and variance vary with latitude. They give up to 250 years warning before AMOC collapse, after ~550 years of monitoring. Future work is needed to clarify suggested dynamical mechanisms driving critical slowing down as the AMOC collapse is approached.

Introduction

The Atlantic Meridional Overturning Circulation (AMOC) is a key component of the global climate system, responsible for a large fraction of the 1.3 PW northward heat transport in the Atlantic basin (Johns *et al.*, 2011). Numerical modelling experiments suggest that without a vigorous AMOC, surface air temperature in the North Atlantic region would cool by around 1–3 °C, with enhanced local cooling of up to 8 °C in regions with large sea-ice changes (Jacob *et al.*, 2005, Vellinga & Wood, 2002). Substantial weakening of the AMOC would also cause a southward shift of the inter-tropical convergence zone, encouraging Sahelian drought (Brayshaw *et al.*, 2009, Vellinga & Wood, 2002), and dynamic changes in sea level of up to 80 cm along the coasts of North America and Europe (Levermann *et al.*, 2005, Vellinga & Wood, 2008). Theoretical arguments (Stommel, 1961), numerical models of varying complexity (Hawkins *et al.*, 2011, Lenton *et al.*, 2009, Manabe & Stouffer, 1988, Rahmstorf *et al.*, 2005) and evidence from palaeoclimate proxy records (Broecker *et al.*, 1985, Clark *et al.*, 2002) support the existence of two stable AMOC states—'on' and 'off'. A reduction in density of the surface waters of the North Atlantic (through an increase in freshwater input or surface warming) can inhibit the formation of deep water and weaken the AMOC. In some model states, a weakening of the AMOC can result in an increase of freshwater transport into the Atlantic, resulting in a positive feedback. Numerical model projections suggest that the AMOC is likely to weaken over the 21st century (Cheng *et al.*, 2013), but the likelihood of an abrupt collapse is very uncertain (Kriegler *et al.*, 2009, Zickfeld *et al.*, 2007), partly because most state-of-the-art climate models used for future projections cannot yet correctly simulate past

abrupt climate changes (Drijfhout *et al.*, 2013, Valdes, 2011). This has generated interest in the possibility that generic early warning signals could exist before abrupt AMOC transitions (Held & Kleinen, 2004, Kleinen *et al.*, 2003, Lenton *et al.*, 2009), which might be diagnosed directly from data. However, existing studies of this issue have been restricted to models of low or intermediate (Held & Kleinen, 2004, Lenton *et al.*, 2009, Livina & Lenton, 2007) complexity.

For a low-order dynamical system approaching a threshold where its current state becomes unstable and it transitions to some other state, one can expect to see it become more sluggish in its response to small perturbations (Held & Kleinen, 2004). This phenomenon of 'critical slowing down' (CSD) can be detected in time series as increasing autocorrelation over time, measured by estimating the AR(1) coefficient (see Methods). Variance is also expected to increase (Carpenter & Brock, 2006) (noting that this is not independent of lag-1 autocorrelation). Here we analyse data from simulations of the fully coupled climate model FAMOUS (Smith *et al.*, 2008a). This is a lower resolution version of HadCM3, one of the models used in the IPCC Fourth Assessment Report. We look for early warning indicators in both the annual and decadal mean time series at 33.75°S–58.75°N (at every 2.5°). The time series of AMOC strength are cutoff at 800 years, before AMOC collapse, and are detrended using a Kernal smoothing function with a bandwidth of 400 years, before calculating the CSD indicators using a window length of 400 years (as described in Chapter 1). Throughout our analysis, Kendall's τ rank correlation coefficient is used as a measure of tendency of the indicators. A value of τ=1 implies an indicator is always rising, τ=−1, always decreasing and τ=0 having no trend.

Results

The AMOC in FAMOUS

The representation of the AMOC in FAMOUS is comparable to a variety of other fully coupled climate models (Figs 2.1, 2.2), all of which are of a higher spatial resolution, and several of which feature in the CMIP3 and CMIP5 data sets used in the IPCC assessments. The AMOC in FAMOUS is broadly similar to that in HadCM3 in terms of mean strength and decadal variability (Smith *et al.*, 2008a). Higher frequency and interannual variability in FAMOUS behaves similarly to other models when compared with observations (Balan Sarojini *et al.*, 2011). For variability on time scales longer than interannual, a comparison with observations is difficult due to the short observational record (currently around 10 years). Figure 2.1 shows the mean AMOC streamfunction averaged over multi-centennial control simulations of FAMOUS and five other coupled AOGCMs. Figure 2.2 shows that, although there is some inter-model variation, the magnitude of AMOC variability in FAMOUS on multi-decadal timescales is in the range of other models, both in terms of magnitude and pattern of variance. Furthermore, the mean transport at 26°N in FAMOUS compares well with the mean transport estimated from the RAPID/MOCHA/WBTS array (Rayner *et al.*, 2011) over the years 2004–2012, where the AMOC is currently being monitored (both are ~17.5 Sv) (Smeed *et al.*, 2014).

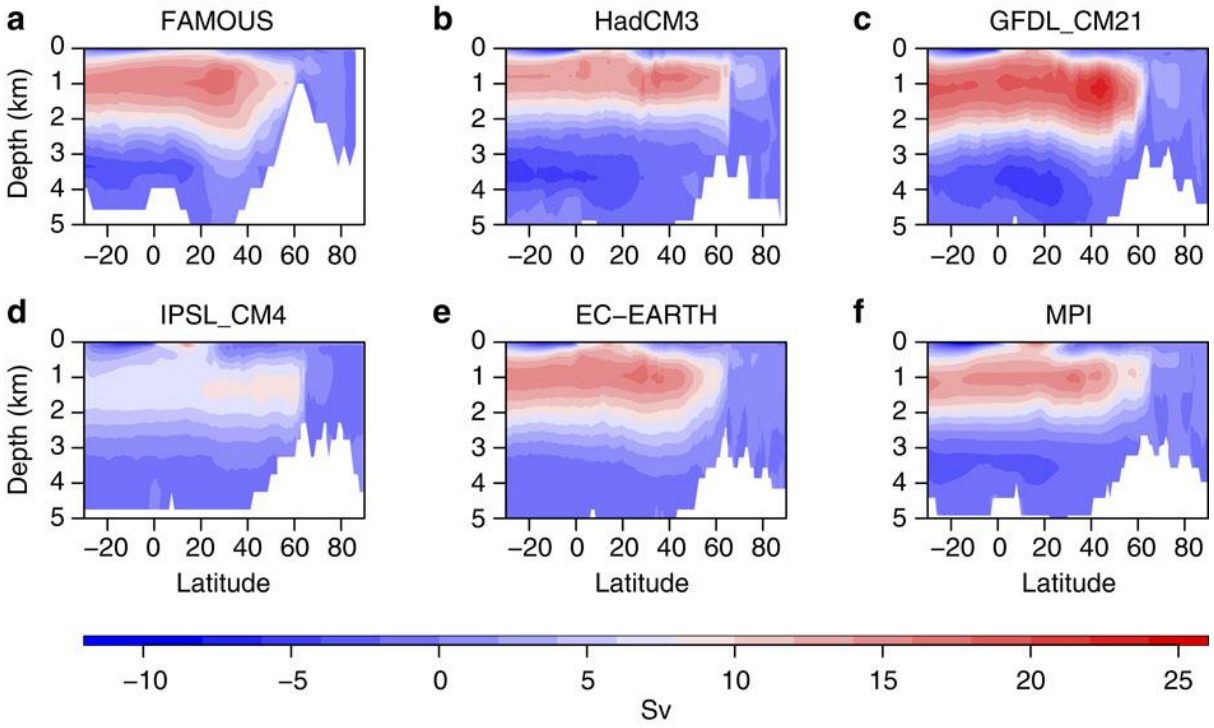

Figure 2.1: Mean AMOC streamfunction in FAMOUS and five other coupled AOGCMs. The streamfunction fields obtained from multicentennial control simulations are shown for (a) FAMOUS, (b) HadCM3, (c) GFDL_CM21, (d) IPSL_CM4, (e) EC_EARTH and (f) MPI.

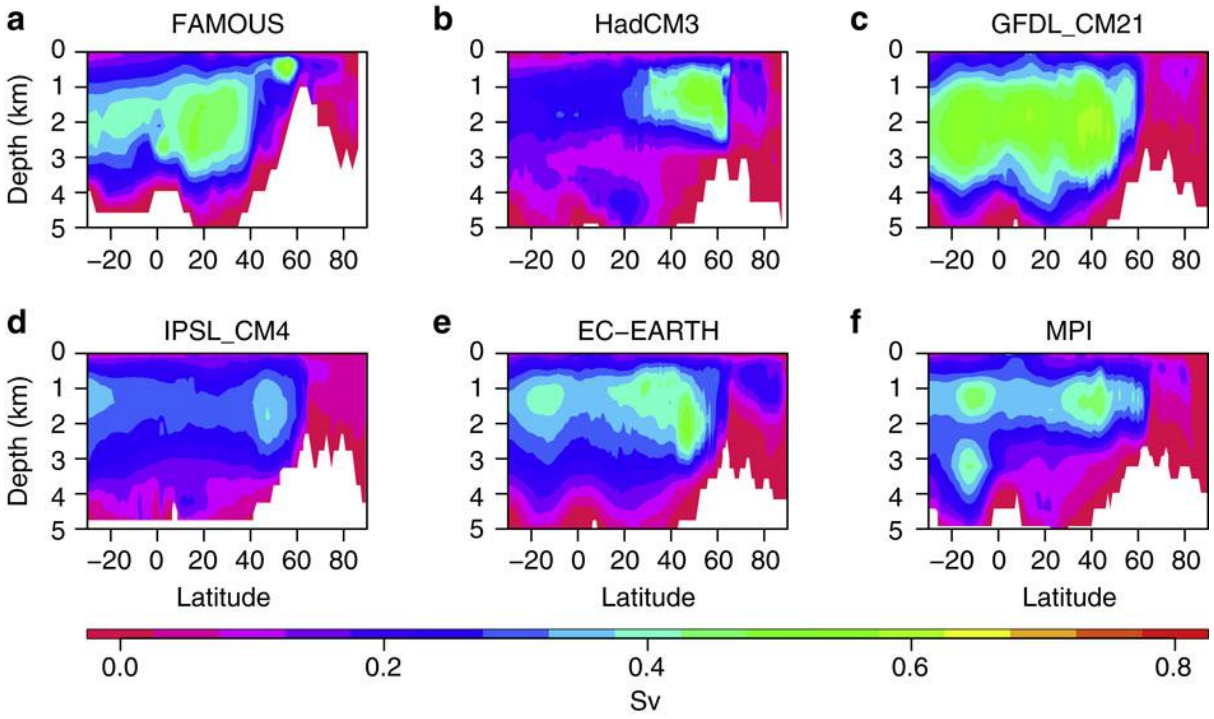

Figure 2.2: Local standard deviation in AMOC streamfunction on time scales longer than 50 years. The standard deviation in streamfunction is shown for (a) FAMOUS, (b) HadCM3, (c) GFDL_CM21, (d) IPSL_CM4, (e) EC_EARTH and (f) MPI.

FAMOUS has been subjected to a hosing experiment (Hawkins *et al.*, 2011), where freshwater forcing is applied in the North Atlantic between 20°N and 50°N (and compensated with a spatially uniform salt flux to conserve global mean salinity). This acts to reduce the density of the surface waters, inhibiting the formation of North Atlantic Deep Water, and weakens the AMOC. This forcing is gradually increased at a rate of 5×10^{-4} Sv yr^{-1}, eventually causing the AMOC to transition into the 'off' state after about 800 years of simulation, at a freshwater input of 0.4 Sv. The freshwater forcing is then removed at the same rate, allowing the AMOC to recover, with associated hysteresis (Fig. 2.3). Data from the simulation is saved at annual resolution, and we analyse just the

upper branch of forcing towards AMOC collapse. We also make use of equilibrium simulations in our analysis (Fig. 2.3, black points and Results). These are initialized at particular values of forcing before the collapse occurs and run to equilibrium, either remaining in the ‘on’ state or eventually transitioning to the ‘off’ state.

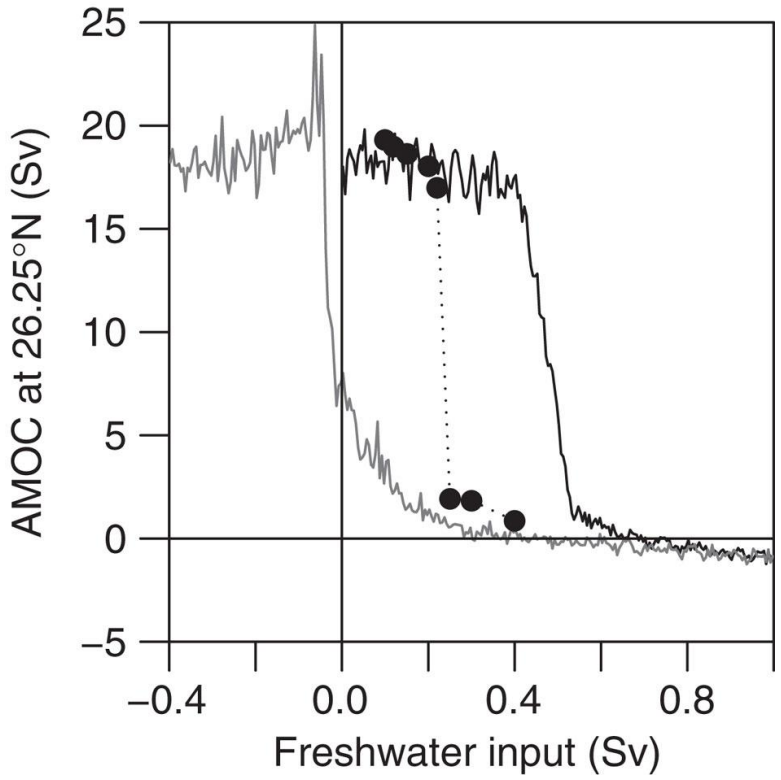

Figure 2.3: Hysteresis of AMOC in the FAMOUS model. AMOC transport (Sv) at 26.25°N and ~1,000 m depth is plotted as a function of imposed freshwater input. The solid black line shows the decadal mean AMOC during a transient experiment with freshwater input increasing from 0–1 Sv over 2,000 years. The solid grey line shows the same for freshwater input decreasing from 1 to −0.4 Sv at the same rate. The filled circles show the equilibrium AMOC transport reached during a series of constant-forcing simulations that are initialized from the corresponding point of the transient (increasing forcing) simulation. Adapted from Fig. 2 of Hawkins *et al.* (2011).

Early warning signals of AMOC collapse at annual resolution

We began our search for early warning signals of AMOC collapse at 26.25°N (Fig. 2.4a), near where an estimated reconstruction of the AMOC is currently monitored by the RAPID-WATCH/MOCHA/WBTS array in the real ocean

(Rayner *et al.*, 2011). Slowing down can be seen by eye in the detrended fluctuations of AMOC strength at this latitude (Fig. 2.4b). With a sliding window length of 400 years (black lines in Fig. 2.4c,d), AR(1) is found to be rising (with τ=0.79), with a less strong rise in variance (τ=0.39). These increasing trends in the indicators are generally robust to varying window length and detrending bandwidth (Fig. 2.4c–f). A decrease in variance seen when detrending using a low bandwidths (Fig. 2.4f) indicates a shift in power from high to lower frequencies that is consistent with CSD.

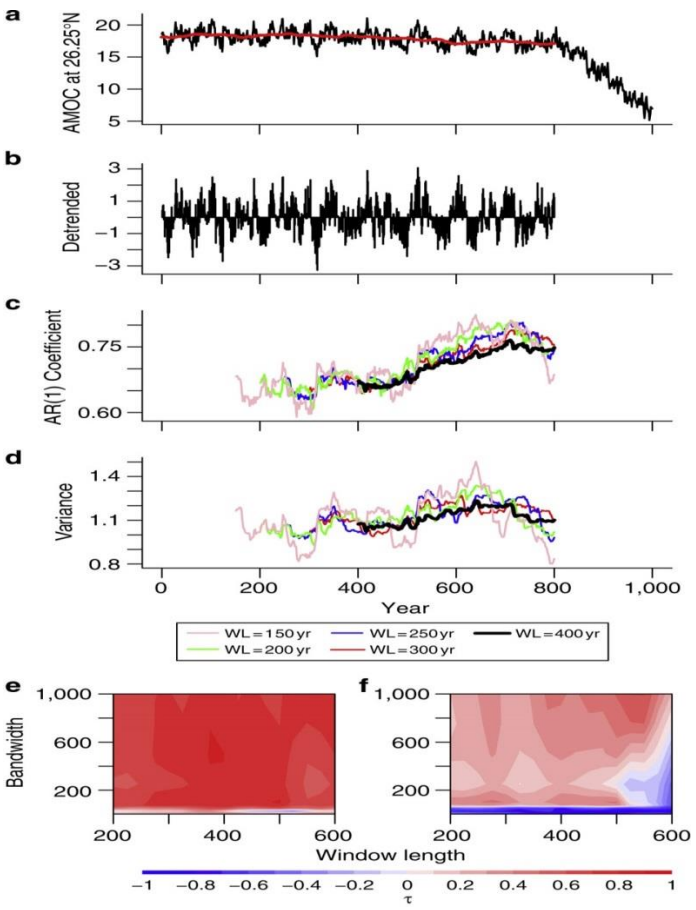

Figure 2.4: Early warning signals of AMOC collapse in the transient hosing experiment. Time series shown is from 26.25°N and ~1,000 m depth. (a) Annual time series up to red vertical line is detrended by a Kernal smoothing function (shown by smooth red line). The early warning signal analysis is carried out on (b) the residuals from this. Indicators of (c) AR(1) coefficient estimation and (d) variance are calculated as described in main text and Chapter 1 and plotted at the end of the window used to estimate them. Examples in (c) and (d) are shown for window lengths of 400, 300, 250, 200 and 150 years while using a detrending bandwidth of 100 years. Sensitivity analysis to determine how robust the indicators are to varying window length and detrending bandwidth is shown as contour plots of tendency, measured by Kendall's τ (see Methods), in the window length-bandwidth plane for (e) AR(1) coefficient estimation and (f) variance.

We next looked for early warning signals of AMOC collapse at a wide range of latitudes in the model Atlantic basin (33.75°S–58.75°N), which show varying AMOC strength (Fig. 2.5a,b, where 26.25°N is indicated by the black line). At all latitudes we find increasing AR(1) and variance (Fig. 2.5c,e), with Kendall's τ values of 0.40–0.92 and 0.39–0.85, respectively (Fig. 2.6a,b). The absolute values of the indicators vary with latitude, hence their increase can be seen more clearly when we look at their percentage change over time (Fig. 2.5d,f). There appears to be a degree of meridional coherence to the indicators based on the annual data; the most robust upward trends in AR(1) are found towards the southern boundary of the Atlantic and in the high northern latitudes (Fig. 2.6a), whereas the most robust upward trends in variance are found just north of the equator (Fig. 2.6b).

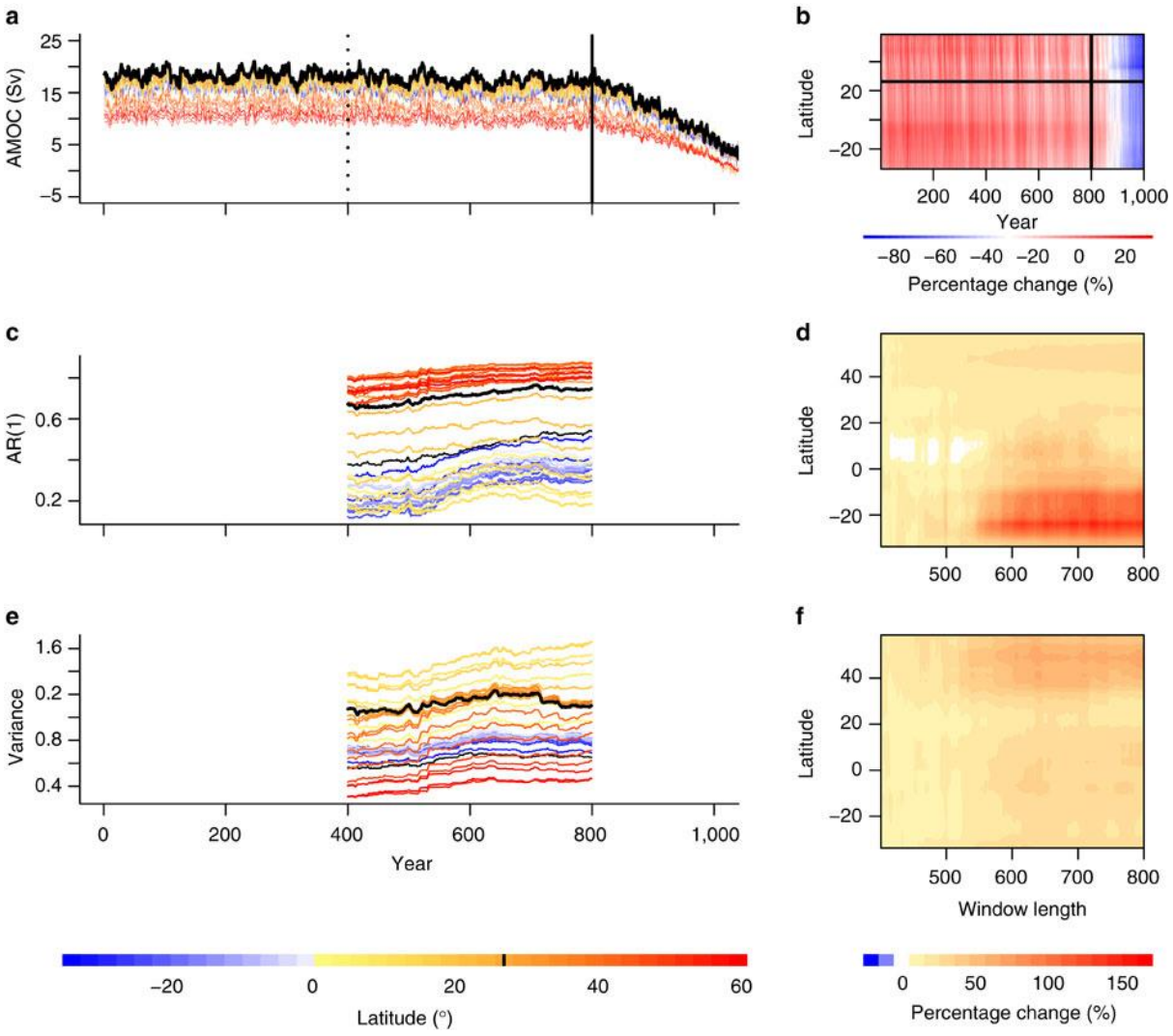

Figure 2.5: AMOC collapse and candidate early warning signals at each latitude. (a) Annual time series of AMOC (Sv) at each latitude are cut before collapse begins at 800 years (solid vertical line). (b) Time series are also shown as a contour plot in the time-latitude plane. They are then detrended and the analysis carried out on the residuals (see Methods, example in Fig. 2.4). A sliding window length of 400 years is used (dotted vertical line in (a) marks the end of the first window) to estimate candidate early warning signals (see Methods): (c,d) AR(1) coefficient, and (e,f) variance. Time series are coloured according to their latitude and 26.25°N (as in Fig. 2.4) is shown in black and contour plots of the percentage change are also shown.

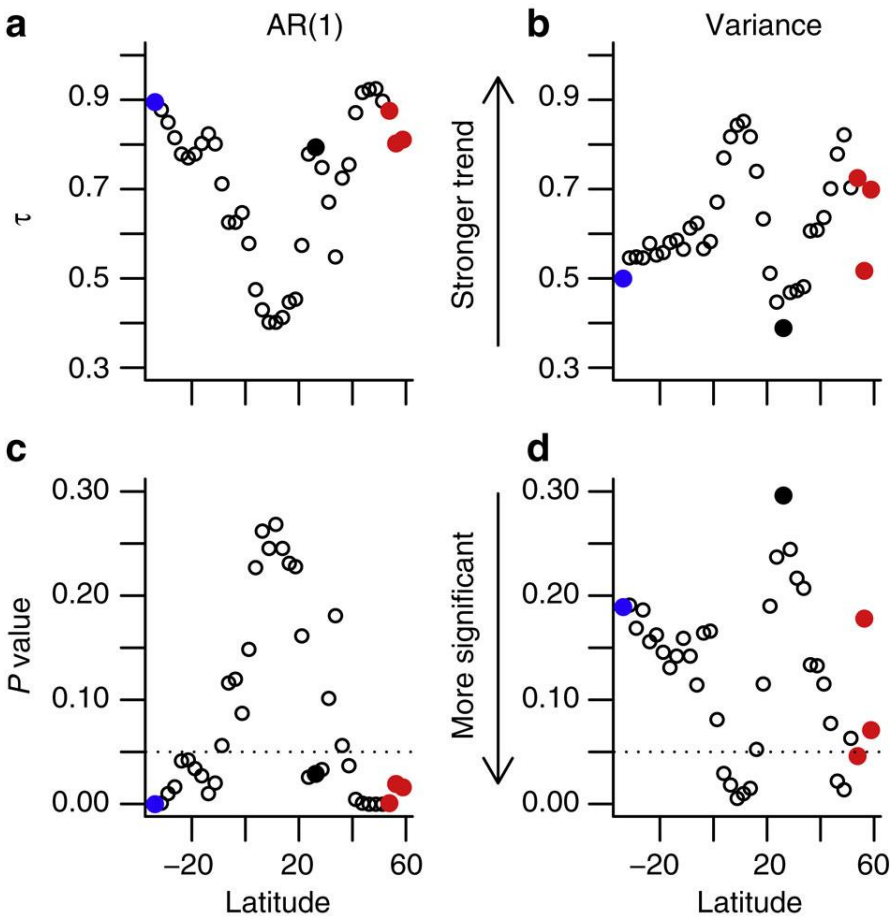

Figure 2.6: Tendency and significance of early warning indicators as a function of latitude. Kendall's τ values are calculated to determine the tendency of estimated (a) AR(1) coefficient and (b) variance indicators (see main text and Chapter 1). Significance of results, using bootstrapped null model ensembles to determine P-values (see main text and Methods), are plotted for (c) AR(1) and (d) variance. Any P-values below the dotted horizontal line are significant at the 95% level ($P<0.05$). Black filled-in points correspond to 26.25°N (where the time series analysed in Fig. 2.3 is obtained). Approximate locations of the OSNAP and SAMOC monitoring arrays are shown in red and blue, respectively.

We test the significance of the early warning signals found by comparing them to a null model. A null model ensemble of 1,000 members was created using a

bootstrapping method to create time series at each latitude that have similar characteristics—such as mean and variance—to the original time series (see Methods). Comparing the transient run to the null model, a hypothetical early warning system located at 26.25°N, near the latitude of the RAPID/MOCHA/WBTS array, provides an AR(1)-based early warning signal that is significant at the 97% level (P=0.029 at 26.25°N, Fig. 2.6c). Variance provides a less reliable early warning signal at this latitude (P=0.296 at 26.25°N, Fig. 2.6d). Looking across latitudes the most reliable early warning signals from AR(1) (Fig. 2.6c) are found in the mid-high northern latitudes (P<0.05 at 38.75–58.75°N with P<0.01 at 41.25–53.75°N) and at the southern boundary of the Atlantic (P<0.05 at 11.25–33.75°S with P<0.01 at 28.75–33.75°S), consistent with the locations where the largest values of τ are found. In the sub-polar North Atlantic, the AR(1) early warning signal is significant at the 99.9% level (P<0.001 at 51.25°N). The most reliable early warning signals from variance (Fig. 2.6d) are found in the equatorial North Atlantic (P<0.05 at 1.25–13.75°N), and in parts of the sub-polar gyre (P<0.05 at 46.25–48.75°N), again where the largest τ values were found. Overall, rising variance is a less reliable early warning indicator than increasing AR(1).

Early warning signals of AMOC collapse using decadal means

When we test for early warning signals of AMOC collapse using the decadal means of the overturning circulation (Fig. 2.7), the statistics are expected to be poorer than from annual data due to using a time series with fewer points (Dakos *et al.*, 2008). Nevertheless we still observe significant signals for both AR(1) and variance. Values of τ for variance are better than those of AR(1) with

a range of 0.58–0.94 (Fig. 2.7b) compared with 0.28–0.84 (Fig. 2.7a). However, the latitudinal pattern of significance is different to that of the annual resolution indicators. Reasonably significant AR(1) signals are found in the tropics (generally $P<0.05$ for 33.75°S–36.25°N; Fig. 2.7c) but not in the North Atlantic (41.25°N and northward), whereas significant variance signals are found in the southern tropics and North Atlantic ($P<0.05$ for 33.75°S–21.25°N and 41.25–58.75°N and $P<0.01$ for 33.75–16.25°S, 43.75–53.75°N and 58.75°N; Fig. 2.7d).

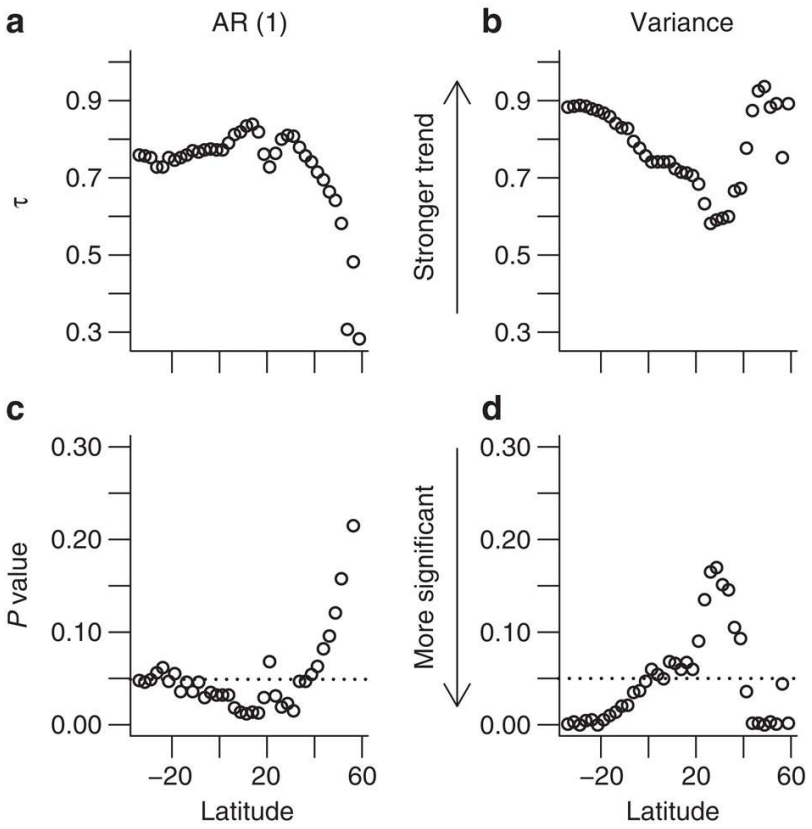

Figure 2.7: Tendency and significance of indicators for decadal resolution time series. As in Fig.2.6, Kendall's τ values are calculated to determine the tendency of estimated (a) AR(1) coefficient and (b) variance indicators (see main text and Chapter 1). Significance of results, using bootstrapped null model ensembles to determine P-values (see main text and Methods), are plotted for (c) AR(1) and (d) variance. Any P-values below the dotted horizontal line are significant at the 95% level (P<0.05).

Comparison with signals from constant forcing simulations

As a further test of the significance of early warning indicators, the results from the transient run (at annual resolution) are contrasted with results from accompanying equilibrium simulations. The equilibrium simulations are initialized from specific points of the pre-collapse phase of the transient

simulation, with the freshwater hosing maintained at a constant level (Hawkins *et al.*, 2011) until the system reaches equilibrium (black points in Fig. 2.3). In five of these equilibrium runs (at freshwater hosing values of 0.1, 0.12, 0.15, 0.2 and 0.22 Sv) the AMOC remains in the 'on' state for the full duration of the simulation (several thousand years in some cases). In another three equilibrium runs (freshwater hosing values of 0.25, 0.3 and 0.4 Sv) the AMOC eventually collapses (after hundreds of years). As none of these equilibrium runs is subject to a change in freshwater forcing, we do not expect them to show early warning signals due to there being no change in the stability of the underlying state.

For each of these equilibrium runs, trends in the estimated AR(1) coefficient and variance are obtained at each latitude (see Methods), with the results shown in Fig. 2.8. As expected from the null models (Fig. 2.6a,b) the indicators from the equilibrium runs span a range of upward and downward trends (black and grey points in Fig. 2.8a,c). However, the trends in both AR(1) and variance are higher in the transient run than those from the equilibrium runs (red points in Fig. 2.8a,c) with a mean τ=0.71 for AR(1) in the transient run compared with τ=−0.17 across the equilibrium runs and τ=0.62 for variance in the transient run compared with τ=−0.16 for the equilibrium simulations. The distribution of τ values found in the transient runs for both AR(1) and variance (Fig. 2.8b,d) is significantly different from the distribution of τ values from the equilibrium runs when using a Mann–Whitney U-test (in both cases $P<2.2 \times 10^{-16}$). These results provide further evidence that the early warning signals in the transient run are real and have not occurred by chance.

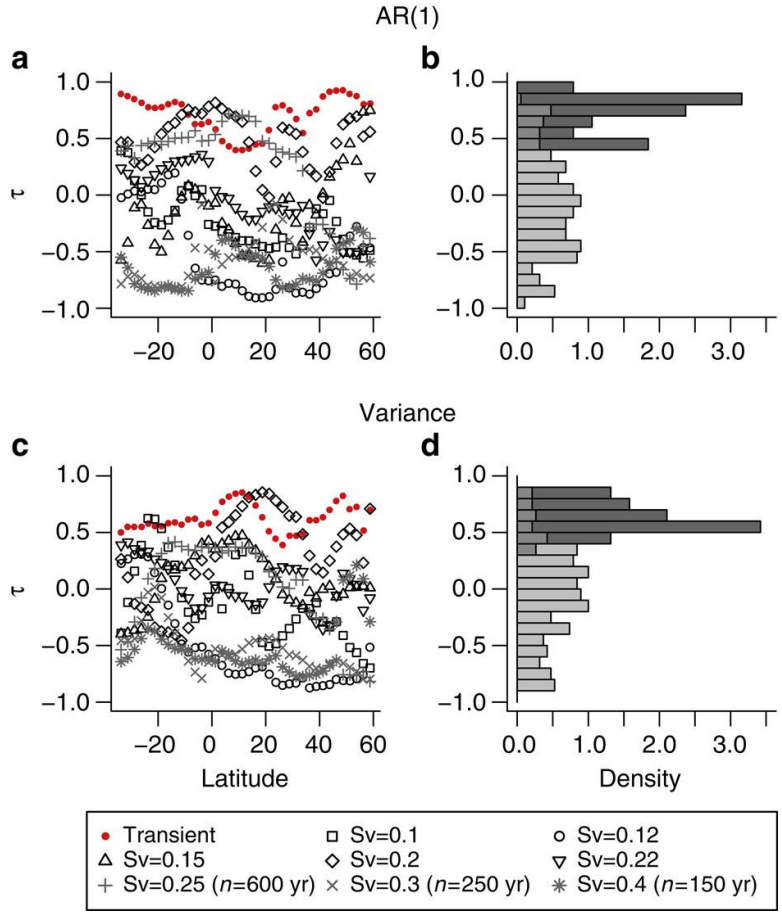

Figure 2.8: Comparing trends in early warning indicators from forced and equilibrium runs. Histograms comparing Kendall's τ values from the annual resolution transient run with those from the equilibrium runs, combining the results across all latitudes (see main text and Methods) for (a,b) AR(1) coefficient estimation and (c,d) variance indicators. (a,c) Results from the transient run are shown in red, from equilibrium runs where time series equal to transient run length (n=800 years) could be obtained are shown in black and from equilibrium runs where a tipping point occurred preventing a long enough time series are shown in grey. Results are summarized in vertical histograms (b,d) respectively. In both cases, the lighter histogram is composed of τ values from the equilibrium runs and the darker τ values from the transient run with the intermediate shading implying the histograms overlap.

Length of time for indicators to become significant

To determine the length of time series needed to observe a significant signal and thus how far in advance the approach towards the tipping point might be predicted, we extend our use of the null models by testing against them on an increasing amount of data (Fig. 2.9; see Methods). We find that the signals begin to become significant ($P<0.05$, red in Fig. 2.9) after ~550 years of simulated data, 250 years before the tipping point occurs. Using two different window lengths (50 and 400 years), we find that for both AR(1) (Fig. 2.9a,c) and variance (Fig. 2.9b,d), the indicators become significant after using very similar lengths of data and in the same latitudinal regions (those seen in Fig. 2.6c,d).

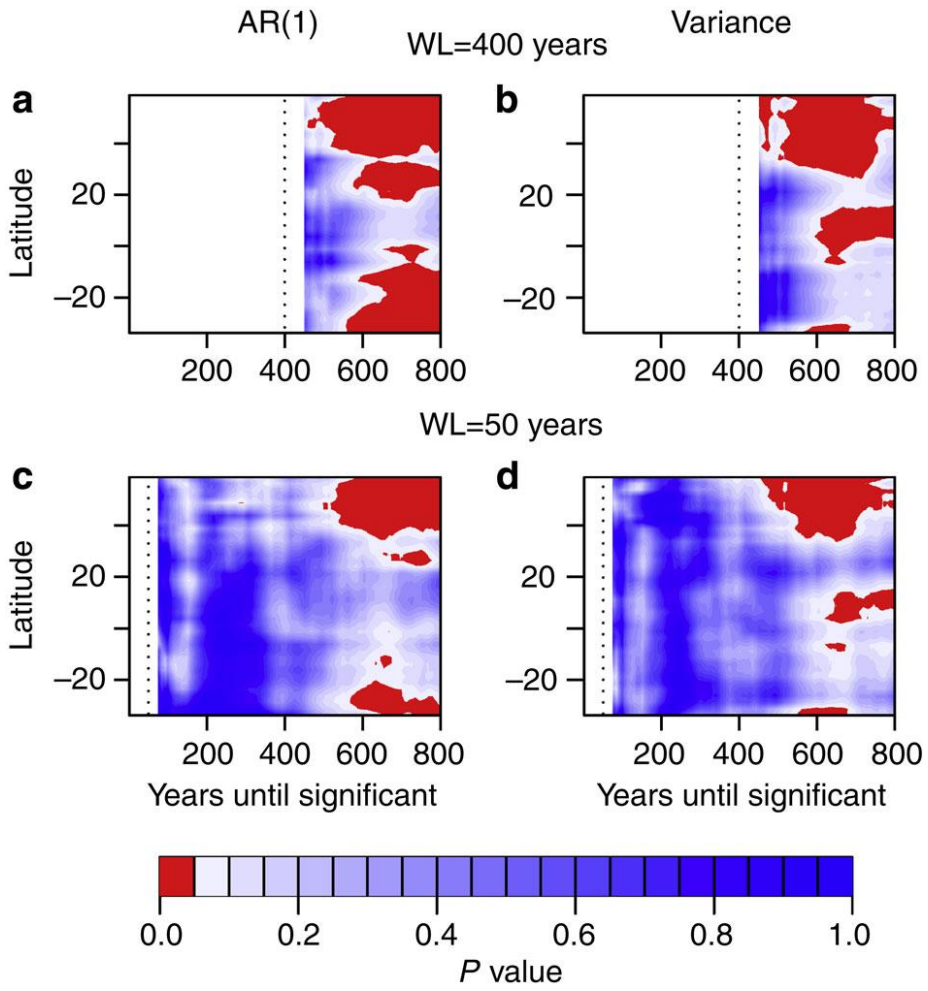

Figure 2.9: Time to significance of early warning indicators. Contour plots of significance as a function of latitude and the length of time series used are calculated as described in main text and Methods for window lengths of (a,b) 400 years and (c,d) 50 years (both shown by dotted line), using the annual resolution time series. Significance for each window length is tested after 50 and 25 years, respectively, to allow the indicators to be long enough to test significance on. Areas not shaded are where significance is not calculated due to either the length of time series at that point being less than the window length used (to the left of the dotted line), or the length of the indicator time series is less than 50 or 25 years depending on the window length used. Red shading suggests results are significant at 95% confidence ($P<0.05$) with blue shading not significant at this level.

Discussion

Our results reveal generic early warning signals for a collapse of the AMOC in a fully coupled atmosphere-ocean general circulation model: the most realistic simulation of the climate system in which this type of signal has been tested. The hosing experiment carried out in FAMOUS involved a relatively slow, linear forcing. Nevertheless, comparison of the transient simulation with the equilibrium runs (Fig. 2.3) shows that the AMOC was forced fast enough to shift it away from equilibrium, such that it lagged the forcing (that is, collapse is delayed in the transient simulation). The theory of CSD is derived for systems close to equilibrium, yet it still seems to work in this case where the timescales of the forcing and the internal dynamics of the AMOC are comparable. In reality, anthropogenic forcing of the AMOC may be faster and more non-linear than simulated here. It is believed that recent freshwater forcing, over approximately the last 50 years, has increased by 0.026 Sv (Lenton *et al.*, 2008), which is comparable to the 0.05 Sv per century increase used to force FAMOUS here. However, anthropogenic forcing may increase faster in the future. It needs to be examined whether a more realistic forcing scenario can still produce early warning signals, or whether it eliminates them, as it does for another climate tipping element (Boulton *et al.*, 2013). If anthropogenic forcing is faster than the intrinsic timescale of the ocean, then the early warning signals should not work as well as the system will not be near to equilibrium.

Our calculation of the length of time it takes for the early warning signals to become significant uses time series at annual resolution. In reality, palaeoclimate reconstruction of the AMOC would be required to gain enough

data to begin to determine if early warning signals are significant with enough time before collapse to be useful. Although existing palaeo reconstructions of the AMOC are at coarser temporal resolution than annual, our results also show that significant signals can be observed at a decadal resolution and thus potentially could also appear in these reconstructions (Dakos *et al.*, 2008). Also the AMOC will not have been subjected to anthropogenic forcing for most of the palaeo reconstruction era, which could be beneficial for observing a signal once the forcing begins. It has also been shown that the warm phase of the Atlantic multidecadel oscillation (AMO) coincides with a strengthening of AMOC and the cool phase, a weakening (Wang & Zhang, 2013) and multidecadal sea surface temperature (SST) variations are closely related to the AMOC in GCMs. The AMO has been reconstructed using SST records (including in-filling) since 1856 at a monthly resolution (Kaplan *et al.*, 1998) and using tree-ring palaeo data, it has been reconstructed at an annual resolution from 1567 (Gray *et al.*, 2004). These reconstructions could act as a proxy for AMOC to test these early warning signals on, although caution should be used when comparing mean data at annual resolution from the model and annual SST reconstructions.

Dynamical systems theory suggests that CSD occurs due to the weakening of a restoring (negative) feedback as a tipping point is approached, causing an increase in the time taken for the system to recover from perturbations. In the case of the AMOC collapse in FAMOUS, the existence of CSD signals suggests that the gradual freshwater forcing is causing a negative feedback to weaken. An important stabilizing feedback on the AMOC involves changes in meridional heat transport (Rahmstorf & Willebrand, 1995). A weakening of the AMOC leads to a reduction in northward ocean heat transport, causing a cooling of the high latitude North Atlantic and associated increase in density, which promotes

a recovery of the circulation through increased deep water formation. However, as the freshwater forcing is applied in the transient experiments analysed here, the AMOC undergoes a gradual weakening (as can be seen in Fig. 2.3) before the collapse. In contrast to salinity anomalies, surface ocean temperature anomalies are strongly damped by atmosphere-ocean heat fluxes. This means that when the AMOC is weaker, with slower northward advection of surface water masses, the increased transit time allows increased damping of the temperature anomalies, weakening the negative feedback from the AMOC itself. This theoretical explanation for the CSD has perhaps the broadest applicability, but many previous studies have provided evidence for other, more detailed, restoring feedback mechanisms that are responsible for controlling the time scale of decadal-centennial variability in coupled AOGCMs. These include an ocean-only mode excited by atmospheric variability, in which heat and salinity transport both play a role and the overturning and gyre circulations interact (Delworth *et al.*, 1993, Dong & Sutton, 2005); a coupled ocean-atmosphere mode, in which AMOC variations trigger dynamical feedbacks in the atmosphere that act to oppose the AMOC anomaly (Timmermann *et al.*, 1998); feedbacks that involve shifts in the inter-tropical convergence zone, leading to salinity anomalies in the tropical Atlantic, which feed back onto the AMOC strength as they are advected northwards (Vellinga & Wu, 2004); and feedbacks that involve links with the Arctic (Hawkins & Sutton, 2008, Jungclaus *et al.*, 2005). To properly understand the precise dynamical reason for CSD in FAMOUS, further work will be required to identify the dominant negative feedback that controls the time scale of AMOC variability in this model, and identify how it is influenced by freshwater forcing. As there is still debate about the key feedbacks that stabilize the AMOC in different AOGCMs, generalization

of these results to other models and the real world remains an important challenge.

The early warning signals in the annual resolution data are most reliable in the high northern latitudes and towards the southern boundary of the Atlantic. Current monitoring of the AMOC occurs at 26.5°N, where in this model, early warning signals are somewhat less reliable. However, there are already plans to monitor the AMOC in the sub-polar North Atlantic with the OSNAP (Overturning in the Subpolar North Atlantic Program) monitoring array (~55–60°N, red in Fig. 2.6). There are also proposals for a SAMOC (South Atlantic Meridional Overturning Circulation) array located in southern boundary of the basin (~34.5°S, blue in Fig. 2.6). Our results based on annual resolution data suggest that these could be the best locations to try to diagnose trends in the dynamical stability of the AMOC. However, the latitudinal results are rather different with decadal averaged data (Fig. 2.7) and may well vary from model to model. Thus, at this stage we can only conclude that early warning signals are likely to be latitude dependent and therefore monitoring at more than one location may increase the likelihood of observing a robust early warning signal.

Latitudinal variation in the reliability of early warning indicators might potentially be understood in terms of the latitudinal characteristics of natural AMOC variability. There are several dynamical components to the AMOC (Hirschi *et al.*, 2007, Lee & Marotzke, 1998), whose relative contributions differ with latitude and have been diagnosed in a (100-year) control simulation of HadCM3 (Sime *et al.*, 2006) (from which the FAMOUS model we use is derived). To leading order, the meridional velocity across a zonal section can be dynamically split into Ekman and geostrophic components. By definition, the Ekman

component is surface-intensified and directly driven by the zonal mean wind stress. The geostrophic component can be further decomposed into baroclinic (vertical shear) and barotropic (depth-independent) contributions. The barotropic component arises due to an interaction between sloping topography and the component of the flow that is constant with depth. In an idealized basin with vertical sidewalls, the barotropic component of the overturning circulation would be zero. In reality, variations in ocean depth across the zonal section cause vertically constant flow to project onto the meridional overturning circulation. For example, a northward depth-independent flow in a shallow part of the section (for example, near the boundaries) and a southward depth-independent flow at longitudes where the ocean is deeper, when zonally and vertically integrated, would produce a net positive contribution to the meridional overturning transport. The remaining, baroclinic, component arises through thermal wind balance associated with zonal density gradients across the basin. There are several physical mechanisms that control basin-wide density gradients, including coastal wind-driven upwelling and downwelling (Köhl, 2005), local buoyancy forcing (Hirschi *et al.*, 2007) and changes in the formation rate and transport of remote water masses. The latter is of particular relevance to the present study, as a change in the density and transport of North Atlantic Deep Water (NADW) as it spreads southwards along the western part of the basin will be reflected in a change in the zonal density gradient and therefore the vertical shear and the baroclinic component of the AMOC (Hodson & Sutton, 2012, Johnson & Marshall, 2002, Kawase, 1987, Zhang, 2010).

CSD occurs because a restoring feedback is weakening as a bifurcation-type tipping point is approached. This negative feedback involves the component of the AMOC that is thermohaline-driven and acting on multi-decadal to centennial

time scales. It is likely that this will be reflected in the baroclinic (vertical shear) component associated with zonal density gradients, as this is the component in which changes in the density and extent of NADW are likely to be most strongly visible. Within the South Atlantic and at some latitudes in the northern North Atlantic, this baroclinic component dominates the AMOC (Sime *et al.*, 2006), and we speculate that this could be consistent with the enhanced significance of the early warning signal based on annual data from these locations. In the low-latitude North Atlantic, in contrast, there is a large influence from other components of AMOC variability, particularly that associated with strong, depth independent flow over sloping topography (Sime *et al.*, 2006), which may help to explain the reduced significance of the early warning signal there. However, a clear distinction between the dynamical drivers of the AMOC components remains elusive, and several studies suggest that the baroclinic and barotropic components can be closely linked. For example, the barotropic component can also be influenced by zonal density gradients, and it is clear that wind forcing plays a crucial role in sustaining the AMOC (Timmermann & Goosse, 2004).

In addition, it remains to be established whether the latitudinal variation in relative dominance of the dynamical components in HadCM3 (Sime *et al.*, 2006) holds on longer (centennial) time scales, and in other models. Further work will be required to establish whether the latitudinal variation in early warning signal reliability exists in other models and to fully understand the dynamical reasons behind this.

While preparing this manuscript, another study was published that explores a different method for detecting early warning of AMOC collapse in the same model (Feng *et al.*, 2014). In that study, the indicators of CSD theory appear to

fail to provide early warning of AMOC collapse. However, their analysis was unusual in looking for temporal spikes in the indicators and averaging the data over latitudes.

Generic early warning indicators can complement system-specific stability indicators, such as the sign of the freshwater transport by the AMOC at the southern boundary of the Atlantic (de Vries & Weber, 2005, Hawkins *et al.*, 2011, Huisman *et al.*, 2010, Rahmstorf, 1996). This 'F_{ov}' indicator may reveal whether the AMOC is in a bistable regime and thus give some indication of whether a sudden collapse is possible. The early warning signals discussed in the present study may complement such a bistability indicator by providing information about the approach of the system towards the tipping point. Historical reconstruction of variations in AMOC strength (Balmaseda *et al.*, 2007, Wanamaker *et al.*, 2012), for example, based on fluctuations in North Atlantic SSTs (Delworth & Mann, 2000), will also be needed to establish natural variability and any trends up to the time of monitoring. Nevertheless, our results suggest that plans for new AMOC monitoring arrays could have a previously unrecognized value in helping establish whether the climate system is being pushed towards AMOC collapse.

Methods

Data

Data for the overturning are taken from the coupled climate model FAMOUS, which has been subjected to a hosing experiment, causing the AMOC to collapse. The time series consist of the annual mean meridional overturning transport at ~1,000 m depth at each latitude. We determine that (at all latitudes) the AMOC begins to collapse after 800 years and so only use data up to this point in our analysis. For the majority of our analysis, we detrend using a Kernal smoother with a bandwidth of 100 years with the early warning signals tested on the residuals. To test the robustness of the indicators at 26.25°N, we vary the window length and detrending bandwidth used (Fig. 2.4e,f).

Null models and significance

We use null models to determine the significance of the early warning signals. A null model is created for each latitude, consisting of 1,000 members. Each ensemble member is the same length as the original time series pre-collapse (800 years) and is created by a bootstrapping method, sampling from the residuals of the time series (with replacement) once it has been detrended as described above. This ensures that each ensemble member has the same statistical properties as the original time series (that is, mean and variance), but the memory of the system is destroyed. Because of this, signals are not expected to be observed in the null model allowing us to explore the probability of the signal occurring by chance.

To determine a P-value, we observe the proportion of the null model members that exhibit a higher Kendall's τ value when AR(1) and variance are tested using the same window length (400 years). The P-value is the probability of observing a signal as strong as that which we observe in the model by chance and any signal that has a P-value <0.05 is considered significant.

Time to significance

To determine the length of data needed to observe a significant signal, we use the null models as described above on a smaller section of the indicator's time series. This section is then increased, 1 year at a time while testing the significance of the indicator each time. When using a window length of 400 years to obtain the indicators, we begin to test the time it takes for them to become significant after 50 years of indicator time series, and, for a window length of 50 years, we use 25 years. The time to significance is the number of years of indicator time series needed for the P-value to be <0.05, plus the window length, giving the total number of years of data needed.

Equilibrium runs

For the equilibrium runs (see main text) that have no AMOC collapse, indicators are tested on the last 800 years of simulated data, again using a window length of 400 years and a detrending bandwidth of 100 years. This is to ensure that the influence of the previously increasing forcing is as small as possible. For equilibrium runs with AMOC collapse before the end of the run, the time series between the initialization from the transient simulation and the point of collapse

are used (600, 250 and 150 years for the three runs in which this occurs) using window lengths equal to half the length of the time series but the detrending bandwidth maintained at 100 years.

Chapter 3

Early Warning Signals of Simulated Amazon Rainforest Dieback

This chapter is based on Boulton, C. A., Good. P., and Lenton, T., M., 'Early warning signals of simulated Amazon rainforest dieback', 2013, Theoretical Ecology, 6, 3, 373-384

Abstract

We test proposed generic tipping point early warning signals in a complex climate model (HadCM3) which simulates future dieback of the Amazon rainforest. The equation governing tree cover in the model suggests that zero and non-zero stable states of tree cover co-exist, and a transcritical bifurcation is approached as productivity declines. Forest dieback is a non-linear change in the non-zero tree cover state, as productivity declines, which should exhibit critical slowing down. We use an ensemble of versions of HadCM3 to test for the corresponding early warning signals. However, on approaching simulated Amazon dieback, expected early warning signals of critical slowing down are not seen in tree cover, vegetation carbon or net primary productivity. The lack of a convincing trend in autocorrelation appears to be a result of the system being forced rapidly and non-linearly. There is a robust rise in variance with time, but this can be explained by increases in inter-annual temperature and precipitation variability that force the forest. This failure of generic early warning indicators led us to seek more system-specific, observable indicators of changing forest stability in the model. The sensitivity of net ecosystem productivity to temperature anomalies (a negative correlation) generally increases as dieback approaches, which is attributable to a non-linear sensitivity of ecosystem respiration to temperature. As a result, the sensitivity of atmospheric CO2 anomalies to temperature anomalies (a positive correlation) increases as dieback approaches. This stability indicator has the benefit of being readily observable in the real world.

Introduction

In recent years, research into the field of tipping points and their predictability has yielded several suggestions for generic early warning signals of an approaching bifurcation-type tipping point (Lenton, 2011, Scheffer *et al.*, 2009). The most general behaviour of a dynamical system approaching a bifurcation (where an attractor loses its stability) is that it becomes more sluggish in its recovery from short-term fluctuations (Wissel, 1984). This is termed 'critical slowing down', and occurs because negative feedback in the system (which keeps it in a given attractor) begins to be overwhelmed by positive feedback (which can propel a transition between attractors)—or more mathematically speaking, the leading eigenvalue governing the decay rate of fluctuations tends toward zero (from a negative value). This critical slowing down behaviour should manifest itself as increasing autocorrelation in time (and possibly space), which can be readily measured. It is also generally expected to cause a rise in variance (Carpenter & Brock, 2006), which requires less data to detect a signal. However, there are special conditions under which rising variance does not occur, or cannot be detected (Dakos *et al.*, 2012).

To date, these proposed generic tipping point early warning indicators have been tested in palaeo-data approaching past abrupt climate changes (Dakos *et al.*, 2008, Livina & Lenton, 2007), and in simple and intermediate complexity climate models approaching forced tipping points (Lenton, 2011), but not in the full-complexity models used for climate projections, e.g. by the Intergovernmental Panel on Climate Change. Furthermore, existing model tests of early warning indicators have generally concentrated on the case study of a

collapse of the thermohaline circulation of the Atlantic, and they have used very slow forcing relative to the rate at which humans are interfering with the climate system (Held & Kleinen, 2004).

Here, we set about to test generic early warning indicators of an approaching tipping point in a complex climate model, forced in a realistic way, which exhibits an iconic example of a potential tipping point response to climate change; dieback of the Amazon rainforest. The model we use is the Hadley Centre climate model version 3, known as 'HadCM3'. Amazon dieback was first predicted when an offline vegetation model was forced with climate change from HadCM3 (and its predecessor HadCM2) (White *et al.*, 1999). Then dieback was found in the fully coupled but lower resolution HadCM3LC model under future forcing (Cox *et al.*, 2000). The equilibrium behaviour of HadCM3LC was later found to have an even stronger non-linear response of Amazon forest cover to temperature (Jones et al. 2009). Unlike reality, 57 different versions of HadCM3 now exist with different settings of key physical model parameters and therefore different climates (Lambert *et al.*, 2013). We show an example of Amazon dieback under future forcing (with the SRES A1B emissions scenario) in one of these model versions in Fig. 3.1. First, there is a decline in net primary productivity (NPP) (Fig. 3.1a), then vegetation carbon (Fig. 3.1b), then broadleaf tree fraction (Fig. 3.1c), the latter beginning around 2060. Below, we make use of all 57 model versions to study multiple realisations of approaching Amazon dieback, and thus begin to examine the statistical reliability of proposed early warning indicators.

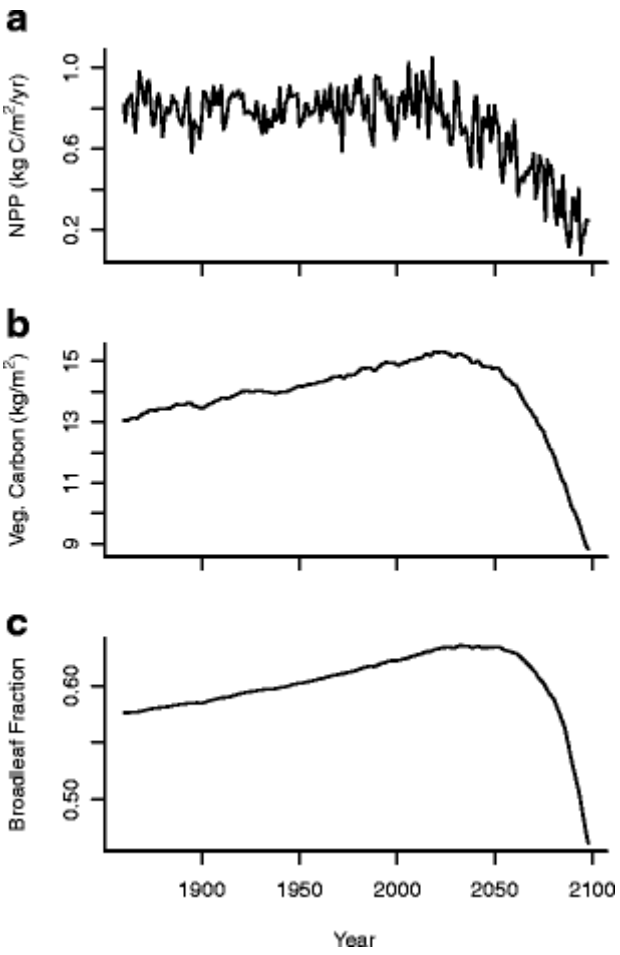

Figure 3.1: An example of Amazon dieback simulated in one of the 57 versions of HadCM3 showing the full time series (1860–2100) of (a) net primary productivity (NPP), (b) vegetation carbon and (c) broadleaf fraction, averaged annually over the region.

Amazon dieback presents an important case study for several reasons. The Amazon rainforest is a critical component of the global carbon cycle, acting as a large store of carbon and typically a significant carbon sink, with the notable exception of recent drought years when it switched to become a carbon source (Lewis *et al.*, 2011, Phillips *et al.*, 2009). The Amazon has been identified as a potential tipping element in the Earth's climate system (Lenton *et al.*, 2008),

partly because alternative attractors for the vegetation–climate system of the region are thought to exist (Oyama & Nobre, 2003, Salati & Vose, 1984), and partly because in some future simulations, a large fraction of the Amazon dies back fairly abruptly (within decades) (Cook & Vizy, 2008, Cox *et al.*, 2004, White *et al.*, 1999). Experts gave an average 20% chance of tipping the Amazon (at least half of its current area is converted from year-round forest due to climate change) if global warming is between 2 and 4 °C by 2200, and a 70% chance if warming exceeds 4 °C (Kriegler *et al.*, 2009). However, the future of the Amazon rainforest is highly uncertain. The observational record shows a lengthening of the dry season, attributed to anthropogenic forcing altering the Walker circulation of the atmosphere in the tropics (Vecchi *et al.*, 2006). In HadCM3, this drying trend continues into the future, and together with warming, this overwhelms the tendency of rising atmospheric CO_2 to protect the forest by increasing the efficiency of photosynthesis (Cox *et al.*, 2004). However, future projections with other general circulation models of the climate (GCMs) give very different precipitation trends over the region in the future (Li *et al.*, 2006). Hence, the change in climate and vegetation of the Amazon in HadCM3 is an extreme result among existing models, with a transition to seasonally dry forest in the Eastern Amazon region considered more realistic (Malhi *et al.*, 2009).

Whilst testing generic early warning signals on the Amazon rainforest is new, there have been previous attempts to better understand the Amazon's vulnerability in terms of observable variables, such as sea surface temperatures (SSTs) in the tropical Pacific (Cox *et al.*, 2004) or the North Atlantic (Cox *et al.*, 2008). The vulnerability of modelled tropical forest cover has been assessed with respect to temperature and dry-season length (DSL) (Good *et al.*, 2011). An approximately linear boundary in the temperature–DSL plane separates

forested from un-forested tropical (20°N–20°S) grid cells. This boundary exists in both the HadCM3LC (Good *et al.*, 2011) and HadGEM2-ES (Good *et al.*, 2013) climate models, which both include the Top-down Representation of Interactive Foliage and Flora Including Dynamics (TRIFFID) vegetation model. Looking across a wider range of models, inter-annual anomalies in atmospheric CO_2 due to anomalies in tropical land carbon storage can be related to anomalies in tropical temperature, due to, e.g. drought or El Niño events (Cox *et al.*, 2013).

For dieback of the Amazon to display tipping point behaviour and corresponding early warning signals, there must be positive feedback in the dynamics of forest loss. Furthermore, for early warning signals to show up in a model study, these positive feedbacks must be captured in the model. A key positive feedback that is (to varying degrees) captured in existing climate models, including HadCM3, is between vegetation and rainfall. Essentially, the forest recycles water to the atmosphere through transpiration and this promotes further precipitation, which supports the forest (Betts, 1999, Salati & Vose, 1984). With the prevailing wind travelling inland, this promotes the existence of forest in parts of the Amazon basin farthest from the Atlantic coast. This positive feedback may be strong enough to produce alternative stable states of vegetation cover in parts of the Amazon (Oyama & Nobre, 2003). However, such bi-stability is not a necessary condition for the existence of early warning signals (Kéfi *et al.*, 2013, Lenton *et al.*, 2008).

It has not been previously established whether the HadCM3 model exhibits bi-stability of vegetation cover in parts of the Amazon region. We examine the governing equation for tree cover below, which suggests there are two

equilibrium solutions even without coupling to the climate. Consistent with this, we find evidence of bi-modality of tree coverage in the full climate model, when looking across different Amazon grid cells (Fig. 3.2). The state without trees is typically dominated by C3 grasses. The two states could arise purely from the competition dynamics in the vegetation model, but they may be reinforced by vegetation-rainfall feedback. Regardless of this, even if there was only one equilibrium solution for tree cover in the model, it is already known that Amazon tree cover (averaged across the region) shows a strong non-linear response to temperature (Jones *et al.*, 2009). Furthermore, such a strong non-linear response should mean that generic early warning signals of approaching dieback are present, as long as the model is forced slowly and subject to low amplitude stochastic variability (such that it remains close to its equilibrium behaviour) (Kéfi *et al.*, 2013). A potential caveat here is that the model Amazon system exhibits inertia such that it lags the climate forcing by several decades. This raises the question of whether proposed early warning indicators might reveal the equilibrium (committed) behaviour of the forest rather than the transient (observed) change, or whether they may fail altogether.

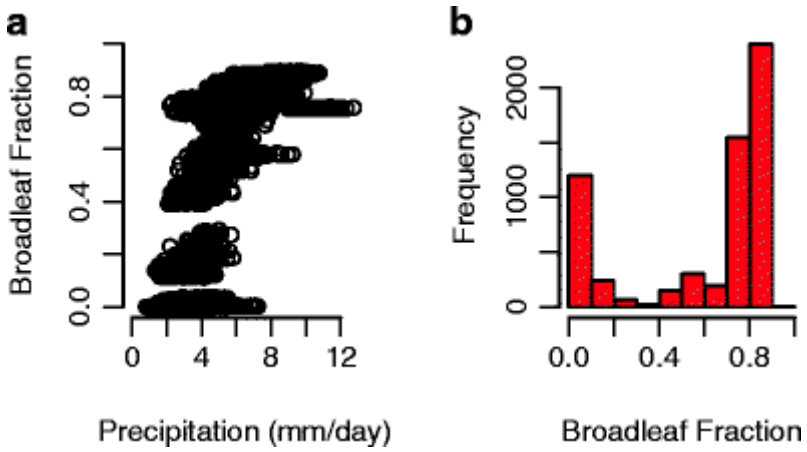

Figure 3.2: Bi-modality of tree cover looking across Amazon grid cells in the same version of HadCM3 as in Fig. 3.1: (a) broadleaf fraction as a function of mean annual precipitation for all Amazon grid cells over the first century of the run (61 points each with 100 years), (b) a histogram of the distribution of tree cover at the grid cell scale when averaging over variations in precipitation.

Data and methods

HadCM3 Earth system ensemble

Our data was obtained from a ‘perturbed physics’ ensemble of versions of the HadCM3 model called HadCM3-ESE (Earth system ensemble). The ensemble contains 57 members (i.e. model versions) where key parameters have been perturbed within boundaries suggested by experts (Lambert et al. 2013). These parameters are grouped according to their role in the Earth system, whether they are within the atmosphere (n = 32 parameters) (Collins *et al.*, 2011), ocean (n = 15) (Collins *et al.*, 2007), sulphur cycle (n = 8) (Lambert *et al.*, 2013) or carbon cycle (n = 8) (Booth *et al.*, 2012). Each of the 57 ensemble members contains a combination of changes to these four subsystems, determined by a Latin hypercube sampling process to maximise the spread of atmosphere and carbon cycles used. We are restricted to the data that has already been saved from these existing model runs as it is extremely computationally expensive to rerun the model.

The different ensemble members (i.e. model versions) are all subject to the same forcing scenario spanning 1860–2100, with historical forcing up to 2000 and the Special Report on Emissions Scenarios A1B scenario thereafter (Nakicenovic *et al.*, 2000). The forcing comprises emissions of carbon dioxide, other greenhouse gases and aerosols, with the model determining their concentrations and the resulting climate effects interactively. A1B can be viewed as a ‘middle of the road’ scenario with an economic rather than an environment focus and a balanced energy usage, as opposed to being fossil

fuel intensive for example. The resulting 57 different model runs behave differently thanks to the perturbations to the physics between model versions.

We define the Amazon region in the model as 40–70°W and 15°S–5°N, which comprises 61 land grid cells. In a particular ensemble member, some of these grid cells may not be forested in 1860, but that allows us to capture forest growth in these regions should it occur later on in the time series. Within a single chosen ensemble member (shown in Figs. 3.1 and 3.2), we examined the behaviour of key variables in each of the 61 individual grid cells. When looking across the whole ensemble of model versions, we averaged over this spatial information, on an annually averaged timescale, to produce a single time series for each variable of interest in each of the 57 ensemble members. Comparing the two approaches allowed us to check whether averaging over the whole region affected our results, compared to using grid points individually, noting that the conditions that destabilise the Amazon rainforest are unlikely to be uniform across the region.

Three key output variables from the land surface scheme and the TRIFFID dynamic global vegetation model (Cox, 2001) are the focus of our time series analysis: broadleaf fraction (BL), vegetation carbon (VC) and NPP. Two 'driver' time series over the Amazon region, temperature and precipitation, are also analysed. In our search for more process-based indicators, we also consider net ecosystem productivity (NEP), by subtracting the soil (heterotrophic) respiration flux from NPP, and atmospheric CO_2 (over the Amazon region), which contains both a long-term, global forcing trend and short-term (inter-annual) variability that reflects changes in NEP of the region.

For each ensemble member, we examined the time series of BL to see if there was a decline (often following a steady rise). If this was seen, we considered the inflection point at which BL starts to decline as the beginning of dieback in the model and only use data up to that point in our main analysis. Across the 57 ensemble members, 31 show dieback starting before 2100. We cannot rule out that dieback will occur later in the other 26 ensemble members, particularly given that committed change of the forest is much greater than transient change in this model. All 57 members are used despite the fact that some models only have small forest coverage, since any amount of forest could exhibit dieback. Models that do not exhibit dieback give time series which are 240 points long, whereas those that exhibit dieback have an average time series of 204 points on which to test for early warning signals.

For the single ensemble member where we consider spatial information, we first determine if there is enough forest in each grid point in the Amazon region to be considered for analysis. If there is a broadleaf fraction less than 0.1 at the start of the time series which does not show any growth, then we ignore this grid point. This leaves 49 (of 61) grid points that contain sufficient forest, and in 43 of these 49 grid points, there is dieback under future forcing. In this case, the average time series length for analysis is approximately 215 points.

Early warning indicators

We use a kernel smoothing function with a bandwidth of 10 years or points (Dakos *et al.*, 2008) to detrend the time series of BL, VC and NPP. Then we use a sliding window length of half the time series (prior to dieback, if it starts) in which we derive AR(1) and variance as potential early warning indicators. We

also derive skewness in the same sliding window but from the original rather than the detrended data (Guttal & Jayaprakash, 2008). Increasingly positive or increasingly negative skewness can be considered early warning signals because the sign of skewness depends on the position of the attractor being approached. In the analysis of these three time series, we expect to observe skewness becoming more negative over time.

To express trends in the indicators, we use Kendall's τ rank correlation coefficient (Dakos *et al.*, 2008), which ranges between 1, for an indicator that is always increasing, to −1 for an indicator that is always decreasing. A τ of 0 implies there is no net trend in the indicator and that it increases as much as it decreases.

To provide a null model of the behaviour of these generic indicators under stable boundary conditions, we make use of the control runs of each ensemble member. There were ~70 years of control run available for each member where the forest was stable and no emissions were imposed.

We calculate AR(1), variance and skewness on these samples using the methods described above. Then we compare these to the indicator trends observed from a comparable length time series prior to dieback in each model (or the last ~70 years if no dieback occurs).

We also test two process-based, system-specific stability indicators. The first of these assesses how the sensitivity of NEP to temperature anomalies changes over time. The second looks for changes in the sensitivity of atmospheric CO_2 variations of the Amazon region to temperature anomalies. These indicators are motivated by the idea that variations in tropical land carbon storage are caused by tropical temperature anomalies. Respiration is prescribed as an exponential

function of temperature whereas photosynthesis is a peaked function of temperature (which allows the possibility that an increase in temperature could cause decreases in photosynthesis if the optimum temperature has been passed). Therefore at higher temperatures, a given increase in temperature should give rise to a greater decrease in NEP and a correspondingly larger addition of CO_2 to the atmosphere.

To calculate these indicators, we detrend the time series of NEP or CO_2 and temperature again using a kernel smoothing function with a bandwidth of 10 years. Then within a sliding window length of 25 years, we estimate the gradient of the best fit (linear regression) line of NEP or CO_2 as a function of temperature, and use the result as an indicator. We use a smaller window length than in other analyses to better capture the effect of events such as El Niño.

Broadleaf fraction model

We use a simplified version of the TRIFFID model (Cox 2001) to better understand the dynamics:

$$\frac{dV}{dt} = P\hat{V}(1 - V) - GV \qquad (3.1)$$

Where *V* is equal to the broadleaf fraction, *G* is a disturbance coefficient (0.004/year) and $\hat{V}$ is either the value of *V* or 0.1 if *V* falls below 0.1. *P* is the productivity, in dimensionless area fraction units. There is a non-linear response

of the broadleaf fraction (*V*) to changing productivity (*P*), with one equilibrium at *V* = 0 and another equilibrium solution:

$$V^* = 1 - GP \qquad (3.2)$$

The equilibrium *V ** has an eigenvalue of *G-P*, which is negative for typical values of *P* found in the ensemble members. As *P* is reduced, the movement of the equilibrium is non-linear as the eigenvalue approaches zero. Eqn. 3.1 can be rearranged to the normal form of a transcritical bifurcation and it has been confirmed elsewhere that such bifurcations exhibit generic early warning signals (Kuehn, 2011). The two stable states observed here should translate into the full GCM version used in the ensemble although complicated due to the calculation of *P*. The $\hat{V}$ parameter in Eqn. 3.1 prevents the vegetation from becoming negative. However at values of productivity *P* we use to test the model, this does not alter the fact we are approaching the bifurcation and so should observe early warning signals.

To explore how this model behaves in conditions observed in the full climate model, we ran three 500-member ensembles of the simplified model of tree cover (Eqn. 3.1): (1) a 'null' model ensemble where there is no forcing on *P* (*P* = 0.9) and a constant noise level (σ = 0.003), (2) a 'linearly forced' ensemble where *P* is reduced linearly from 0.9 to 0.004 (the value of *G* and hence the transcritical bifurcation point) whilst keeping the same noise level, and (3) a 'non-linearly forced' ensemble where *P* is kept constant for the first 180 years at 0.9 and then reduced linearly from 0.9 to 0.004 over the final 60 years and

noise level is increased from $\sigma = 0.001$ to $\sigma = 0.006$ linearly across the time series to mimic the increase in variance in temperature and precipitation observed in the full model. This third ensemble is an attempt to recreate the conditions seen in the full climate model. In all cases, we analysed the first 200 years of each run, because in cases (2) and (3), dieback never begins before this and generally just after.

Results

HadCM3 Earth system ensemble

In the example model run, dieback starts to occur in the second half of this century around 2060 (Fig. 3.1). Prior to this, when looking across spatial locations over the first 100 years of the model run, we see evidence of two states for broadleaf tree cover, with one mode around 0.8 and another at 0–0.2 (Fig. 3.2). When analysing the results from individual spatial locations within this model run, the majority show dieback, but they do not consistently show the signal of critical slowing down (i.e. rising AR(1) and rising variance) prior to dieback. Typically, there are decreases in AR(1) for broadleaf fraction and vegetation carbon (Fig. 3.3a, d) and a tendency toward increases in AR(1) for NPP (Fig. 3.3g). All three variables typically show increases in variance (Fig. 3.3b,e,h). There are predominantly negative skewness trends observed for broadleaf fraction and vegetation carbon (Fig. 3.3c,f), but no clear skewness trends for NPP (Fig. 3.3i).

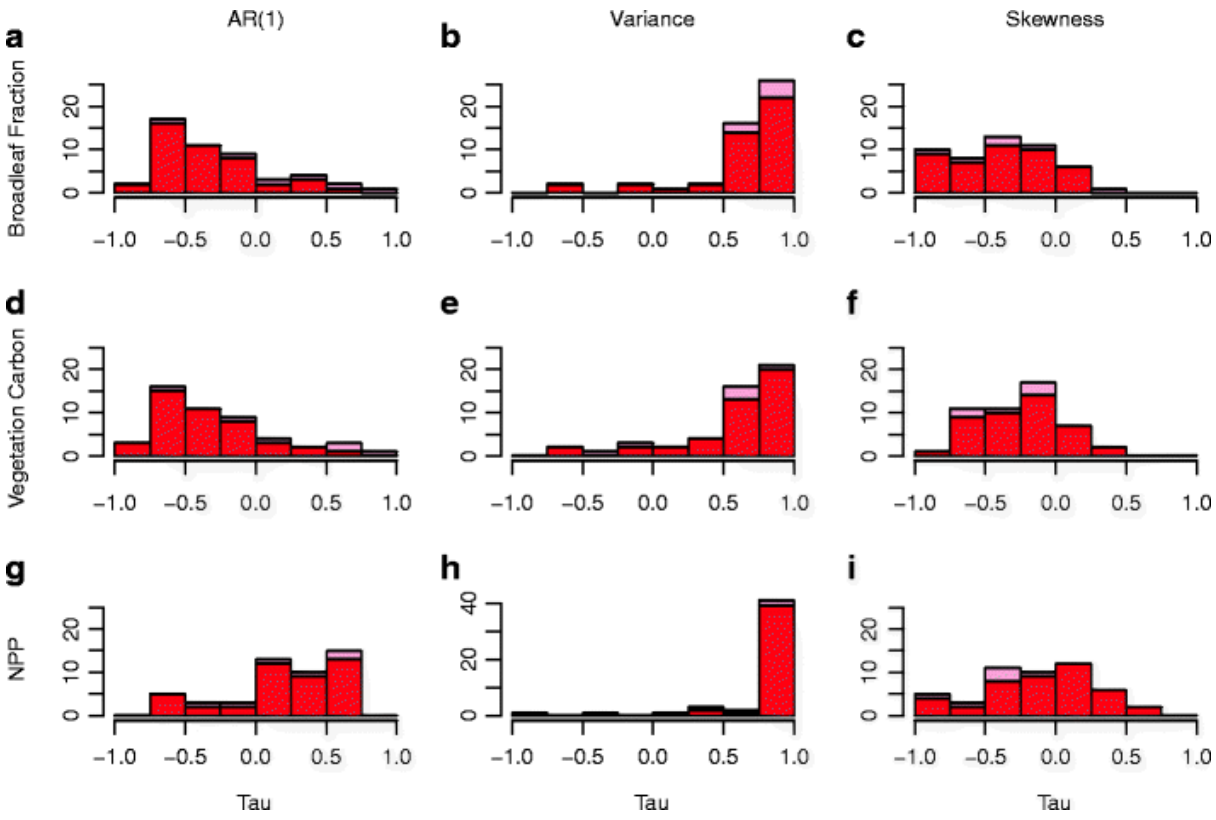

Figure 3.3: Trends in generic early warning signals across different grid cells of the version of HadCM3 in Figs. 3.1 and 3.2. Annually averaged time series for broadleaf fraction (a–c), vegetation carbon (d–f) and NPP (g–i) are used for each spatial grid point where there is sufficient forest (grid points which have a broadleaf fraction less than 0.1 at the start of the time series without showing growth are ignored). Note that in (h), the y-axis has been extended due to the majority of the tau values being in the last bin. In each histogram, the darker bars refer to those grid cells which show dieback starting prior to 2100 whereas the (stacked) lighter bars are the cells without dieback before 2100.

We analyse the spatially averaged behaviour of this example model run (as in Fig. 3.1) in Fig. 3.4. One reason why early warning signals might fail in this and other cases, based on Eqn. 3.1, is that NPP does not start to decline until shortly before dieback begins (where the data is cut off for most of the

analyses). Hence in this instance, we analyse the full time series of broadleaf fraction (Fig. 3.4a,b), vegetation carbon (Fig. 3.4f,g) and NPP (Fig. 3.4k,l), showing where NPP begins to decline around 2020 (vertical line in the time series of Fig. 3.4). Once NPP starts to decline, critical slowing down in tree cover would be expected according to Eqn. 3.1. Consistent with this, broadleaf fraction shows a rise in AR(1) and variance and a negative trend in skewness (Fig. 3.4c,d,e). Vegetation carbon shows an overall decline in AR(1), rise in variance and negative trend in skewness (Fig. 3.4h,i,j). NPP shows no clear trend in AR(1), rising variance and a trend to negative skewness (Fig. 3.4m,n,o). However, the encouraging results for broadleaf fraction should be treated with caution as the data set is only one sample of a wide range of signals that the different versions of HadCM3 can produce. Furthermore, in the real world, we are interested in warning signals before dieback begins.

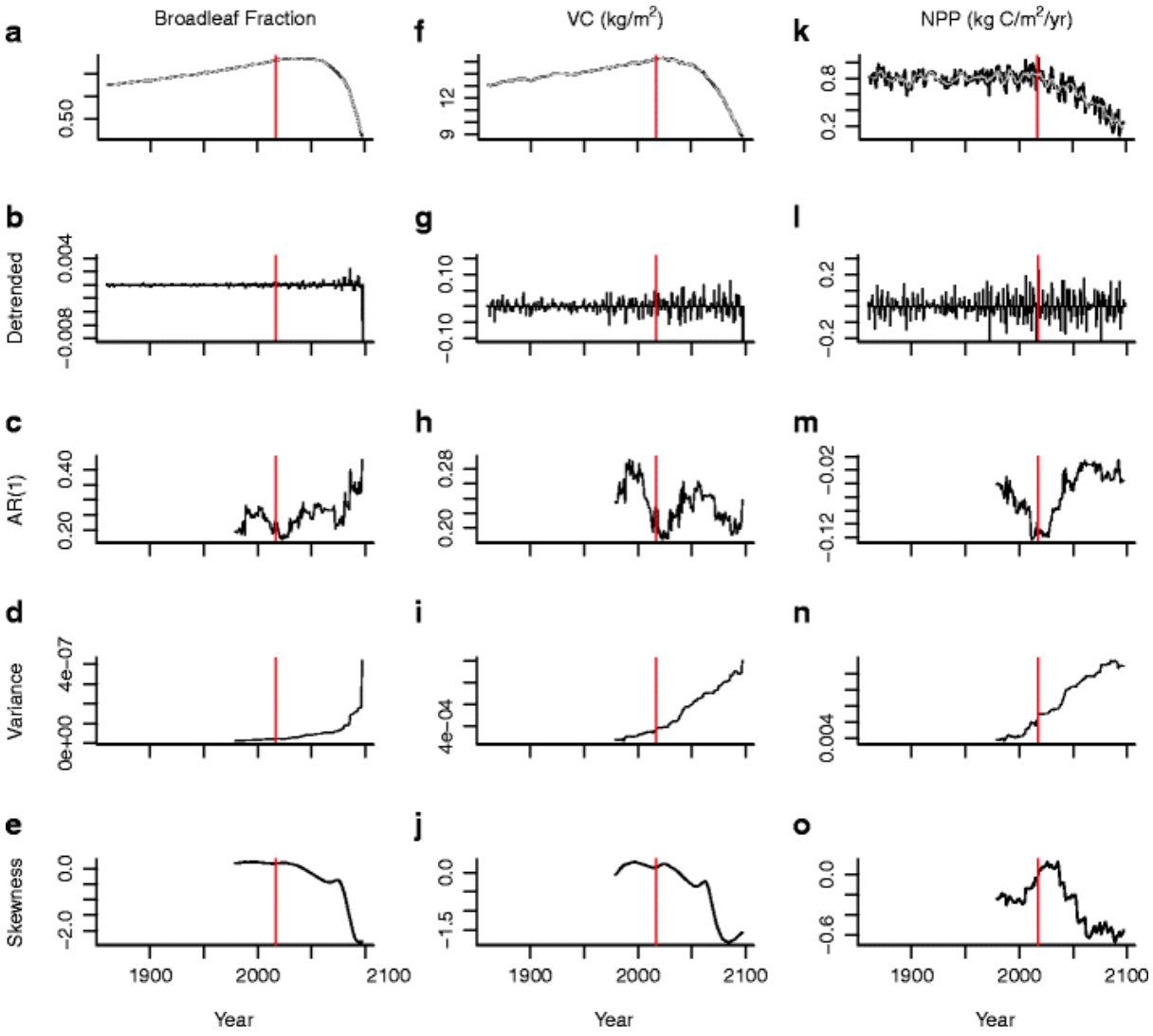

Figure 3.4: Early warning signals observed for the three key output variables from the example time series shown in Figs. 3.1–3.3. Spatial averages of broadleaf fraction (a–e), vegetation carbon (f–j) and NPP (k–o) are analysed for the full-time series rather than up to observed dieback, in order to examine the full effect of NPP decline that begins around 2020. The start of this decline in productivity is shown in all the time series with a vertical line. The smoothing time series used to derive the residuals (the detrended time series) are shown in the top row (a, f and k) over the original time series.

Looking across the ensemble of model versions using spatially averaged data, generic early warning indicators do not show the signal of critical slowing down (i.e. rising AR(1) and rising variance) prior to the start of Amazon dieback. Fig. 3.5 shows histograms of Kendall τ results for all the ensemble members when testing the AR(1) coefficient, variance and skewness as indicators on the broadleaf fraction, vegetation carbon and NPP time series. AR(1) of broadleaf fraction, vegetation carbon and NPP typically decline, but results with no trend or some increase are also found (Fig. 3.5a,d,g). The strongest result is rising variance in all three variables (broadleaf fraction, vegetation carbon, NPP) (Fig. 3.5b,e,h), although occasional downward trends are seen. Broadleaf fraction and vegetation carbon are typically increasingly negatively skewed over time (Fig. 3.5c,f), whereas there is a hint of NPP becoming increasingly positively skewed over time (Fig. 3.5i). In all individual cases examined, there is a switch of sign of skewness of NPP over time. Differences between the results for ensemble members that show Amazon dieback starting before 2100 (dark histograms in Fig. 3.5) and those that do not (lighter stacked histograms) are modest. There is a slightly stronger tendency of decreasing AR(1) in models with no dieback (Fig. 3.5a,d,g) and arguably trends in skewness are slightly stronger (Fig. 3.5c,f,i).

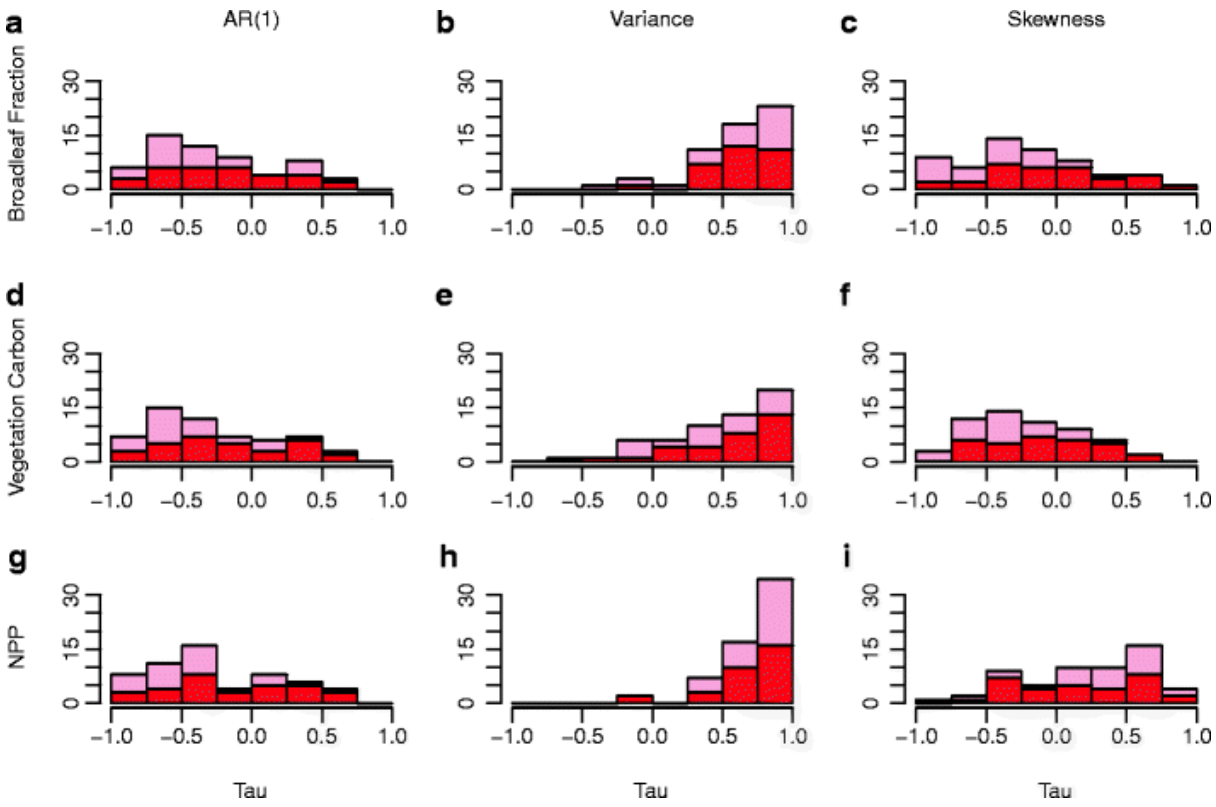

Figure 3.5: Trends in proposed generic early warning indicators across the ensemble of 57 models, from analysis of: (a–c) broadleaf fraction, (d–f) vegetation carbon and (g–i) NPP time series averaged over the Amazon region. Trends in AR(1), variance and skewness are expressed as Kendall τ values. The resulting histograms give an indication of range of indicator trends across the ensemble and their robustness. In each histogram, the darker bars refer to those models which show dieback starting prior to 2100 whereas the (stacked) lighter bars are the runs without dieback before 2100.

When compared to the control runs where no forcing is imposed on the system, which we are treating as a null model (Fig. 3.6), the most prominent change in the indicators are an increase in variance in broadleaf fraction, vegetation carbon and NPP (Fig. 3.6b,e,h). The distribution of trends in AR(1) (Fig. 3.6a,d,g) and skewness (Fig. 3.6c,f,i) do not appear significantly different under

climate forcing to the control runs of the models. However, we stress here that the time series are short and longer time series would be needed to attempt a statistical test on the significance of the results.

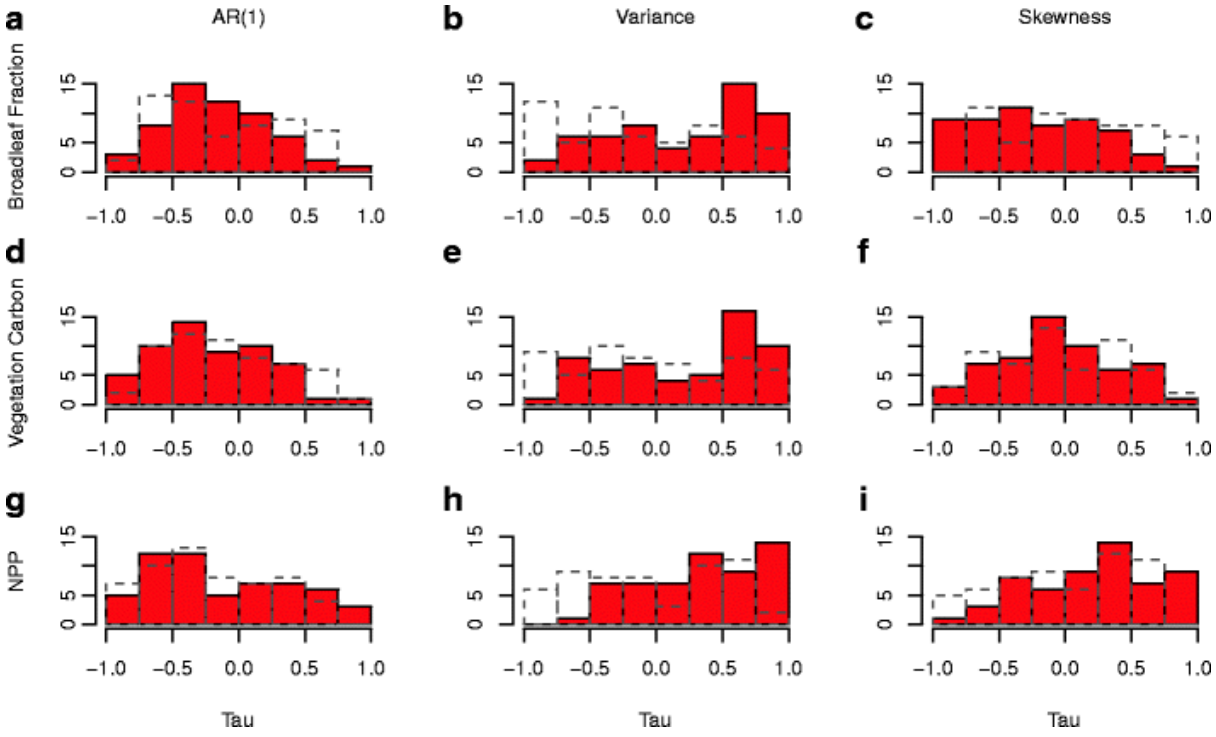

Figure 3.6: Trends in generic early warning signals for time series found in control versions of each ensemble member (without forced under an emissions scenario) (dotted histograms) compared to the early warning signals found in the original ensemble members using the same length time series as in the control run time series (~70 years). Broadleaf fraction (a–c), vegetation carbon (d–f) and NPP (g–i) time series were averaged over the Amazon region each year.

The indicator trend of increasing variance that we do consistently observe could be due to corresponding trends in the environmental variables forcing the forest

(Fig. 3.7). Indeed, we find that temperature generally shows strongly rising variance (Fig. 3.7b), together with a tendency toward declining AR(1) (Fig. 3.7a), and perhaps a slight shift to negative skewness (Fig. 3.7c). Precipitation also tends to show increasing variance or in a smaller number of cases decreasing variance (Fig. 3.7e), but no clear trend in AR(1) (Fig. 3.7d) or skewness (Fig. 3.7f). The increases in variance of temperature and precipitation can be mostly attributed to the model producing more frequent or extreme El Niño events under climate forcing, although we should note that temperature and precipitation are not pure external forcing variables—they are also affected by feedback from the forest. We also note that a strong enough increase in variance (as seen for temperature) can cause a decrease in AR(1) because it makes neighbouring points in a time series less alike.

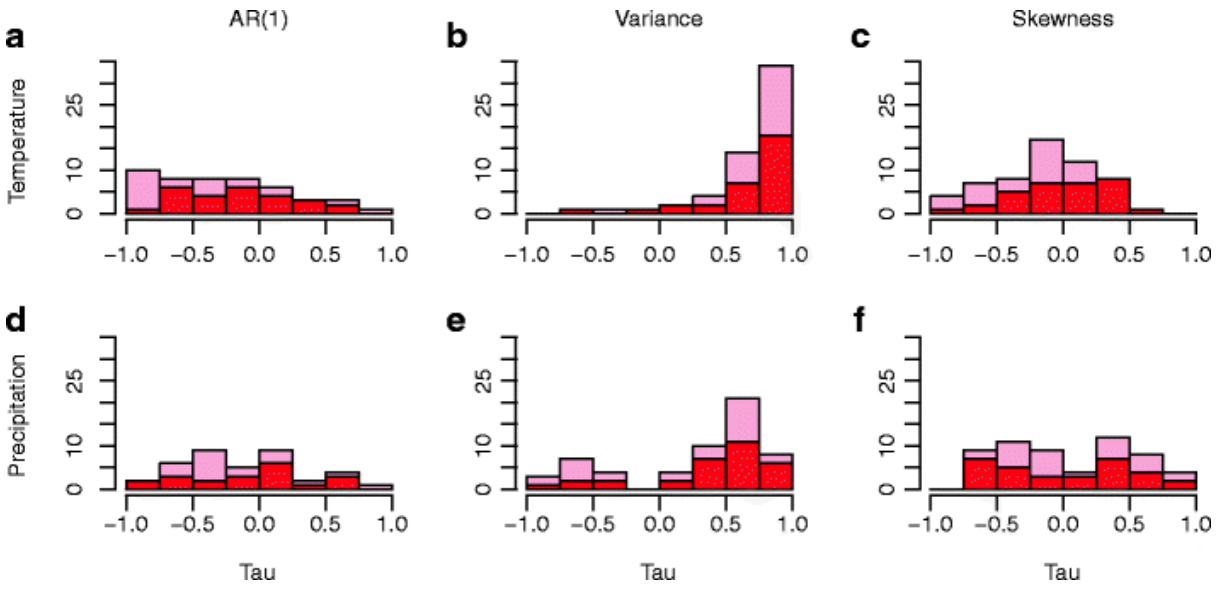

Figure 3.7: Trends in the environmental variables of (a–c) temperature and (d–f) precipitation averaged over the Amazon region, across the ensemble of 57 models. Again trends in AR(1) estimation, variance and skewness are measured as Kendall tau values with the darker bars of each histogram referring to ensemble members in which Amazon dieback starts before 2100 and the (stacked) lighter bars those in which it does not.

Considering more process-based indicators of forest stability, we find an expected negative correlation between (inter-annual) changes in temperature and (inter-annual) changes in NEP—in other words, warming suppresses NEP (Fig. 3.8a)—and this explains why CO_2 anomalies are positively correlated with temperature (Fig. 3.8b)—as carbon is then released from the forest to the atmosphere. These sensitivities generally become stronger with time across the ensemble—which means a negative trend in dNEP/dT (Fig. 3.8h) and a positive trend in dCO_2/dT (Fig. 3.8i). In an example run (Fig. 3.8c–e), the trends in these indicators are present at least a century before dieback begins (Fig. 3.8f,g). These results are seen regardless of whether a model shows dieback before

2100 (stacked bars in Fig. 3.8h,i). Whilst it appears the dCO_2/dT sensitivity shows a slightly more robust increasing tendency in the models that do not exhibit dieback by 2100, a Mann–Whitney U test reveals that the two distributions are not statistically significantly different at 5% significance ($p = 0.066$).

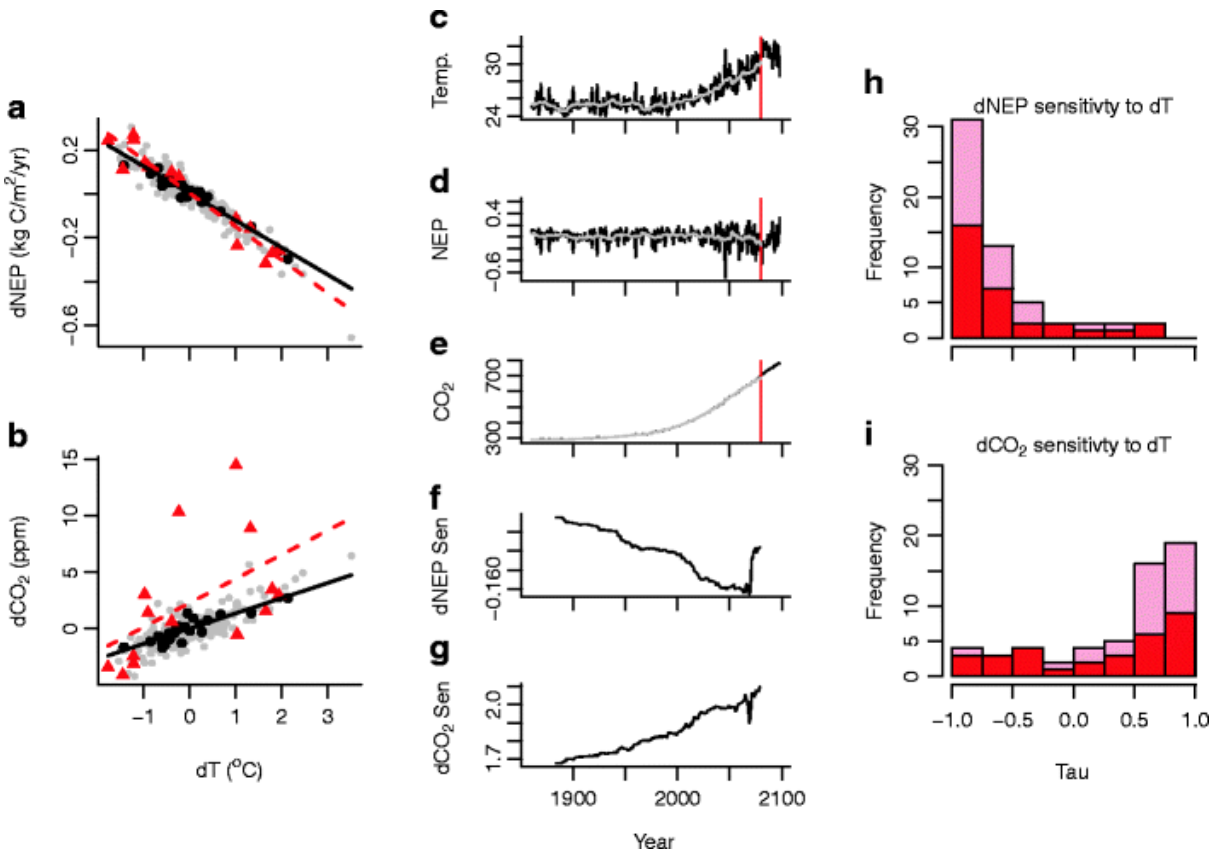

Figure 3.8: Changes in the sensitivity of NEP and CO_2 to temperature anomalies over time as early warning indicators: (a–g) shows results for an individual ensemble member, (h,i) summarises results across the ensemble. In (a) and (b), temperature anomalies (dT) are plotted against a NEP (dNEP) and b CO_2 (dCO_2) anomalies, after detrending the time series with bandwidth = 10. Using a window length of 25 years, the first window's points are shown as dark circles and the last using triangles with the corresponding linear regression lines plotted using a solid line and dotted line respectively. The original time series of temperature (in °C) (c), NEP (in kilograms of carbon per square metre per year) (d) and atmospheric CO_2 (in ppm) (e) are shown along with the detrended time series (lighter colour). Only using data prior to the start of dieback in broadleaf fraction (shown by the vertical line in c–e), the gradient of the regression lines for both (f) dNEP/dT and (g) dCO_2/dT in sliding windows of 25 years are plotted. Observing the whole ensemble, histograms show the Kendall's τ of the resulting

trends in sensitivity of (h) dNEP/dT and (i) dCO_2/dT. In both cases, the darker bars of each histogram refer to ensemble members that show dieback starting by 2100 whereas the (stacked) lighter bars to not.

Broadleaf fraction model

On examining the behaviour of the simplified model of broadleaf fraction, the null model ensemble produces no significant early warning signals as expected (Fig. 3.9a–c). In the linearly forced model ensemble, we do find the expected early warning signal of increasing AR(1), together with a slight tendency toward increasing variance, and a tendency toward negative skewness (Fig. 3.9d–f). However, in the non-linearly forced ensemble, the early warning signal from AR(1) is eliminated (with no clear trends in this indicator), whereas variance is strongly increasing reflecting increasing amplitude of the forcing noise, and there are no clear trends in skewness (Fig. 3.9g–i). These results suggest that the fairly rapid and non-linear forcing of the full climate model may be responsible for eliminating the expected early warning signal of rising autocorrelation.

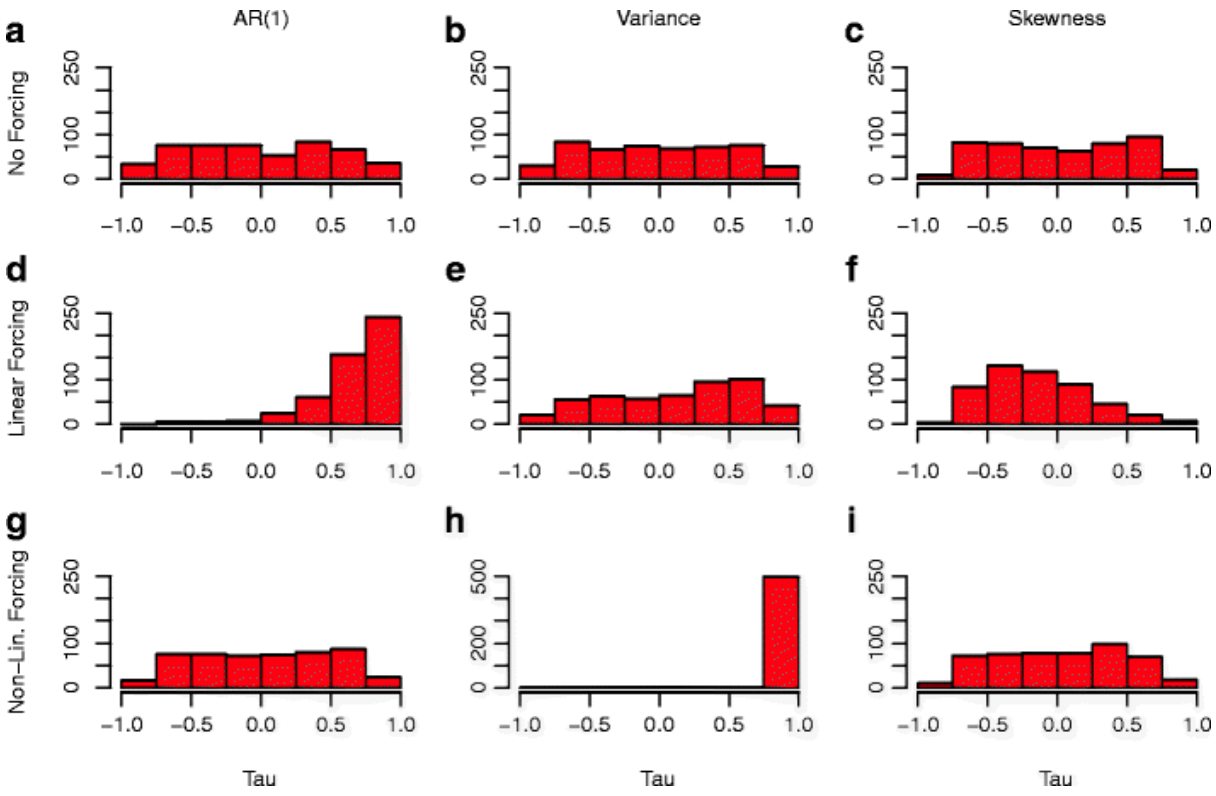

Figure 3.9: Results using a simplified version of the governing TRIFFID equation used in HadCM3 to determine broadleaf fraction (Eqn. 3.1). Kendall's τ values from early warning signals for the three 500 member ensembles with no forcing (a–c), linear forcing (d–f) and non-linear forcing (g–i) are shown as histograms. Full details on each of the three ensembles are given in the main text.

Discussion

We have analysed time series of the Amazon rainforest in an ensemble of versions of the HadCM3 model, to test both generic and system-specific indicators of an approaching tipping point. Despite Amazon dieback beginning in over half of the ensemble members before 2100, the expected generic early warning signals of an approaching bifurcation-type tipping point are not consistently present. In particular, the expected signal of increasing autocorrelation is missing. There is a robust increase in variance in the models but this may be attributed to increasing variance in key forcing factors for the forest, notably increasing temperature variability driven by El Niño.

The failure to observe generic early warning signals could occur for several reasons. The simplest would be if there was no 'critical transition' in the model world. However, we have several reasons to believe that the model is capable of displaying critical slowing down. As already noted, the equilibrium response of Amazon forest cover to temperature is strongly non-linear (Jones *et al.*, 2009), and such responses should show early warnings signals even if there is no bifurcation (Kéfi *et al.*, 2013). Furthermore, the underlying equations suggest that as NPP declines, and with it forest cover, a transcritical bifurcation is approached. Although the system is a long way from the bifurcation when NPP starts to decline and then dieback begins, it is moving towards the bifurcation and therefore the corresponding leading eigenvalue is increasing toward zero, which should produce critical slowing down. Hence, the explanation for the lack of early warning signals could be that the conditions for observing them—namely that the model is forced slowly and subject to low amplitude, additive

stochastic variability (such that it remains close to its equilibrium behaviour)—are violated. Indeed, the full climate model is subject to fairly rapid and non-linear forcing, and when we impose such a forcing in a simplified model of broadleaf fraction, the expected early warning signal of rising autocorrelation is eliminated.

It is also worth noting that we are able to measure generic early warning indicators in our ensemble with relative ease compared to the real world. Observations of time series would be hard to obtain, especially to the degree of accuracy we would need to test, for example, increases in variance in broadleaf fraction. In this case, observational error would be larger than the measures of variance we generally observe in the time series derived from the models runs.

One critical feedback that is missing from the models analysed here and could contribute to the presence of early warning signals in reality is a positive feedback between vegetation state and fire. Essentially, the presence of trees suppresses fire encouraging their dominance over grasses, whereas the presence of grasses promotes fire, preventing the establishment of trees. This local scale feedback could be responsible for creating alternative attractor states (forest or savannah) across large parts of the Amazon today (Hirota *et al.*, 2011, Staver *et al.*, 2011). Analysis of satellite vegetation cover data suggests a range of precipitation for which forest and savannah states are both stable in the Amazon region, with the forest becoming less ‘resilient’ (i.e. the basin of attraction becomes shallower) as precipitation declines toward a critical threshold around 1,800 mm/year (or 5 mm/day) (Hirota *et al.*, 2011, Staver *et al.*, 2011). Corresponding spatial locations where there is predicted to be forest-savannah bi-stability are in the southern or south-eastern Amazon, particularly

in Bolivia (Hirota *et al.*, 2011). There are also areas of currently low tree cover in south Brazil that are predicted to be bi-stable (Staver *et al.*, 2011). In recent drought years, a critical transition to a 'mega fire' regime has been observed in a part of the Amazon (Pueyo *et al.*, 2010).

Failures of the generic early warning indicators caused us to seek more process-based indicators of changing forest stability in the model. The sensitivity of NEP to temperature anomalies (a negative correlation) generally increases over time (as dieback approaches). This is readily understood because respiration is prescribed as an exponential function of temperature whereas photosynthesis is a peaked function of temperature (and may even switch regime from rising to declining with temperature). Thus, at higher temperatures a given increase in temperature gives rise to a larger decrease in NEP. Furthermore, the sensitivity of atmospheric CO_2 anomalies over the Amazon region to temperature anomalies (a positive correlation) increases robustly, and both of these quantities are readily observable in the real world. The GOSAT satellite record of regional atmospheric CO_2 measurements (including the Amazon region) only began around 2009. However, global CO_2 anomalies are dominated by variability in tropical land carbon stores and they have been measured for longer.

Although these process-based stability indicators are potentially observable, they can only indicate a tendency of changing resilience of the forest (in this case decreasing). There is no particular, universal threshold value which signals Amazon dieback in the model, perhaps because the physics of each model version is different. For example, the forest of one model version may be resilient to dNEP/dT sensitivity (gradient) of −3, whereas another may show

dieback before this sensitivity is reached. This lack of an absolute indicator of a dieback threshold is also true for generic indicators such as rising variance.

As yet we have not tested any spatial early warning indicators on the data, for example, rising spatial correlation (Bathiany *et al.*, 2013, Dakos *et al.*, 2011). However, the ensemble members in HadCM3-ESE generally show dieback which begins in the north-east of the region and moves towards the centre. In other words, dieback does not occur coherently and simultaneously across the whole region in the model. Hence spatial indicators may fail. Also, we have not performed a grid point by grid point analysis for all the ensemble members, because this would involve individually examining ~3,500 time series. We believe such analysis would be redundant because it is already clear from our analysis of the broadleaf fraction model that due to fast, non-linear forcing of the system, including increasing variance in the driving time series, generic early warning signals—notably rising autocorrelation—tend to fail.

The absence of generic early warning signals of Amazon dieback in the HadCM3 model does not imply they would be absent in the real world. In particular, observational data suggests that there is bi-stability of tree cover at a much finer spatial scale than the model resolves, and this may be due to positive feedbacks between vegetation cover and fire that are not included in HadCM3 (Hirota *et al.*, 2011, Staver *et al.*, 2011). Therefore further research should consider if this can give rise to early warning signals at a finer spatial scale.

Chapter 4

Exploring the Uncertainty of Modelled Amazon Dieback

This chapter is based on Boulton, C. A., Booth, B. B. B., and Good, P. G., 'Exploring uncertainty of Amazon dieback in a perturbed parameter Earth system ensemble' (in prep, previously reviewed at Global Change Biology)

Abstract

The future of the Amazon rainforest is unknown due to uncertainties in projected climate change and the response of the forest to this change (forest resiliency). Here we explore the effect of the uncertainties in climate and land surface processes on the future of the forest, using a perturbed physics ensemble of HadCM3. This is the first time Amazon forest changes are presented using an ensemble exploring both land vegetation processes and physical climate feedbacks in a fully coupled modelling framework. Under three different emissions scenarios, we measure the change in the forest coverage by the end of the 21st century (the transient response), and adapt a concept of 'dry-season resilience', to predict the long term committed response of the forest, should the state of the climate remain constant past 2100. Our analysis of this ensemble suggests that there will be a high chance of greater forest loss on longer timescales than is realised by 2100, especially for mid-range and low emissions scenarios. In both the transient and predicted committed responses, there is an increasing uncertainty in the outcome for the forest as the strength of the emissions scenarios increase. We then decompose the uncertainty into that associated with future climate change and that associated with forest resiliency, finding that it is important to reduce the uncertainty in both of these if we are to better determine the Amazon's fate.

Introduction

There is currently a lot of focus on the future stability of the Amazon rainforest. This is largely due to its role as an important carbon store and current carbon sink in the climate system (Malhi *et al.*, 2008). Significant loss, or dieback, of the rainforest could result in this carbon sink becoming a source, releasing carbon from the forest to the atmosphere and contributing to climate change. The Amazon rainforest is also important for other reasons such as sustaining high biodiversity (Dirzo & Raven, 2003).

General circulation models (GCMs) give some insight into the future responses of the rainforest, projecting climate change forced by emissions scenarios, and (if they include a vegetation model - DGVM) the response of the forest to this. Amazon forest dieback was first simulated in an offline vegetation model forced by climate change occurring in HadCM2 (White *et al.*, 1999). Since then, it has also been found in some coupled GCMs such as HadCM3LC (Cox *et al.*, 2000). Results from the standard version of HadCM3 show much larger dieback compared to simulations from most other GCMs. This is due to strong regional drying and warming that overwhelm the rising atmospheric CO_2 that contributes to increased photosynthesis and thus productivity of the Amazon rainforest (Cox *et al.*, 2004, Good *et al.*, 2011, Good *et al.*, 2013, Huntingford *et al.*, 2013, Malhi *et al.*, 2009). This does not mean the response in HadCM3 is implausible: Shiogama *et al.* (2011) used observational constraints to suggest that the CMIP3 ensemble mean underestimates the most likely level of drying over the central/eastern Amazon. Nevertheless, the differences between current projections suggest that the forest's future is uncertain.

There has been much research into the varied responses of the forest under different GCM projections and dynamic global vegetation models (DGVMs). For example Sitch *et al.* (2008) use a single GCM to test a variety of DGVMs under different emissions scenarios. More recently, Huntingford *et al.* (2013) test the effect of climate change patterns from 22 GCMs which explore changes in land vegetation processes (Booth *et al.*, 2012), whilst using a single DGVM (TRIFFID) (Cox, 2001). These changes are then compared to those from Sitch *et al.* (2008) suggesting that there is a larger uncertainty associated with future emissions scenarios than climate model uncertainty. These works explore uncertainty in the future of the Amazon rainforest by focusing on specific modelled components (e.g. forest resiliency and climate change respectively). Poulter *et al.* (2010) perturb parameter values within the LPJmL DGVM and combine this with an ensemble of 8 GCMs to determine which parameters are most important in reducing uncertainty of future Amazon rainforest response. Galbraith *et al.* (2010) use factorial simulations to determine the effect that certain factors, such as temperature or precipitation changes, have on vegetation carbon in the Amazon region for three DGVMs.

The modelled vegetation in the rainforest (as well as vegetation elsewhere) exhibits inertia, meaning there is a delay in the response of the forest to the climate change that has occurred. The eventual response based on the climate change that has happened up to a certain time is known as the 'committed response' and can take many years to be realised (Jones *et al.*, 2009). The committed response may be calculated using 'equilibrium vegetation' simulations where the climate is held at a constant level and the vegetation is allowed to settle to equilibrium (Cox, 2001, Jones *et al.*, 2009). In a transient scenario (where radiative forcing was steadily increasing), Jones *et al.* (2009)

found that Amazon dieback lagged the committed forest change by around 50 years. In this case, the transient forest response could be considered a lower bound to the potential long term forest loss that would occur in this model without reversing climate change. Understanding this committed response is important in determining the longer term outcome of the forest to emissions over the 21st century as, for example, the area of sustainable forest coverage may be significantly reduced well before transient loss is observed. Huntingford *et al.* (2013) calculate the committed response for the 22 models they test and find that rainforests that are growing in the transient experiment continue to grow slightly whereas rainforests which have 'peaked' and are on a decline show more dieback in their committed response.

The primary controls on the large-scale distribution of committed vegetation under present-day through future conditions are rainfall, temperature and atmospheric CO_2 concentration. Good *et al.* (2011) showed that for HadCM3LC and HadGEM2-ES (Good *et al.*, 2013), while considering tropical land, certain combinations of temperature and dry-season length (the number of months a year that precipitation falls below a certain threshold) promote sustainable forest, whereas there is no forest found in areas which are too warm or dry (i.e. have a long dry-season length). The dry-season length used here is the number of months that produce a water deficit according to maximum cumulative water deficit (MCWD) calculations (Malhi *et al.*, 2009). In these simulations at least, the boundary between sustainable forest and no forest is fairly distinct.

In turn, Amazon rainfall anomalies have been linked to sea surface temperature indices in both the tropical Pacific (Cox *et al.*, 2004, Harris *et al.*, 2008) and Atlantic (Cox *et al.*, 2008, Good *et al.*, 2008, Harris *et al.*, 2008). Both of these

indices are observable in the real world. Furthermore, increased rainfall comes from air that has passed over extensive vegetation suggesting that precipitation changes are also linked to deforestation (Spracklen *et al.*, 2012). Using observed precipitation values in tropical rainforest areas, potential analysis (Livina *et al.*, 2010) has been used to determine how vulnerable certain areas of the forest are (Hirota *et al.*, 2011) which is related to how far away they are from the boundary of not having enough precipitation to sustain themselves.

Dieback of the Amazon rainforest has been considered a tipping point in the Earth system (Lenton *et al.*, 2008) and generic early warning signals based on time-series analysis of variance, autocorrelation and skewness (Lenton, 2011) have also been tested on output of the ensemble of HadCM3 used here (Boulton *et al.*, 2013). However due to the slower dynamics of the system (the committed response of the forest) compared to fast, anthropogenic forcing, the generic early warning signals do not show much promise. Instead, indicators based on the physical processes of the Amazon rainforest appear to be more promising.

Here we explore uncertainty in Amazon forest projections using output from a 57-member perturbed-physics ensemble of HadCM3 (Booth *et al.*, 2013), a GCM whose Amazon dieback in its standard configuration is at the upper end of current projections. Our uncertainties in future climate change and forest resiliency are represented by the physical parameters that are perturbed in the ensemble, allowing the opportunity to determine how sensitive future Amazon forest change is to these. This ensemble explores both land vegetation processes and physical climate feedbacks and represents the first time future Amazon rainforest changes have been analysed with this uncertainty. This is

carried out within a fully-coupled framework meaning there is no mismatch between atmospheric drivers and changes in surface conditions, and allowing the vegetation to feedback on the atmosphere, both locally and globally. As well as analysing the response of the Amazon rainforest by 2100 under 3 emissions scenarios for each ensemble member, we also predict the long term committed change of the forest, which would not be realised for many decades beyond 2100. To do this, we present a novel use of the dry-season resilience method described earlier (Good *et al.*, 2011). Our use of this method determines a lower bound of potential long term forest loss.

Materials and methods

HadCM3-ESE

Our data is obtained from the HadCM3 Earth System Ensemble (HadCM3-ESE) (Lambert *et al.*, 2013), using the TRIFFID DGVM (Cox, 2001) to simulate vegetation. There are 57 model configurations within the ensemble, each containing a different combination of perturbed parameters. The parameters are perturbed within boundaries suggested by experts and grouped according to their role within the Earth system, whether they are part of the carbon cycle (n=8 parameters) (Booth *et al.*, 2012), atmosphere (n=32) (Collins *et al.*, 2011), sulphur cycle (n=8) (Lambert *et al.*, 2013) or ocean (n=15) (Collins *et al.*, 2007). A Latin hypercube sampling method was used to sample a range of combinations of carbon cycle and atmosphere parameters (Lambert *et al.*, 2013). There were originally 68 members, however 11 were removed from the ensemble for failing to simulate reasonable top of the atmosphere (TOA) radiative fluxes during the spin up (outside the bounds in Collins *et al.* (2011)). Ensemble members that failed to simulate the presence of Amazon or boreal forests were also removed (Lambert *et al.*, 2013). The ensemble is driven by emissions profiles expected to give the trajectories explained below (much like Meinshausen *et al.* (2008)). This means that atmospheric greenhouse gas concentrations are prognostic values and vary due to different emergent model sensitivities resulting from the underlying perturbed parameters sampled in these experiments, even under the same emissions scenario. If the direct forcings or concentrations were applied to the ensemble members, it would

prevent the opportunity to explore global feedbacks in the carbon cycle and thus by using emissions profiles, greater uncertainty is explored.

HadCM3-ESE has been run under 3 scenarios, a mitigation scenario RCP 2.6 (van Vuuren *et al.*, 2006, van Vuuren *et al.*, 2007) , a balanced scenario, A1B (Nakicenovic *et al.*, 2000) and a business as usual scenario, RCP 8.5 (Riahi *et al.*, 2007) , as detailed by Booth *et al.* (2013). General comparisons between each scenario's model outputs such as global mean temperature have been shown elsewhere (Booth *et al.*, 2013). Each of the scenarios share a common historical driving dataset from 1860-1950 based on SRES data, after which parallel SRES and RCP historical simulations were run. These form the basis from which SRES A1B (from 1990) and the 2 RCPs (from 2005) were extended. Further details about the experimental setup are described by Booth *et al.* (2013).

HadCM3-ESE was originally created to explore the spread of results possible under HadCM3 dynamics, rather than to determine the effects of individual parameters on changes in vegetation. For this, single parameters would have to be perturbed whilst keeping other constant. However we explored the relationship between the transient responses and land surface parameters perturbed in the ensemble, noting that full inference on the true effect of each parameter is difficult to determine. The parameters concerned with the carbon cycle (Booth *et al.*, 2012) are shown in Table 4.1 with a short description and the ranges they are sampled from. Parameters from the other groups (detailed above) are less influential on forest response and are not included in Table 4.1. Note that some perturbed parameter values are assigned to each plant functional type (PFT) in the ensemble, however Table 4.1 only shows the

ranges for the broadleaf fraction PFT. Full details of the other PFTs can be found in Booth *et al.* (2012).

Parameter	Range (for broadleaf FPT)	Description
f0	0.72-0.95	Stomatal resistance
minLAI	1-4	Minimum Leaf Area Index needed before a PFT competes for space
N_{LO}	0.018-0.1 kgN/kgC	Top leaf nitrogen concentration
Q_{10}	1.5-3.5	Soil respiration
T_{OPT}	27-37 (°C)	Optimum temperature for photosynthesis
Θ_{CRIT}	0.01-0.99	Volumetric soil moisture

Table 4.1: Ranges and descriptions of perturbed parameters in the carbon cycle component of HadCM3-ESE, as detailed in Booth *et al.* (2012)

Estimating the committed forest response: modified dry-season resilience method

The basis of our analysis is to determine climate conditions that sustain forest and to explore the long term committed response of the forest (Jones *et al.*, 2009), due to the changes in these climate conditions over time. Our method is based on that of Good *et al.* (2011).

The method of Good *et al.* as tested on the standard version of HadCM3LC (Good *et al.*, 2011) and HadGEM2-ES (Good *et al.*, 2013), uses annual mean temperature and annual dry-season length (DSL, the number of months in a year that monthly precipitation is below 100mm) from land grid points in the tropics (20°S-20°N), as well as global atmospheric CO_2 concentration as climate drivers that affect the sustainability of the forest.

To determine climate conditions that are suitable for sustaining forest, equilibrium broadleaf tree fraction (BL) is plotted in the temperature-DSL plane for a given model configuration (see Fig. 4.1a for an example using our method). The points are coloured depending on whether there is forest (green, BL > 0.4), no forest (red, BL < 0.05). An intermediate amount of forest (blue, 0.05 < BL < 0.4) is savannah type land that are dominated by grasses. We have also circled points contained within a region we define as the Amazon rainforest (40°-70°W, 15°S-5°N) as the climate changes in these points are what we are most interested in. Fig. 4.1a shows two distinct regions: one where climate promotes sustainable forest growth and a region which does not contain forest. The boundary between the two regions is approximately linear, so is quantified with a linear fit of the form shown in Eqn. 4.1.

$$DSR = DSL + \alpha T + \gamma CO_2 + c \quad (4.1)$$

DSR (in units of months) is dry season resilience, a measure of the resiliency of a grid point to changes in climate. Visually, *DSR* refers to the distance away from the boundary between forest and no forest a grid point is with *DSR*=0 on the boundary itself, suggesting points on the boundary have no resilience to an increase in temperature or dry season length. *DSL* and *T* refer to the dry-season length and temperature of a given grid point whereas CO_2 is the global mean value of atmospheric CO_2. The coefficients α and γ, the temperature sensitivity and CO_2 fertilisation coefficient respectively are to be determined along with the constant *c*. With this formulation, we are able to make statements such as 'if DSL were to increase by a month, then temperature would have to decrease by α for the grid point to have the same resilience'. The parameters α, γ and *c* in Eqn. 4.1 are dependent on the parameters perturbed within the ensemble and as such there is uncertainty associated with them, which we will later decompose.

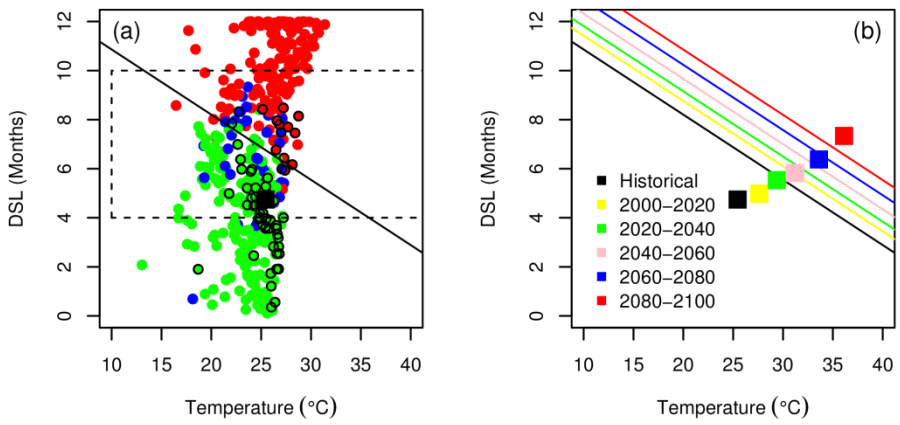

Figure 4.1: Estimating dry-season resilience (DSR) for a typical ensemble member. (a) The historical (1860-1950) mean temperature and dry-season length (DSL) is observed for all tropical grid boxes (20°S -20°N) which are then plotted in the temperature-DSL plane. The colour of each grid boxes' point is green for 'Forest' (*BL* > 0.4), blue for 'Intermediate' (0.4 < *BL* < 0.05) or red for 'No Forest' (*BL* < 0.05). The *DSR*=0 line (as described in main text) is shown by a black line. Circled points are those contained with the region 40°-70°W, 15°S-5°N (the Amazon region) and the black square is the mean state of the Amazon forest (green circled points). Dotted lines refer to the region used to calculate parameters in Eqn. 4.1. (b) Future changes in atmospheric CO_2, temperature and DSL move both the DSR=0 line and the position of the points (represented here by the mean Amazon forest state, black square) are tracked in 20 year averages over the 21st century.

Good *et al.* originally estimated these parameters on equilibrium runs, where the vegetation has settled to equilibrium under a constant climate. The

parameters are calculated through the use of an algorithm that minimises the number of grid points that are on the wrong side of the boundary.

To fit the parameters for this ensemble, we adapt the above method. Initial equilibrium vegetation simulations were not available, due to computational expense associated with the large ensemble size (this would involve carrying out 57 additional full GCM experiments for each of the 3 future scenarios explored in this study). Using the fact that the 3 scenarios used the same historical simulation from 1860-1950, as well the climate staying relatively stable during this time, we treat this as a 'quasi-equilibrium' early industrial state to begin our analysis from. For each land grid point in the tropics (20°N-20°S) within each configuration, we calculate the average temperature, dry season length and the average broadleaf (BL) fraction over these 90 years. We also extract the 1860-1950 mean global CO_2 (ppm) value for each ensemble member.

Our modification of the original DSR method is to use a logistic regression fit to estimate the parameters in Eqn. 4.1, focusing around the transition from forest to no forest by using only grid points with temperature, $T > 10°C$ and $4 < DSL < 10$ (Fig. 4.1a – dotted line) and fitting the line to where $BL=0.025$, the midpoint of the blue, intermediate values of forest in Fig. 4.1a. This standardised method of computing α and c is much more efficient than using the original method to determine them for all 57 ensemble members.

An important caveat here is that by using equilibrium runs in their analysis, Good *et al.* (2011) were able to infer the value of γ, the CO_2 fertilisation coefficient (equal to 0.0043) from HadCM3LC, the standard model with the original parameter values by running a parallel model with double the

atmospheric CO_2 concentration. We use their value in our analysis as we do not have the simulations required to estimate this fertilisation coefficient for each individual configuration. These extra runs would have allowed us to have two values for atmospheric CO_2 from which we would be able to infer the fertilisation coefficient through the use of our logistic regression fit each time. Instead we are making the simplification that the CO_2 fertilisation effect does not vary between simulations, although it is important to note that the true value of γ in each instance is dependent on the parameters perturbed for each configuration. We note that only the fertilisation coefficient γ is kept constant across all configurations and that the fertilisation effect itself will differ depending on the global atmospheric CO_2 concentration.

After determining α and *c* for each configuration, Eqn. 4.1 allows a prediction of whether broadleaf forest is sustainable or not at each location for each year based on its DSR value (Eqn. 4.1), given the prevailing climate. First we calculate the number of points in the Amazon region that are below the *DSR*=0 line in our quasi-equilibrium (1860-1950) state in each of our configurations. We then calculate the number of points that are below the line using the 2080-2100 average from each simulation. In other words, we are using DSR as a method of extrapolation to estimate the state of the committed Amazon rainforest without running a corresponding equilibrium run for each ensemble member. A prediction on the post-2100 equilibrium state of a grid point is based on the equilibrium state of a grid point with similar climate in the quasi-equilibrium state. Due to the CO_2 fertilisation effect, increases in atmospheric CO_2 cause the boundary line (*DSR*=0) to move upwards. Consequently moderately increased temperatures and DSLs can sustain forest under the higher atmospheric CO_2 values. An example of these changes over the 21st century is

shown in Fig. 4.1b. We use the 20 year averages to eliminate year-to-year variability. The difference between the 1860-1950 and 2080-2100 values gives us our prediction of committed change. The combined result of configurations from each emissions scenario gives us a measure of uncertainty of the future behaviour of the Amazon rainforest and its committed response to 21st century climate change.

Results

Transient responses

Time series of the transient responses of the Amazon rainforest up to 2100 in HadCM3-ESE are shown in Fig. 4.2a as the number of grid boxes with exhibit forest (BL > 0.4) in each ensemble member. Proportional changes shown in Fig. 4.2b. These responses are calculated as the proportional change in the number of Amazon region (40°-70°W, 15°S-5°N) grid points that exhibit forest (i.e. *BL* > 0.4) between 2000 and 2100.

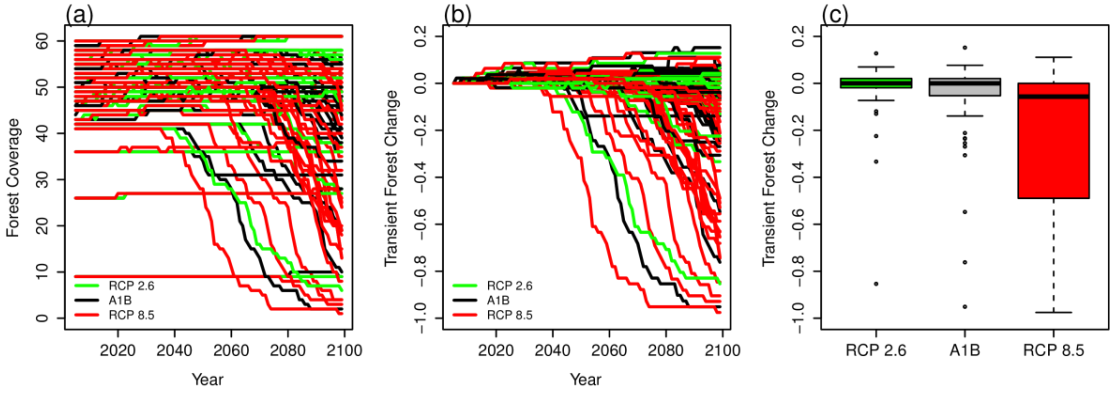

Figure 4.2: Transient changes in number of grid boxes containing Amazon forest (BL fraction > 0.4 within the region 40°-70°W, 15°S-5°N) in HadCM3-ESE compared to historical (1860-1950) Amazon forest coverage. (a) Time series of forest coverage (number of grid boxes) for each individual member of HadCM3-ESE. (b) Time series of these transient proportional changes for each individual member of HadCM3-ESE. (c) Box and whisker plots for each scenario showing the median, inter-quartile range and minimum and maximum values (ignoring outliers, black circles).

Unlike the standard configuration of HadCM3 (Cox *et al.*, 2000), the majority of the simulations show little change by the end of the 21st century (Fig. 4.2a,b). However there are simulations which show dieback at similar levels to that of the forest in the standard model and even greater. This shows that the large dieback in the standard model is not the mean result for the HadCM3 model. However, it is still within the envelope of uncertainty provided by this ensemble.

When partitioning the transient responses by scenario (Fig. 4.2c), there is an increasing uncertainty in the forest state at 2100 with increasing strength of emission scenarios. With the RCP 2.6 mitigation scenario, we see that the mean transient response is no change to the forest cover with a few simulations showing dieback, giving a negatively skewed distribution. For A1B simulations, while the mean response still suggests no change, it is clear there is more of a tendency for forest loss to be exhibited than occurs under mitigation. With RCP 8.5, the mean response decreases slightly to a loss of around 5%. The uncertainty however is a lot greater. As well as having more members which show loss and dieback, there are also more simulations that have forest growth than the other two scenarios.

The simulation with the largest dieback that occurs in RCP 2.6 shows signs of forest loss by 2040 and does so in all three scenarios. For the forest to dieback so soon in the century suggests that in some cases, the configuration of perturbed parameters can cause forests that are already very near the threshold of dieback under present day conditions.

When determining if any of the perturbed parameters were linked to forest loss, we found the strongest relationships were found between forest change and T_{OPT} (the optimum temperature for photosynthesis) and *minLAI* (a competition

parameter specifying the minimum leaf area index a plant functional type needs before it begins to compete for space). Using results from RCP 8.5, which have the largest spread of transient responses, analysis on the combination of T_{OPT} and *minLAI* on forest change (Fig. 4.3) shows low values of T_{OPT} and high *minLAI* for broadleaf preconditions dieback. Members with a T_{OPT} of greater than 32°C show no extreme dieback (although less extreme loss is still observed, Fig. 4.3a). Likewise, members with a *minLAI* less than 2.5 show no extreme dieback (Fig. 4.3b), whereas members with stronger dieback have a T_{OPT} less than 32°C and a *minLAI* greater than 2.5 (Fig. 4.3a,b). However other factors such as changes in climate that would stress the forest, as well as the values of other parameters not explored, will determine if dieback does occur. Although there are less members which show dieback under the A1B scenario, the boundaries for T_{OPT} and *minLAI* seem consistent (Fig. 4.3c,d). This further strengthens the argument that although other factors such as climate change, which is not as strong in the A1B scenarios, drive dieback, low T_{OPT} combined with high *minLAI* is a precondition. The values of *minLAI* and T_{OPT} in the standard configuration, 3 and 32°C respectively (Cox, 2001), are near the thresholds that precondition dieback (Fig. 4.3). This could explain, at least partially, why dieback is observed in the standard model, but not in the majority of the ensemble.

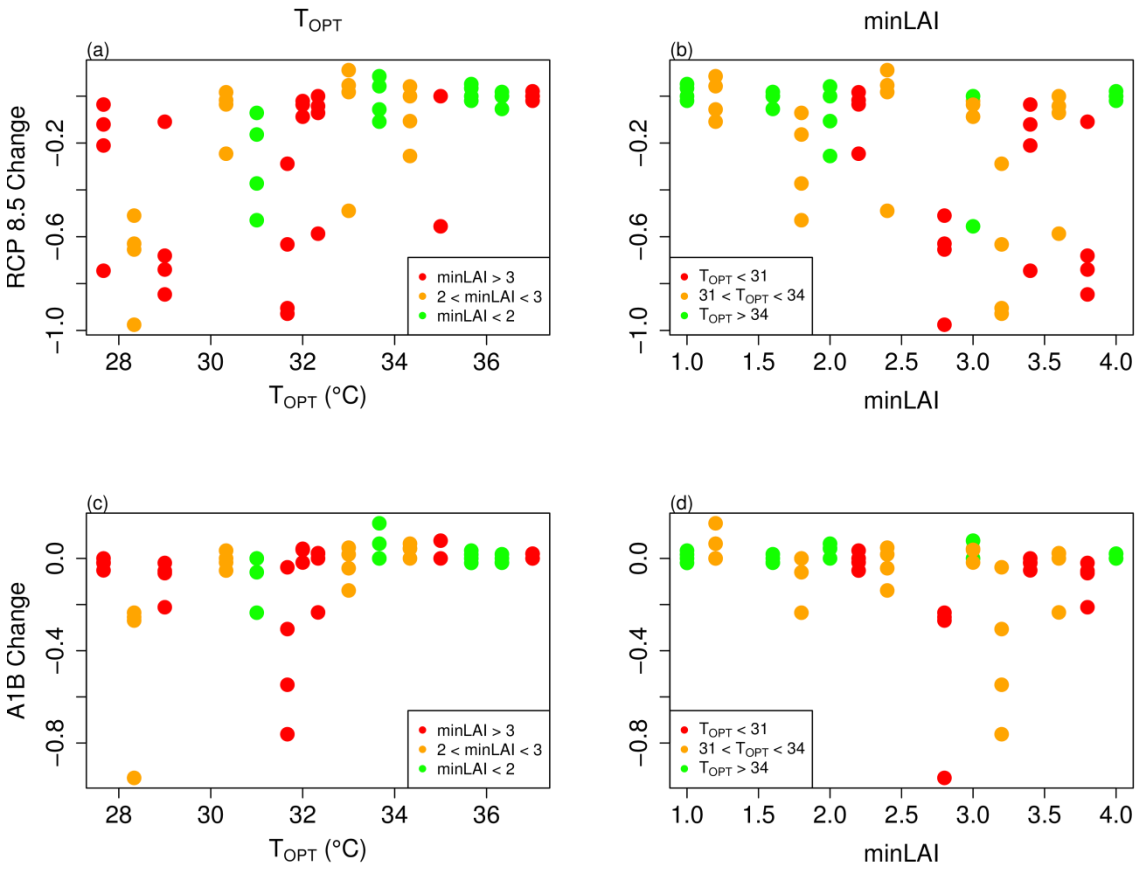

Figure 4.3: The effect of perturbing parameters on transient forest change by 2100. Proportional forest change observed in ensemble members under scenarios (a,b) RCP 8.5 and (c,d) A1B scenarios are plotted against the (a,c) T_{OPT} and (b,d) *minLAI* values of each member. The colours of points show the value of the parameter not plotted.

Committed response predictions

To compare the transient responses (those in Fig. 4.2) to our predictions of the committed responses (calculated using our modified DSR method), we present the results in the form of cumulative density functions (CDFs, Fig. 4.4).

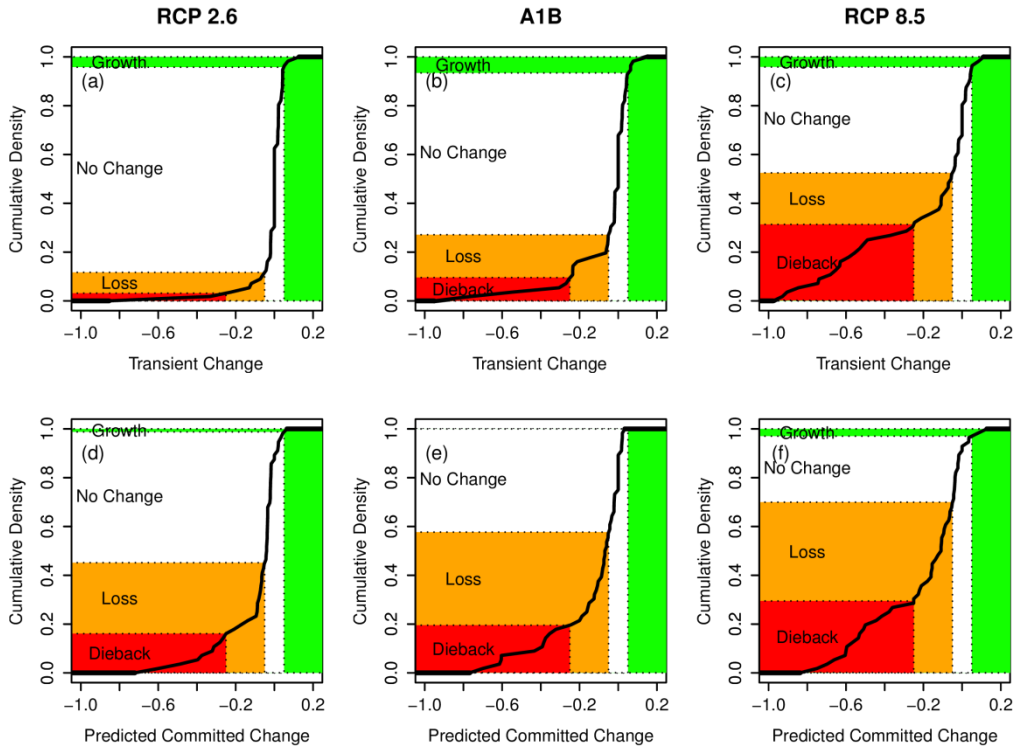

Figure 4.4: Summary CDFs of the Amazon rainforest fractional changes in grid boxes deemed forest for ensemble members of HadCM3-ESE. Transient responses observed by 2100 for scenarios (a) RCP 2.6, (b) A1B and (c) RCP 8.5 are shown above predicted committed responses using the DSR method for (d) RCP 2.6, (e) A1B and (f) RCP 8.5. Coloured regions show proportion of models which show changes we class as ‘Dieback’ (red, < -25%), ‘Loss’ (orange, >-25%,<-5%), ‘No Change’ (white, >-5%, <5%) and ‘Growth’ (green, >5%).

In all three scenarios, our prediction of committed change suggests there is more uncertainty in the eventual outcome of the forest with a higher chance of

further forest loss than is realised by 2100 (the transient response). For example under RCP 2.6, the mitigation scenario, there is fairly robust response of 'No change' (forest remains within 5% of its original size) by 2100 and 12.2% of ensemble members show some degree of 'Loss' (>5% decrease) (Fig. 4.4a). However 45.6% of models predict a committed 'Loss' (>5% decrease) or 'Dieback' (>25% decrease) (Fig. 4.4d). Similar results are observed for the other two scenarios (A1B, 28.1% and 57.9%; Fig 4.4b,e, and RCP 8.5, 52.6% and 70.1%; Fig. 4.4c,f). However fewer models predicting large committed forest loss are found under the mitigation scenario.

In both the transient and predicted committed responses of the forest, stronger emissions scenarios (increased CO_2 emissions for example), lead to an increasing uncertainty in the resulting forest change with more of a tendency towards forest loss. However like the transient response, there is also more forest 'Growth' (>5% increase) predicted as a committed response under the stronger RCP 8.5 scenario than the predicted committed response of the other scenarios. This suggests more spread and thus more uncertainty in future outcome of the forest under stronger emissions scenarios. This uncertainty is also noted by the gradient of the CDFs as steeper gradients suggest less uncertainty.

Decomposing uncertainty

To begin to determine causes in the spread of committed responses predicted, we decompose the uncertainty into that associated with climate change, and that associated with forest resiliency (the coefficient α, the temperature sensitivity, and *c* in Eqn. 4.1, previously calculated individually for each of the 57 configurations). This analysis is carried out on the RCP 8.5 scenario members as out of the three scenarios, this had the largest predicted committed spread (Fig. 4.4).

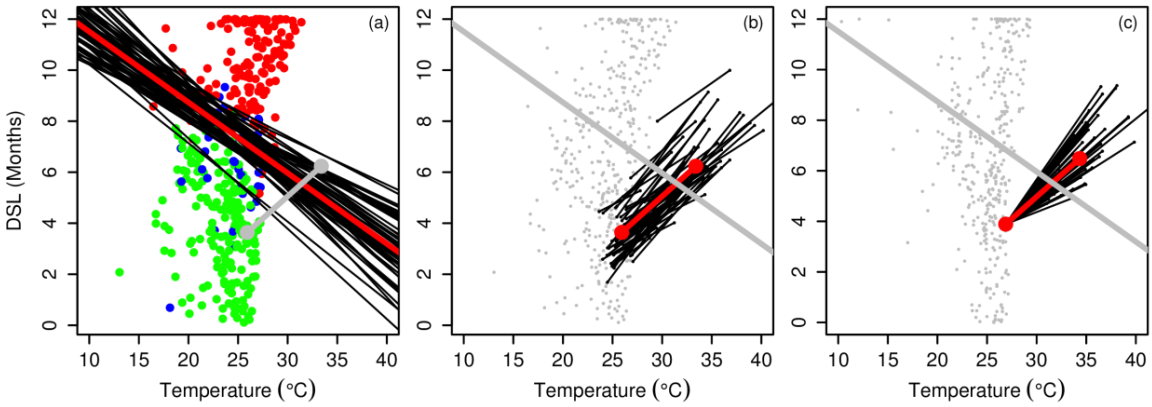

Figure 4.5: Graphical representations of how (a) forest resiliency, (b) climate and (c) climate (including observations) are constrained. In all cases black lines represent values from individual ensemble members, red lines represent the values used when the variable is constrained and grey lines represent how the other variable is constrained. Points shown in the background are from (a,b) a typical ensemble member or (c) observations. The same point (in the Amazon region) is used when demonstrating constraining climate and climate (including observations).

Decomposing the uncertainty is achieved by keeping one set of parameters (either climate change or forest resiliency parameters), constant whilst allowing the other to vary and repeating the analysis used to predict the committed response. The uncertainty associated with the climate change component is explored by fixing the values of α and c in Eqn. 4.1 for each model to the ensemble mean values. This holds the forest resiliency constant. Fig. 4.5a shows the average forest resiliency (red *DSR*=0 line). As the climate is still allowed to change, movement of the grid points over the 21st century as well as movement of the *DSR*=0 due to increasing atmospheric CO_2 will occur. Similarly, uncertainty associated with forest resiliency is explored by fixing the climate at each location in each model to the ensemble mean (Fig. 4.5b shows the average climate change for an example grid point – see red line).

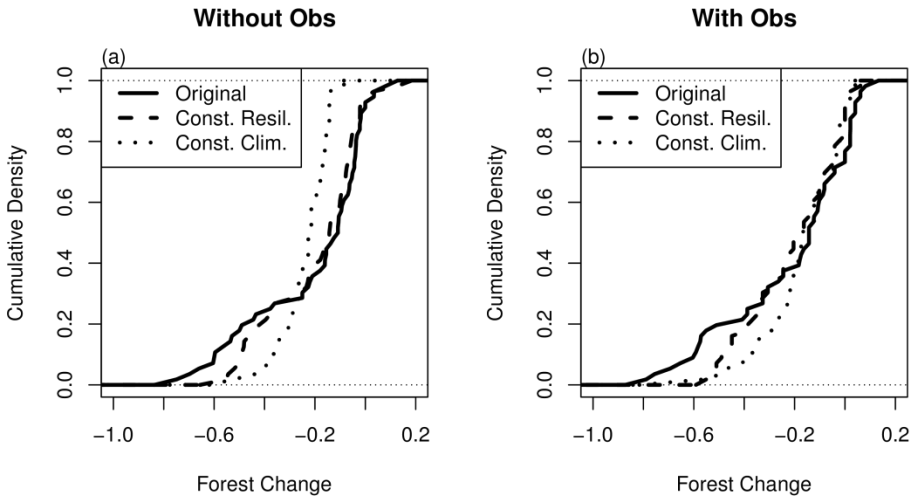

Figure 4.6: CDFs showing predicted committed Amazon forest change for the RCP 8.5 scenario. Committed change is predicted using (a) the 1860-1950 modelled state (temperature and DSL) and (b) real world observations (see main text). Committed change is also predicted whilst keeping resilience parameters constant (dashed lines) and climate change constant (dotted lines). Constraining one variable allows the uncertainty in the other to be explored (as described in text).

Compared to our overall prediction uncertainty (Fig. 4.6a – solid line), we find that our uncertainty due to climate change, under the A1B emissions scenario is similar (Fig. 4.6a – dashed line). This suggests the largest proportion of overall uncertainty is explained by uncertainties in the climate, compared to uncertainty in forest resiliency (Fig. 4.6a – dotted line), which has less of a spread of results (but still shows some uncertainty).

We further constrain our uncertainty in climate change by using real world observations of temperature from the CRUTEM3 dataset (Brohan *et al.*, 2006) and DSL from the GPCC precipitation dataset (Schneider *et al.*, 2014) to use as

starting positions for each grid point. Due to observational constraints, rather than using 1860-1950 as our quasi-equilibrium state, we instead use a 1950-1980 average. Using real world observations eliminates the uncertainty associated with the starting position of each grid point in the temperature-DSL plane. Then the equivalent of the 2080-2100 mean state of a grid point is achieved by adding the climate change in the model between 1860-1950 and 2080-2100 onto the real world observations (Fig. 4.5c). Using these real world observations, we decompose the uncertainty again as we have described above. By doing this, we are able to compare how our uncertainty in forest resiliency compares to our uncertainty in future climate change, eliminating uncertainty in what we already know about recent past climate.

When initialising our analysis using the real world observations as the starting climate (Fig. 4.6b – solid line) we again find that our uncertainties associated with future climate change are still large (Fig. 4.6b – dashed line). Indeed, fixing the starting climate has a rather small effect on the range of projections in this ensemble. However there is more spread in forest outcome while exploring the uncertainty in forest resilience (Fig. 4.6b – dotted line) than previously (Fig. 4.6a). This suggests that forest resiliency is important to understand as well as future climate change.

Discussion

We explore the future Amazon forest response to uncertainties in both land vegetation processes and physical climate feedbacks. These suggest a range of transient forest responses consistent with uncertainties in current climate model parameters. This shows that the result of ~60% dieback from the standard HadCM3 model (Cox *et al.*, 2000) is not the most typical result for this model structure. The values of *minLAI* (3) and T_{OPT} (32°C) in the standard configuration (Cox, 2001) are near the thresholds that precondition dieback (Fig 4.3). This combination of parameters could partially explain the dieback observed in the standard model (Cox *et al.*, 2000).

In the majority of cases, perturbing parameters describing the physics and vegetation processes of the model, leads to forests that are much more resilient to future climate change over the next century than in the standard version. However large changes can still occur, especially under strong emissions scenarios. The spread in results we find compared to the standard configuration highlights the importance of fully exploring both parameter and future emission scenario uncertainty, as well as trying to reduce it.

Our modified use of Good *et al.*'s DSR framework allows us to make predictions of committed change of the forest based on the emissions scenarios up to 2100. Aside from making this prediction, we also hope the methods described in this paper could progress future work towards a more 'system-specific' indicator or framework rather than the more generic early warning signals which have been shown to fail in this instance.

We note here that these predictions of committed change are ‘lower bounds’, meaning that more loss is likely to occur than we predict. Our assumption of the regression model we fit being linear breaks down at higher temperatures since this becomes a limiting factor in forest sustainability when we run a subset of the ensemble members to equilibrium (Fig. 4.7). We note in this case that the emissions are held at constant 2100 levels but the temperature and DSL change from their 2100 values as they lag the emissions. The threshold for when this change in α begins to become significant is dependent on the optimum temperature (T_{OPT}) for photosynthesis in the model configuration (one of the major uncertainties in future tropical forest response (Booth *et al.*, 2012, Matthews *et al.*, 2007). Nevertheless, the technique presented here represents a computationally efficient method of estimating the lower bound to simulated forest loss on the basis of the historical and future GCM climate and forest coverage. Future work could involve adding a non-linear temperature term into Eqn. 4.1 whilst exploring higher temperatures in true 2100 equilibrium runs for example.

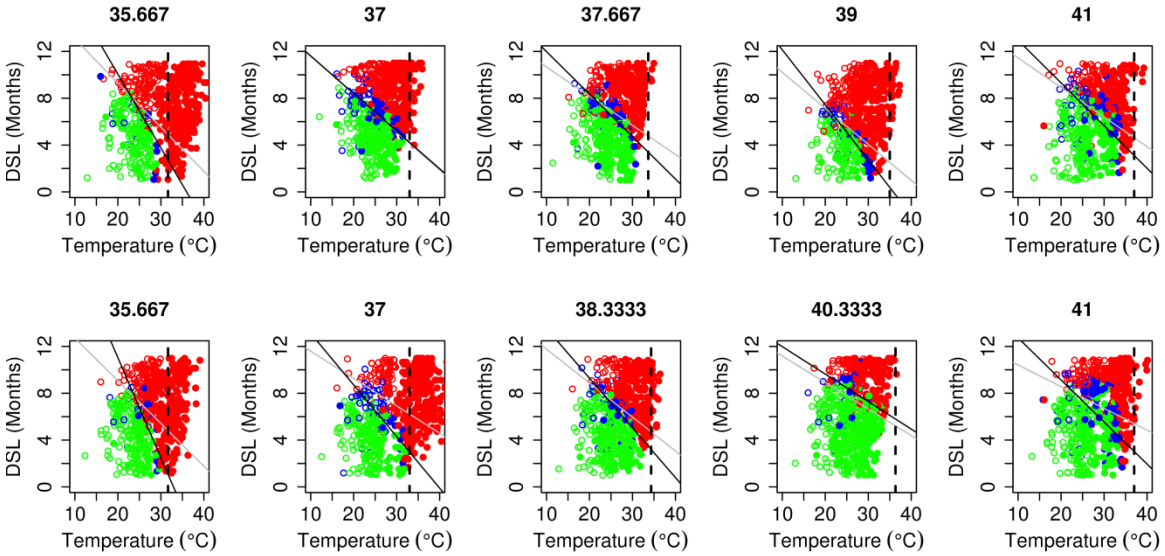

Figure 4.7: The effect of the parameter T_{OPT}, the optimum temperature for photosynthesis on the DSR=0 line for the 10 equilibrium runs (see main text). Plots are ordered in increasing value of T_{OPT}. Hollow points are data from the transient historical (1860-1950) and filled points are from the equilibrium run. The DSR=0 line is shown when fitted only with the transient data (grey) and with all the data (black).

The DSR framework provides a simple metric that can quantify why different models show markedly different responses. For example, there is a huge range of DSLs found within the ensemble, when using both the member mean forest (Fig. 4.8a) and all the forest grid points across all members (Fig. 4.8b). Furthermore, these DSL values are highly correlated with their corresponding MCWD values (Malhi *et al.*, 2009) (r=0.898 for the 1860-1950 state and r=0.963 for the 2080-2100 state when using the ensemble forest mean values, Fig. 4.8a), suggesting that using this more simple measure of using the number of months the forest is under water stress, rather than the amount it is stressed by, in our calculations is beneficial. Given the uncertainty in current DGVM

estimates, the DSR framework gives insight into moisture and temperatures constraints are, and thus could do for other models. Furthermore, the DSR framework could allow the relative contributions of temperature and DSL changes to forest loss, to be calculated.

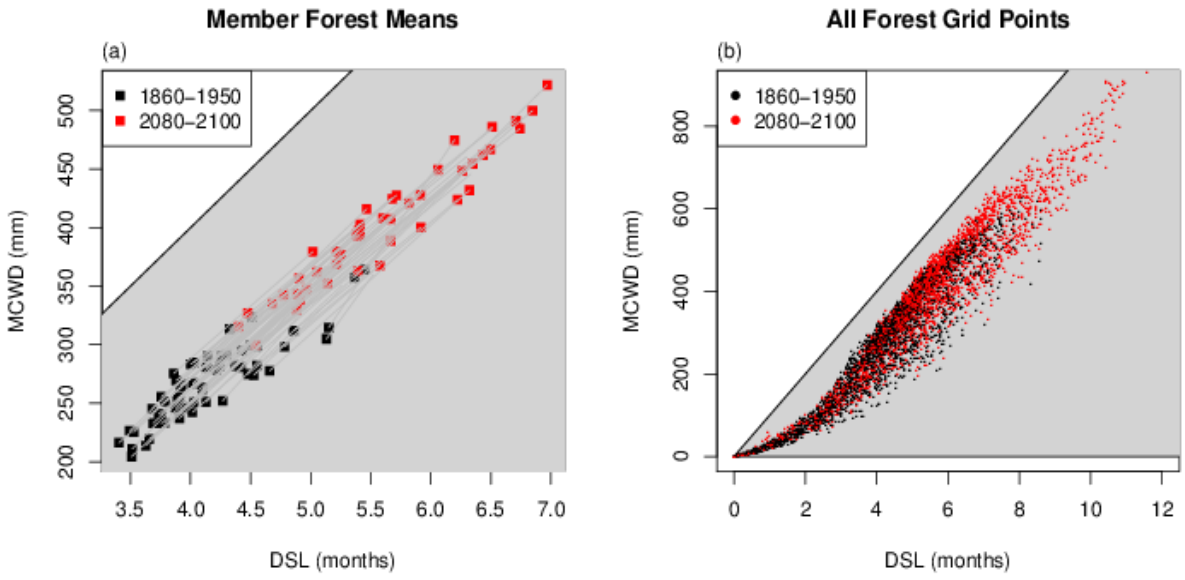

Figure 4.8: A comparison of DSL and MCWD in the HadCM3-ESE under the A1B scenario. (a) Ensemble member forest means and (b) all forest grid point across all ensemble members' DSL values are plotted against their MCWD values for the 1860-1950 mean state (black) and the 2080-2100 mean state (red). In (a), ensemble members' states are connected by a dark grey line. The light grey background shows the envelope of possible values for DSL and MCWD.

As mentioned in our results, as emissions scenarios get stronger, there are a larger proportion of ensemble members which show loss and dieback (both

during the 21st century and in the committed response), suggesting that the increased temperatures and DSLs caused by the increased atmospheric CO_2 overwhelm the CO_2 fertilisation effect. Visually, the movement of the individual grid points in the Amazon region moves towards the boundary between conditions promoting sustainable forest and that unsuitable for forest, faster than the boundary line itself moves. The increases in CO_2 compared to the consequent increases in temperature and dry-season length could be considered as a balance of expansion and risk of collapse and is important to consider when planning mitigation strategy.

As well as an increasing risk of dieback as emissions scenarios get stronger, there is also an increasing uncertainty. Interestingly this could mean that for ensemble members with parameter combinations which lead to slower increases in temperature and DSLs than others, under larger CO_2 increases, forest growth may be observed (green section of the CDF, Fig. 4.4f). This further suggests that reducing uncertainty in forest resiliency (which contributes to these changes in temperatures and DSLs) is important in determining the future response of the forest.

When decomposing the uncertainty in our framework, the climate change component appears to be more important than forest resiliency. However both contribute to the total uncertainty. This is more evident when we use observations as starting climate (comparing Fig. 4.6a and b). It is worth noting that we are assuming that the differences in 1950-1980 and 1860-1950 climates are small relative to future changes, and that the forest stability has not markedly changed as a result. A caveat here, is that we use the CO_2 fertilisation coefficient γ quantified from the standard HadCM3LC model by Good *et al.*

(2011). We are, therefore, only exploring the non-CO_2 fertilisation component of forest resiliency, and subsequently expect this framework to underestimate the importance of the total forest resilience uncertainty, where the impact on CO_2 fertilisation (via changes to parameter γ) would also be accounted for.

To determine how well our ensemble simulates the observations used for the starting climate in our decomposition of uncertainty, we compare them to real-world climate of the Amazon rainforest from the CRUTEM3 (Brohan *et al.*, 2006) and GPCC (Schneider *et al.*, 2014) datasets used previously. By comparing the average Amazon climate state in the temperature-DSL plane and to that of the real-world (Fig. 4.9a), we find that the observations lie within a reasonable range of our simulations as we find models which have Amazon regions state ranges that encompass the observations (Fig. 4.9b,c). Previous work comparing the Amazon region observations to those of members of a multi-model ensemble suggests that models are generally too dry and that accounting for this produces less dieback (Malhi *et al.*, 2009). Our findings suggest that this would have little effect on the predicted committed change of the forest in this ensemble (Fig. 4.6).

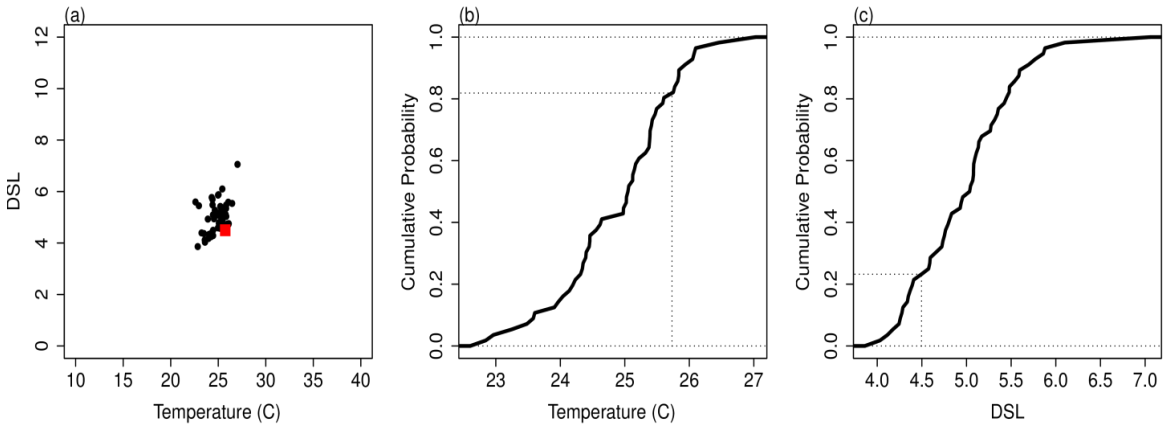

Figure 4.9: A comparison of the starting temperatures and dry-season lengths (DSLs) for members of HadCM3-ESE and the observed climate using the mean Amazon rainforest states. (a) The position of the starting state for the average Amazon rainforest for each ensemble member (black circles) shown alongside the observed average Amazon rainforest state (red square). Cumulative probability distributions are shown for both (b) temperature and (c) DSL for the ensembles with the observed climate shown by dotted lines in each case.

The analysis and consequent results of this work are dependent on the validity of the inertia found by Jones *et al.* (2009) in the standard HadCM3 configuration. Whilst the response time of the forest to natural drought in the real world (such as in 2005) appears to be within months (Phillips *et al.*, 2009), longer term responses to less extreme but more sustained decreases in precipitation or increases in temperature are yet to be determined. Drought experiments (Costa *et al.*, 2010) have shown a slower decrease in tree mortality. This highlights the importance of improving DGVMs such that they are able to create the short term responses to extreme drought, as well as the longer term responses to slow increases in temperature and water stress,

allowing us to reduce our uncertainty in both the forest's transient and committed responses.

Our analysis compliments the work of both Sitch *et al.* (2008) and Huntingford *et al.* (2013). Sitch *et al.* explore uncertainties associated with a number of DGVMs, when used with different emissions scenarios, whereas Huntingford *et al.* explore the uncertainty associated with components of the climate response from a multi-model ensemble which are then used to drive a common DGVM (based on MOSES/TRIFFID). Huntingford *et al.* compare their results with the spread arising from parameter uncertainty in only the land surface response within a fully coupled GCM (HadCM3C, described by Booth *et al.* (2012)). Here we present results from new simulations (Booth *et al.*, 2013) that explore uncertainties in both the land surface/vegetation response and the physical climate simultaneously. This provides the first GCM ensemble where uncertainties in both physical climate and land processes interact within a common experimental framework. Furthermore our approach to determining uncertainty is very different from both of the previous works using our novel dry-season resilience method. This allows us also to begin to determine where the uncertainties lie. Our work also compliments that of Poulter *et al.* (2010), who perturb parameters within a different DGVM (LPJmL) more extensively, and combine this with 8 different GCMs. We provide some uncertainty associated with TRIFFID, as well as having a fully coupled framework where forest changes both locally and globally feedback on the climate response, exploring more uncertainty within this. While Galbraith *et al.* (2010) suggests that TRIFFID is insensitive to a drying climate in regards to changes in vegetation carbon compared to an increasing temperature, Good *et al.* (2011) suggest that both are equally important. Both of these analyses are carried out on the

standard configuration of HadCM3, leaving open questions regarding the importance of these factors in a perturbed physics ensemble.

In conclusion, we have highlighted the uncertainty in projections of the Amazon rainforest due to uncertainties in climate change and land based processes (in an experiment that explores a broad range of vegetation and climate responses) and thus the importance of reducing these to better determine the forest's outcome. We find that 12.2%, 28.1% and 52.6% of ensemble members show some degree of loss (>5% forest coverage loss) under RCP 2.6, A1B and RCP 8.5 emissions scenarios by 2100 but predict that 45.6%, 57.9% and 70.1% respectively are committed to this amount of loss due to 21st century climate change. Our predictions of committed rainforest change show that even under the most intense mitigation, the forest may not be sustainable, despite appearing to be at the end of the 21st century, suggesting that planning beyond 2100 is essential.

Chapter 5

Slowing down of North Pacific climate variability and its implications for abrupt ecosystem change

This chapter is based on Boulton, C. A. & Lenton T. M., 'Slowing down of North Pacific climate variability and its implications for ecosystem regime shifts', 2015, PNAS, 112, 37, 11496-11501

Abstract

Marine ecosystems are sensitive to stochastic environmental variability, with higher-amplitude, lower-frequency – i.e. 'redder' – variability posing a greater threat of triggering large ecosystem changes. Here we show that fluctuations in the Pacific Decadal Oscillation (PDO) index have slowed down markedly over the observational record, as indicated by a robust increase in autocorrelation. This 'reddening' of the spectrum of climate variability is also found in regionally averaged North Pacific sea surface temperatures (SSTs), and can be at least partly explained by observed deepening of the ocean mixed layer. The progressive reddening of North Pacific climate variability has important implications for marine ecosystems. Ecosystem variables that respond linearly to climate forcing will have become prone to much larger variations over the observational record, whereas ecosystem variables that respond non-linearly to climate forcing will have become prone to more frequent 'regime shifts'. Thus, slowing down of North Pacific climate variability can help explain the large magnitude and potentially the quick succession of well-known abrupt changes in North Pacific ecosystems in 1977 and 1989.

Introduction

Sea surface temperature (SST) fluctuations are well known to exhibit 'red' spectra – with increased power at lower frequencies – even when forced by purely 'white' noise from the atmosphere (Frankignoul & Hasselmann, 1977, Hasselmann, 1976). Individual realizations of a stationary red noise process will typically drift one side of the mean and later switch over, resembling an irregular oscillation. The 'Pacific Decadal Oscillation' (PDO) index (Hare & Mantua, 2000, Mantua *et al.*, 1997) describes one such 'oscillation' – it is the time variation of the dominant spatial pattern of sea surface temperatures (SSTs) >20°N in the Pacific, having removed the seasonal cycle and an overall warming trend. Shifts in the PDO between its positive and negative phases – especially around 1977 and 1989 – have been linked to abrupt changes in salmon productivity (Litzow *et al.*, 2014, Mantua *et al.*, 1997), drought regimes in the U.S. (McCabe *et al.*, 2004), changes in Indian summer monsoon rainfall (Krishnan & Sugi, 2003), and a host of other ecological and climatic time series (Hare & Mantua, 2000). Whilst some early work described the PDO itself as exhibiting 'regime shifts' (Hare & Mantua, 2000, Scheffer *et al.*, 2001), implying the existence of alternative climate attractors, subsequent work has found no evidence for nonlinearity in the PDO or other North Pacific climate indices (Hsieh *et al.*, 2005). Instead North Pacific climate variability can be characterized as a linear, stochastic, red noise process (Rudnick & Davis, 2003). North Pacific marine ecosystems have also been widely described as exhibiting 'regime shifts' between alternative attractors in response to stochastic fluctuations in the physical climate system (Hsieh *et al.*, 2005). Whilst there is evidence for non-linearity in some North Pacific ecosystem time series (Hsieh *et*

al., 2005, Hsieh & Ohman, 2006), other populations appear to linearly track the climate forcing, especially if their generation time matches the damping timescale of the forcing (Hsieh & Ohman, 2006). Thus, one "double integration" model for the overall behavior is that SSTs (and indices derived from them such as the PDO) integrate white noise forcing from the atmosphere to produce red noise (Hasselmann, 1976), and ecosystems further integrate this red noise to create time series that are even redder (Di Lorenzo & Ohman, 2013). The redder the ocean climate variability, the more closely a marine ecosystem will track it (Steele *et al.*, 1994), making the ecosystem more prone to large changes driven by the climate variability (Steele *et al.*, 1994).

Here we consider whether North Pacific climate variability has changed over the observational record. A simple generic model for SST fluctuations (Rudnick & Davis, 2003) is the first-order linear autoregressive process:

$$x_t = \alpha x_{t-1} + \sigma\eta_t \qquad (5.1)$$

where the subscript t denotes time, x is the time series, α is the lag-1 autocorrelation (AR(1)) coefficient ($0 < \alpha < 1$, for red noise) and η_t is Gaussian white noise of amplitude σ. Existing work has used a constant value of $\alpha = 0.95$ (corresponding to a de-correlation timescale of 20 months) to mimic the PDO index (Rudnick & Davis, 2003). If, however, there has been a change in the spectrum of SST variability we expect this to manifest as a change in α, with an accompanying change in variance (Carpenter & Brock, 2006, Ditlevsen & Johnsen, 2010). Hence, for each of a series of North Pacific datasets, we estimated the AR(1) coefficient (α) and calculated the variance, in a sliding window moved through the dataset, to look for any trends (see Methods). In

each case, the results were tested against a null model for North Pacific SST fluctuations, with fixed α = 0.95 (Rudnick & Davis, 2003), which was used to generate 10,000 realizations of a series of the same length as the time series being tested (see Methods) – the idea being that a relatively short realization of such a stationary red noise process can display a wide range in trends of lag-1 autocorrelation and variance.

Results

Slowing down of fluctuations is visible in the original PDO index (Fig. 5.1a), with a strongly increasing trend in the AR(1) coefficient (Kendall τ = 0.941, see Methods), from an estimated α = 0.57 to α = 0.95 over the 112.5 year record (Fig. 5.1b, see Methods). Variance also generally increases (Fig. 5.1c), but not as strongly across the whole record (τ = 0.599), due to a downturn in recent decades. The observed trend in AR(1) in the PDO index lies outside the distribution resulting from 10,000 realizations of the null model (Fig. 5.1d). Hence we can reject with high confidence ($p < 0.0001$) the null hypothesis that the observed trend in autocorrelation in the PDO index is the result of a red noise process with fixed α. The increasing trend in variance in the PDO index (τ = 0.599) is less significant and could occur by chance with p = 0.144 (Fig. 5.2a). However, other aspects of North Pacific climate are showing increases in variance (Sydeman *et al.*, 2013), and if we just consider the data up to 1989 it shows a strong increase in variance (τ = 0.909).

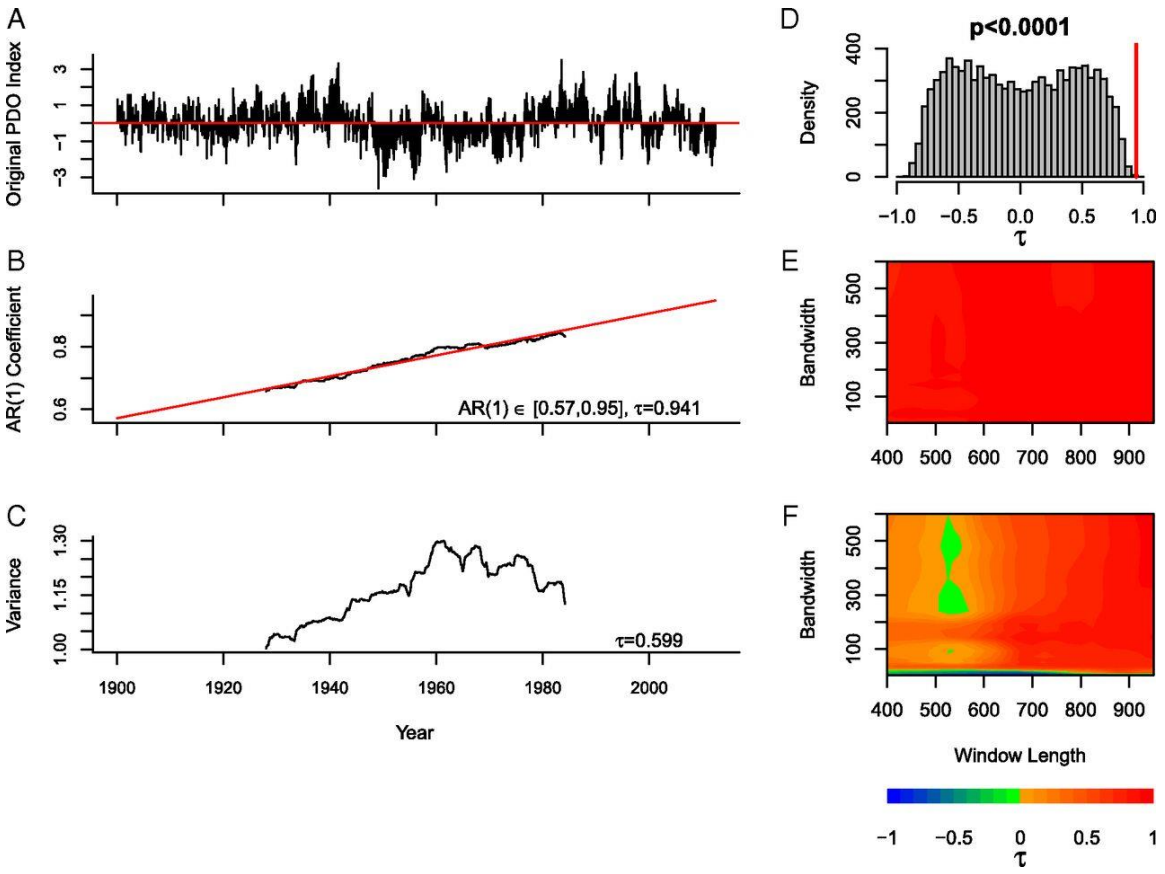

Figure 5.1: Slowing down observed in the Pacific Decadal Oscillation index 1900-2012. (a) The original PDO index. (b) Estimated increase in AR(1) coefficient, using a window length of 675 points (half the series), without detrending (see Methods), results plotted in the middle of the sliding window. (c) Estimated increase in variance. Trends in the indicators are expressed as Kendall τ values. (d) Range in τ values expected from ensembles of 10000 realisations of a null model with fixed α = 0.95. Red vertical line denotes the τ value found in the PDO index (a) ($p < 0.0001$). (e) and (f) are sensitivity analyses of the AR(1) coefficient estimate and variance respectively by testing the value of τ for a variety of window lengths and bandwidths.

Non-stationary behavior can cause autocorrelation to increase hence we examined the effect of further detrending the data before analysis. As the

bandwidth is decreased, the lowest frequencies, including the (multi) decadal 'oscillation' itself, are the first to be removed and the results are limited to ever shorter timescale fluctuations. As a further sensitivity analysis, we also varied the sliding window length in which the indicators are calculated. The positive trend in AR(1) is robust to varying sliding window length and filtering bandwidth (Kendall τ = 0.77–0.98, Fig. 5.1e). The positive trend in variance is also fairly robust (Fig. 5.1f), except when using a very short filtering bandwidth for detrending. This leaves only the highest frequency variability in the index and its variance decreases (e.g. τ = -0.903 with the shortest filtering bandwidth). This is consistent with a shift in power from high to low frequencies, which can also be seen in the changing power spectrum of the data (Fig. 5.3a). Comparing to the null model (also filtered with the shortest bandwidth prior to analysis) the decline in variance at high frequencies has P = 0.002 and we can reject the null model at 5% significance (Fig. 5.2b).

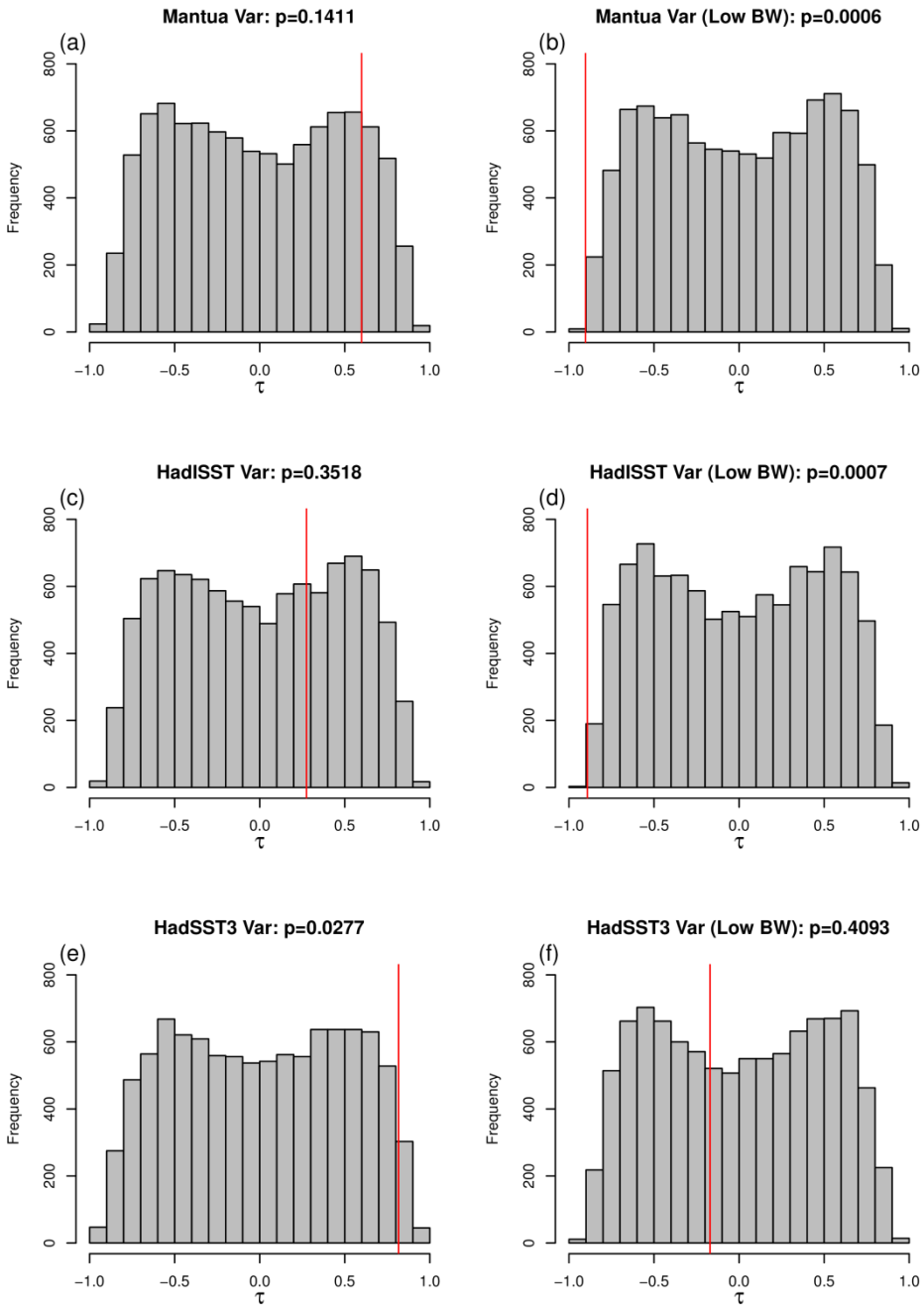

Figure 5.2: Assessing the significance of trends observed in (a,c,e) variance and (b,d,f) variance at high frequencies (having used a low bandwidth for detrending the original data), for: (a,b) Mantua PDO index, (c,d) HadISST average North Pacific index, and (e,f) HadSST3 average North Pacific index. The histogram in each case represents 10000 runs of a null model (described in the main text), the red line is the result from analysis of the corresponding index, and P-values for each hypothesis test are shown above each plot.

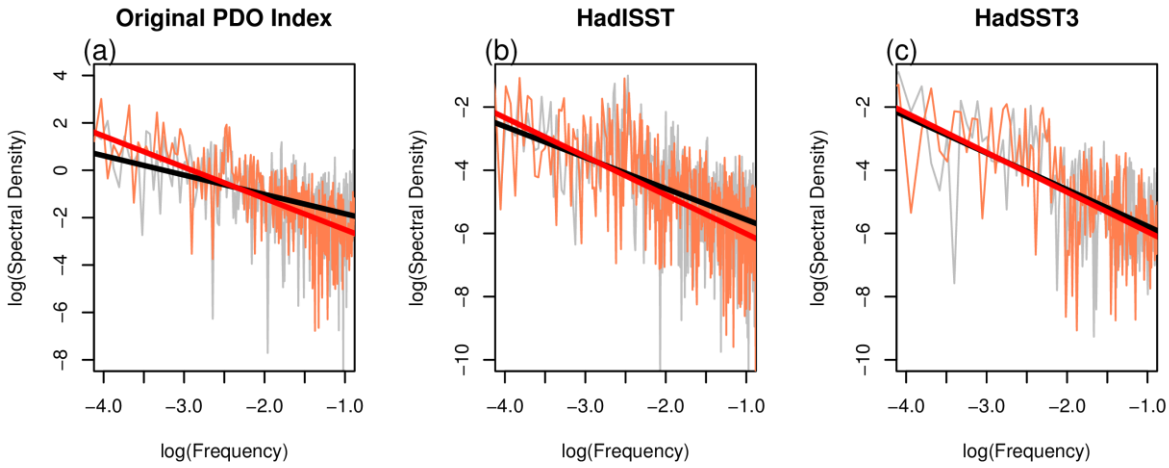

Figure 5.3: Power spectra of SST variability (log-log plots) for: (a) Mantua PDO index, (b) HadISST average North Pacific index, and (c) HadSST3 average North Pacific index. In each case, power spectra for the first half of the data are shown in grey and the second half in orange. A regression line has been fitted through each of these (black and red respectively), showing a shift in power from high to low frequencies in all three time series. There is no overlap between the 95% confidence intervals on the gradient of the regression lines when using the PDO index (a), an overlap of ~4% the width of the 95% confidence intervals for HadISST (b) and an overlap of ~75% the width of the confidence intervals for HadSST3 (c).

To establish whether slowing down has anything to do with the spatial pattern of the PDO we examined the HadISST (Rayner *et al.*, 2003) dataset. Removing seasonal and overall warming trends and then simply averaging HadISST over the North Pacific domain 20-60°N (see Methods), slowing down is again visible by eye (Fig. 5.4a) and confirmed by a strongly increasing trend in the AR(1) coefficient ($\tau = 0.938$; Fig. 5.4b) with estimated change in α from 0.68 to 0.92

(the 95% confidence interval on the change in α hardly alters these results given to 2 significant figures). There is a weaker increasing trend in variance (τ = 0.275; Fig. 5.4c). For the trend in AR(1), a null model of fixed α = 0.95 can be rejected at P < 0.0001 (Fig. 5.4d), whereas the positive trend in variance is not significant (P = 0.352; Fig. 5.2c) again due to a decline in variance in recent decades. The positive trend in the AR(1) coefficient is robust to varying detrending bandwidth and sliding window length (τ = 0.70-0.95; Fig. 5.4e). Variance again decreases at the highest frequencies (τ = -0.892 for the shortest filtering bandwidth, Fig. 5.4f), with P = 0.0007 against the null model (Fig. 5.2d), consistent with a shift in power from high to low frequencies that is seen in the power spectrum (Fig. 5.3b). Thus, slowing down of North Pacific SST fluctuations is not particularly associated with the spatial pattern of PDO variability.

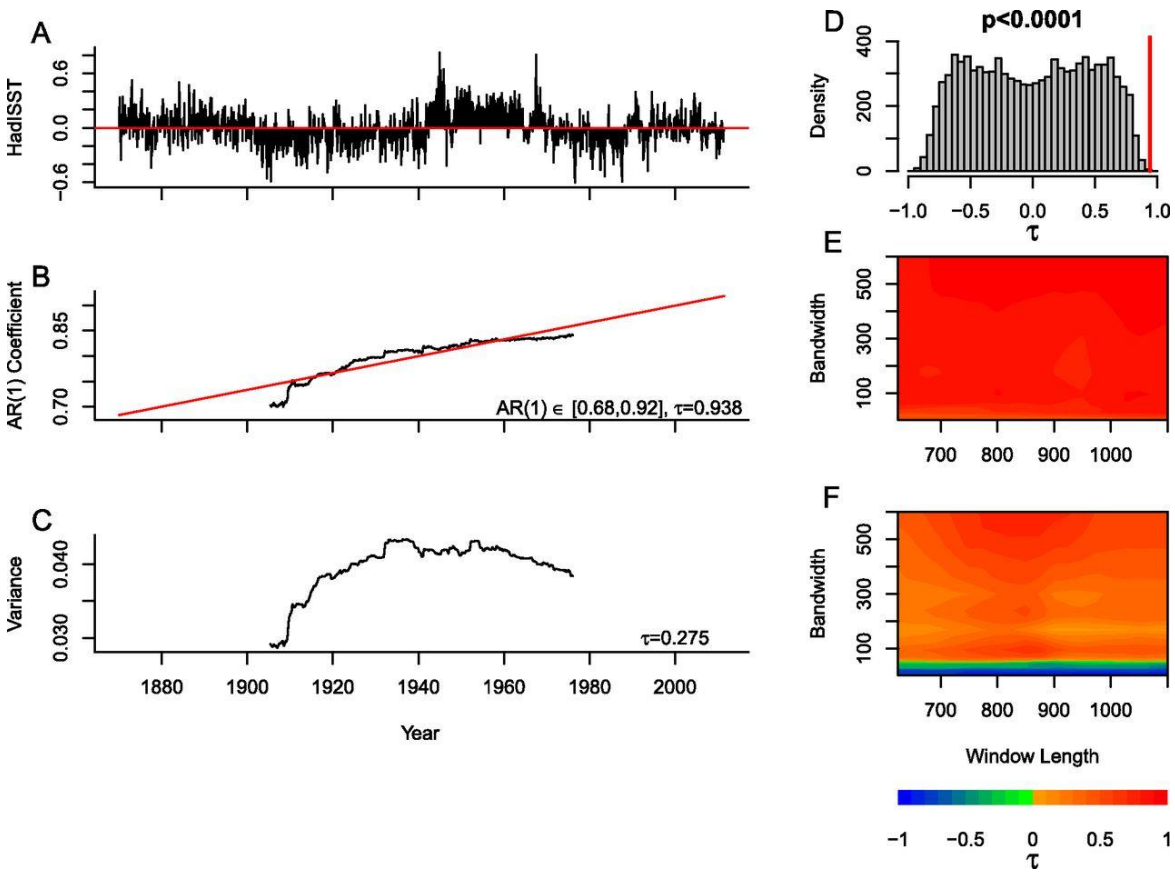

Figure 5.4: Slowing down in average North Pacific reconstructed sea surface temperatures 1870-2011. (a) HadISST data detrended and averaged over the North Pacific (see Methods). (b) Estimated increase in AR(1) coefficient, using a window length of 850 points (half the series), without detrending (see Methods), results plotted in the middle of the sliding window. (c) Estimated increase in variance. (d) Range in Kendall τ values expected from ensembles of 10000 realisations of a null model with fixed $\alpha = 0.95$, with red vertical line denoting $\tau =$ 0.938 found in (a) ($P < 0.0001$). Sensitivity analyses of (e) the AR(1) coefficient estimate and (f) variance by testing the value of τ for a variety of window lengths and bandwidths.

We also analyzed the ERSST v3 dataset (Smith *et al.*, 2008b) and the results for average North Pacific SSTs also show a positive trend in AR(1) over time,

but there are considerable differences in the early part of the two datasets and hence the corresponding AR(1) estimates (Fig. 5.5). Data sampling was generally sparser in the past hence datasets are subject to more infilling further back in time, which in turn could affect measures of autocorrelation. Hence we analyzed the original HadSST3 dataset (Kennedy *et al.*, 2011) without infilling. We focus on the interval 1950 onwards (Fig. 5.6a) because data collection in the North Pacific was sparse before that. Despite the much shorter time interval, we find increasing AR(1) ($\tau = 0.832$; Fig. 5.6b) with α increasing from 0.72 to 0.93, accompanied by an increase in variance ($\tau = 0.816$; Fig. 5.6c). The increase in AR(1) has P = 0.0042 against the null model (Fig. 5.6d) and the trend in variance has P = 0.028 (Fig. 5.2e). Increasing trends in AR(1) and variance are robust to changes in sliding window length and detrending bandwidth (Fig. 5.6e,f). A slight decrease in variance at high frequencies (τ = -0.171 for the shortest filtering bandwidth, Fig. 5.6f), is not significant (P = 0.430, Fig. 5.2f), commensurate with only a slight shift of power to lower frequencies (Fig. 5.3c). Nevertheless, slowing down of North Pacific SST fluctuations has occurred just since 1950.

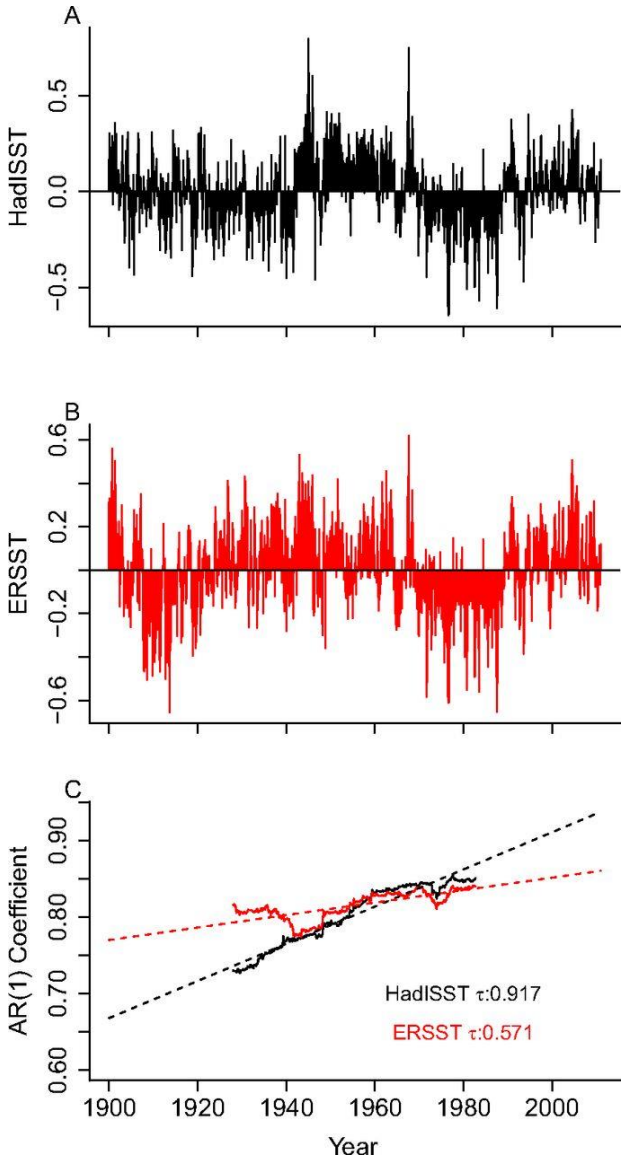

Figure 5.5: Time series of average North Pacific sea surface temperatures from (a) HadISST and (b) ERSST v3 (1900 onwards, with the average annual cycle and quadratic warming trend removed). (c) Estimates of the AR(1) coefficient for each time series using a window length of 675 points (approximately half the time series), plotted at the middle of the time window it is calculated on. A regression line is fitted through each of these AR(1) estimates (dotted line), and Kendall τ values for the trends in AR(1) are given.

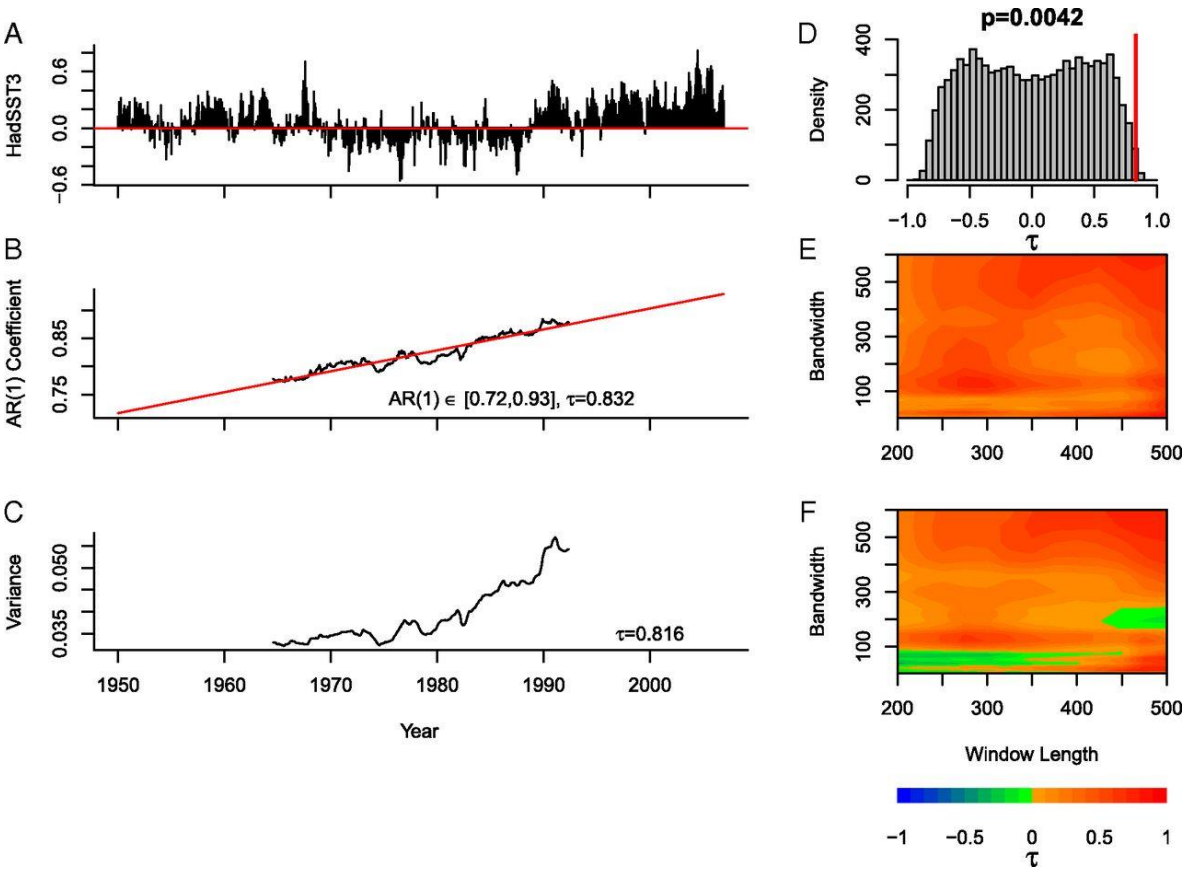

Figure 5.6: Slowing down in average North Pacific sea surface temperature raw data 1950-2006. (a) HadSST3 data detrended and averaged over the North Pacific (see Methods). (b) Estimated increase in AR(1) coefficient, using a window length of 350 points (half the series), without detrending (see Methods), results plotted in the middle of the sliding window. (c) Estimated increase in variance. (d) Range in Kendall τ values expected from ensembles of 10000 realisations of a null model with fixed $\alpha = 0.95$, with red vertical line denoting $\tau =$ 0.832 found in (a) (P=0.0042). Sensitivity analyses of (e) the AR(1) coefficient estimate and (f) variance by testing the value of τ for a variety of window lengths and bandwidths.

A grid point by grid point analysis of HadSST3 data was conducted to determine where the slowing down of SST fluctuations is occurring, revealing that it is

widespread but not universal across the North Pacific domain (Fig. 5.7a,b). Strong slowing down occurs around the basin edge, e.g. off the West Coast of North America. To try and explain this slowing down we consider a simple physical mechanism consistent with known climatic trends, namely deepening of the mixed layer (Carton *et al.*, 2008). In the original model of Frankignoul & Hasselmann (1977) the key environmental variables affecting α are mixed layer depth (*h*) and average wind speed (*U*):

$$\alpha = 1 - kU/h \qquad (5.2)$$

We estimate the constant, *k* (s month^{-1}), for each grid point using the mean values of *U* (from NCEP/NCAR reanalysis (Kalnay *et al.*, 1996)), *h* (from observations (Carton *et al.*, 2008)) and α (from the midpoint of the fitted trend, Fig. 5.7b) (see Methods). Trends in reanalysis (Kalnay *et al.*, 1996) wind speed (*U*) across the domain are mixed and generally small so for simplicity we hold *U* constant, and consider the change in *h* required in Eqn. 5.2 to explain observed trends in α. We compare this to observed changes in mixed layer depth (MLD). Due to limited availability of MLD data we start the analysis in 1960.

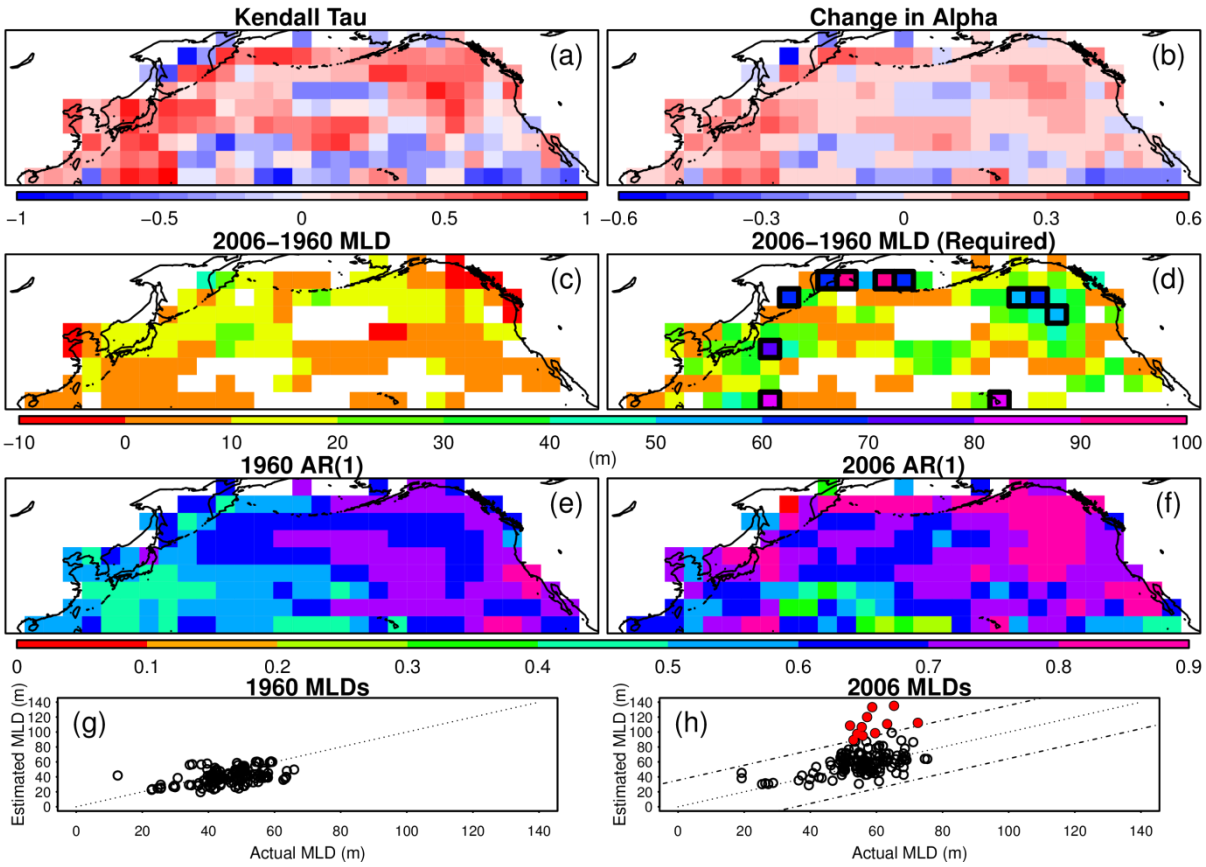

Figure 5.7: Spatial analysis of slowing down in SSTs 1960-2006 and whether it can be explained by increasing mixed later depth. Using individual grid points from HadSST3 (1960-2006) along with mixed-layer depth (MLD) and wind speed data in Eqn. 5.2 (see Methods). Analysis of SST time series at individual grid points: (a) Kendall's τ values for the trend in AR(1) coefficient, and (b) the estimated change in AR(1) coefficient (α) from 1960 to 2006 based on fitting a linear trend. For grid points that exhibit slowing down (increasing α): (c) the observed MLD change (1960-2005), and (d) required MLD change to explain the slowing down signal following Eqn. 5.2. Observed (e) 1960 and (f) 2006 MLD are shown for each grid point. Observed versus estimated (required) MLD in; (g) 1960, and (h) 2006. In (h), outliers that are more than 2 standard deviations (dash-dotted lines) away from fitting the observed value are shown as red points, and the corresponding spatial locations are outlined in (d).

In areas where slowing down (increasing α) is observed, the mixed layer has generally deepened since 1960, typically by 0-20 m (Fig. 5.7c). However, the changes in MLD required to explain the slowing down (Fig. 5.7d) typically exceed those observed (Fig. 5.7c), with a few striking regions where very large increases in MLD would be required (outlined in Fig. 5.7d). A caveat here is that the simple model (Frankignoul & Hasselmann, 1977) chosen cannot explain the power spectrum of SST anomalies in regions strongly influenced by oceanic processes (Reynolds, 1978), including near Japan in the Kuroshio Current. However, the model (Frankignoul & Hasselmann, 1977) is valid (Reynolds, 1978) in areas of strong slowing down in the central and northeast North Pacific (Fig. 5.7d). A further caveat is that MLD is closely related (Kaplan *et al.*, 2000) to the PDO via changes in the strength of the Aleutian Low pressure system, questioning the model assumption that MLD is an independent forcing parameter. Thus, we can partly explain observed slowing down of North Pacific SST fluctuations as due to deepening of the mixed layer (effectively giving the surface ocean a greater heat capacity), but we leave it to future work to fully explain the signal.

To examine how slowing down of North Pacific SST fluctuations might affect the variability in marine ecosystems, we tested how two simple models behave when forced by different levels of red noise, encompassing the range of values of α found from analysis of the PDO index (see Methods). The first is a bi-stable model with two alternative attractors (Eqn. 5.3, see Methods), which represents the concept that some ecosystem time series react nonlinearly to climate forcing and can exhibit true 'regime shifts' between attractors (Hsieh *et al.*, 2005). The specific model chosen is generic and not based on a specific real world system. It is set up to have no bias towards either state and to allow shifts

between states to occur when α = 0.55 in the forcing time series (the lowest value considered). The second model is a 'double integration' model which represents the concept that some ecosystem variables have only one state and linearly track climate forcing, integrating it with their own characteristic time scale (Hsieh & Ohman, 2006). The model is based on one used elsewhere to simulate *Nyctiphanes simplex* (a species of krill) which has a damping timescale of approximately 20 months (Di Lorenzo & Ohman, 2013) (Eqn. 5.4, see Methods). For forcing these ecosystem models, we note that although there is an increase in variance in the climate indices we have analyzed, the increase in amplitude of variability is very modest compared to that generated by increasing α in our AR(1) model. Hence we normalize (dividing by standard deviation) the amplitude of the different levels of red noise (generated by different α) used to force the two idealized ecosystem models. This means that the resulting changes in the ecosystem models are due to the changing memory in climate forcing alone. The range of values of α we explore correspond to damping timescales from ~2.2 months when α=0.55, to 20 months when α=0.95, there being an exponential relationship between α and the timescale, and noting that α=0.95 is often used to mimic PDO variability (Rudnick & Davis, 2003).

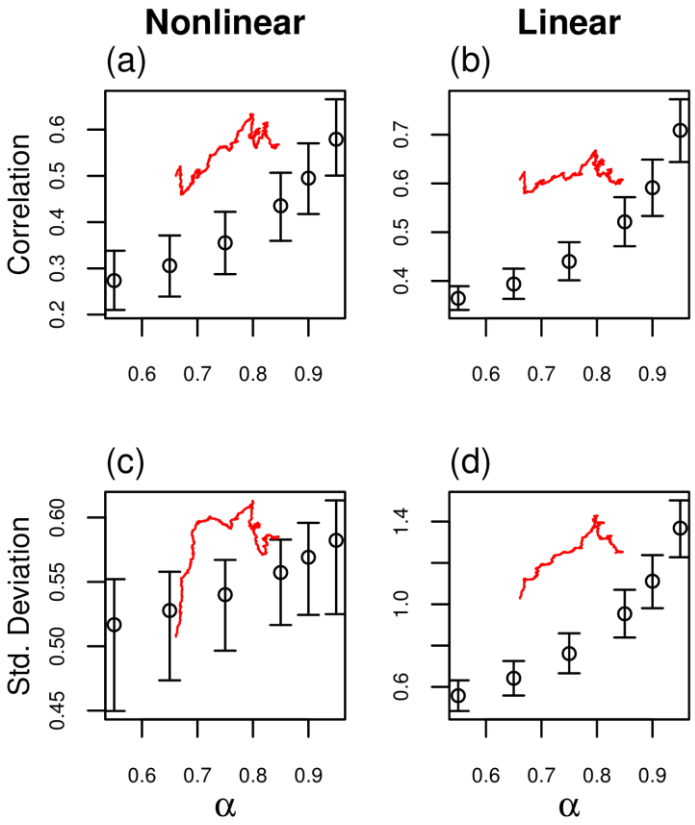

Figure 5.8: The effect of reddening climate forcing on two simple models representing marine ecosystems (see Methods). Correlation between the (a) nonlinear and (b) linear model time series and the forcing time series for different values of α are shown with the 5th and 95th percentiles from 1000 simulations at each value of α. The mean standard deviation of the ensemble for the (c) nonlinear and (d) linear model time series are shown with the 5th and 95th percentiles. In all four plots, the red lines show the same analysis when the original PDO index is used to force the simple models and plotting the α value from this against the other statistics when using a moving window (see Methods).

Increasing the autocorrelation in the climate forcing causes both the nonlinear and linear ecosystem models to become more correlated with the forcing (Fig. 5.8a,b). When forcing the models with the PDO index, the correlation also increases as α increases (red lines). This is expected as the timescale of the ocean is increasing toward the assumed timescale of the ecosystems. As autocorrelation in driving SSTs increases, the standard deviation of variations in both the nonlinear and linear ecosystem models generally increases (Fig. 5.8c,d) and this is much more pronounced in the linear model (Fig. 5.8d). When using the PDO index to force the simple models, we again find that increases in α are generally linked to increases in standard deviation in both ecosystem models (red lines). Examining some specific instances (Fig. 5.9), as autocorrelation in driving SSTs increases (Fig. 5.9a-c), regime shifts in the nonlinear system become more frequent (Fig. 5.9d-f), and the linear system shows increasing deviations from its single equilibrium (Fig. 5.9g-i). Thus, larger ecosystem changes are associated with increasing memory in the climate forcing, especially in the linear ecosystem model, and if the ecosystem response is non-linear, then shifts between different regimes become more frequent.

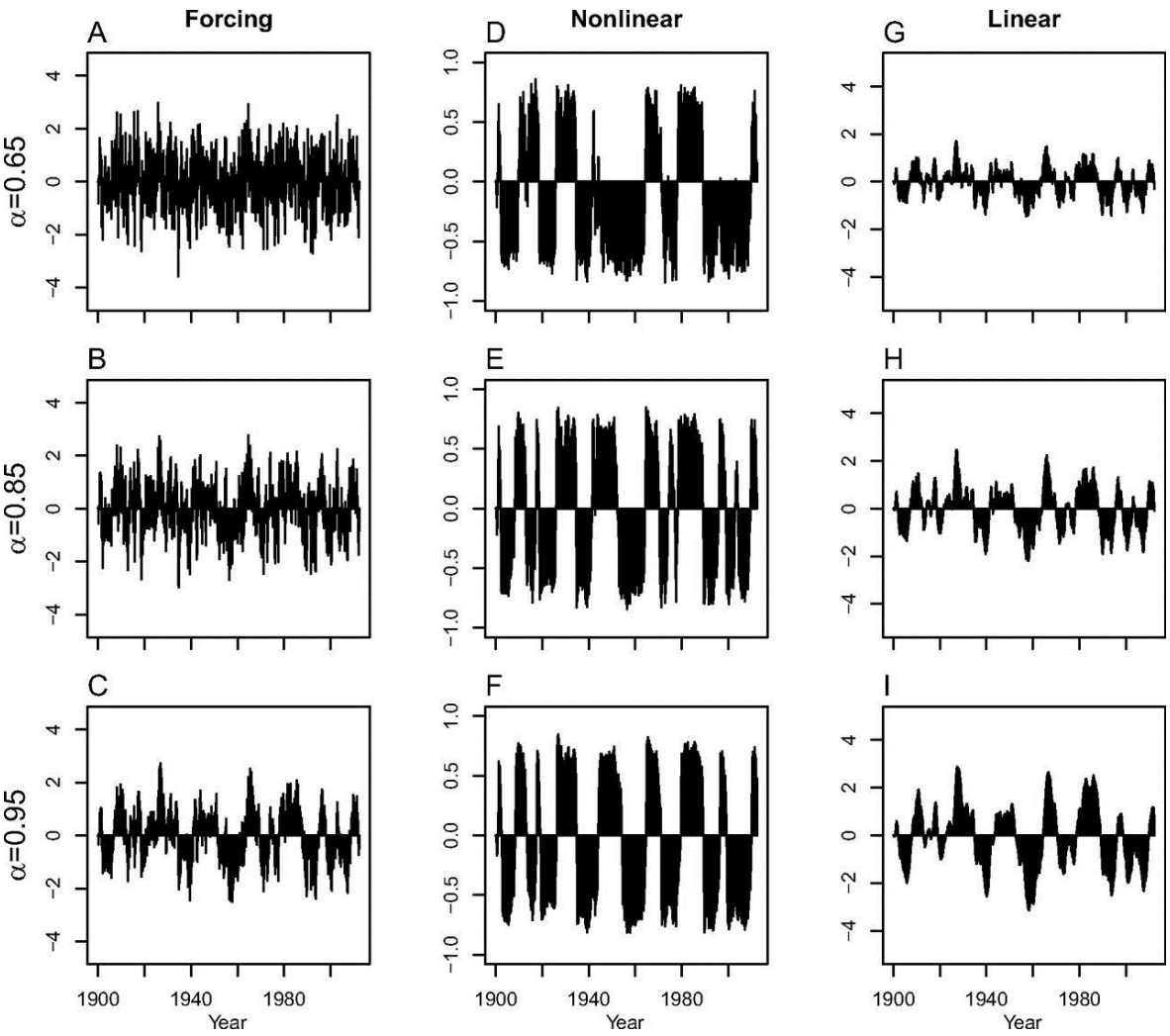

Figure 5.9: Example ensemble members from the forcing of two simple model ecosystems with different levels of climatic red noise by altering the value of α (see Methods). Forcing series (a-c), generated from α values of 0.65, 0.85 and 0.95 respectively, are applied to the nonlinear model, creating time series (d-f), and the linear model, creating time series (g-i).

Discussion

In summary, we detect strong slowing down (‘reddening’) of North Pacific SST fluctuations, and of the PDO index constructed from them, over the observational record. Slowing down since 1960 can be at least partly explained by observed deepening of the ocean mixed layer. It represents a systematic change toward lower frequency, somewhat higher amplitude North Pacific climate variability. Two of the resulting transitions in the PDO index, around 1977 and 1989 are well-known to have had significant impacts on a diverse range of ecological and climate systems (Hare & Mantua, 2000, Mantua *et al.*, 1997). Marine ecosystems, both those that have nonlinear dynamics (Hsieh *et al.*, 2005) and those that linearly track climate forcing (Di Lorenzo & Ohman, 2013, Hsieh & Ohman, 2006) are vulnerable to large and sometimes abrupt changes in response to the low-frequency variability in the physical ocean (Steele *et al.*, 1994). Our results suggest that ecosystem variables that respond linearly to climate variability became prone to larger changes over the observational record, as fluctuations in North Pacific SSTs slowed down. Furthermore, those ecosystem variables that respond non-linearly to climate variability became prone to more frequent abrupt ‘regime shifts’. These results may help explain the well-known abrupt changes that occurred in North Pacific ecosystems in 1977 and 1989 (Hare & Mantua, 2000, Mantua *et al.*, 1997). The large size, especially of the 1977 shift, could be seen as a linear response to slowing down in climate variability, whilst the two events in relatively quick succession could be interpreted as a non-linear response to slowing down in climate variability that was less likely to have occurred earlier in the twentieth century.

It is tempting to extrapolate forwards and infer that if the trend towards increasing autocorrelation in the North Pacific ocean were to continue, the propensity for large ecosystem changes would increase. Models are generally poor at simulating observed MLD in the North Pacific (Huang *et al.*, 2014). However, global warming is robustly expected to drive ocean stratification and a decrease in MLD over the North Pacific (Jang *et al.*, 2011), which would tend to reverse the historical trend of slowing down.

Methods

The PDO index (Mantua *et al.*, 1997) is the time variation of the first empirical orthogonal function (EOF) of Pacific SSTs >20°N, derived from the UKMO Historical SST dataset (Parker *et al.*, 1995) (MOHSST) and Reynolds' OI SST datasets (Reynolds *et al.*, 2002, Reynolds & Smith, 1994) (V1 and V2) (Jan 1900–May 2012). We created further indices from the HadISST (Rayner *et al.*, 2003) (Jan 1870–July 2011), HadSST3 (Kennedy *et al.*, 2011) (Jan 1950–Dec 2006) and ERSST v3 (Smith *et al.*, 2008b) (Jan 1900–Dec 2011) datasets using North Pacific grid points (20-60^{o}N) that are complete over the corresponding time spans. The average annual cycle of each grid point was removed, along with a quadratic warming trend (calculated by a regression model fit). Any further detrending used a Kernal smoother of fixed bandwidth, with the bandwidth varied as a sensitivity analysis.

For each resulting series, within a sliding window of half the series, the variance was calculated and the AR(1) coefficient (α) estimated by fitting an auto-regressive model (Eqn. 5.1). The sliding window length was also varied as a sensitivity analysis. The tendency of an indicator to increase or decrease was measured with Kendall's τ rank correlation coefficient (ranging from 1 to -1). A positive value indicates an increasing trend in an indicator, the larger the value, the more robust is that trend. Sensitivity analysis results are given as contour plots of Kendall's τ values.

To determine the significance of our results, we ran bootstrap ensembles of 10000 runs of a null model, using Eqn. 5.1 with fixed α=0.95, to produce series of identical length to each index. Resulting trends in the null model of AR(1) coefficient were calculated using the same window length (half the length of each series) used to generate example indicators from the real PDO index (e.g. Fig. 5.1b,c).

We estimated trends in α for each grid point of HadSST3 (Kennedy *et al.*, 2011) in the North Pacific (20-60^{o}N) that is complete from 1950-2006. The midpoint value of α between 1960 and 2006 (based on linear regression), the mean wind speed from reanalysis, and mean mixed layer depth from points which are 95% complete from 1960-2006, were used to estimate *k* in Eqn. 5.2.

To determine how marine ecosystems could be affected by slowing down in the PDO or SST fluctuations, we forced two alternative simple models with time series of varying red noise.

The first model represents a system with two stable states that will react nonlinearly to forcing (*f*):

$$\dot{y} = -y^3 + \frac{1}{3}y + 0.1f \quad (5.3)$$

This model was chosen to allow the system to sample both states under realistic forcing *f*, but not too frequently.

The second model, which reacts linearly to the forcing, is given by:

$$\dot{y} = -\frac{y}{20} + 0.1f \quad (5.4)$$

The $-\frac{y}{20}$term (a damping timescale of 20 months) equates to $\alpha \approx 0.95$, similar to the dampening timescale of 24 months used elsewhere to simulate *Nyctiphanes simplex* (Di Lorenzo & Ohman, 2013).

The forcing time series are created using Eqn. 5.1 in the main text, setting σ=0.5, with α taking one of 6 different values (0.55, 0.65, 0.75, 0.85, 0.9 or 0.95, spanning those found in the PDO index), and *x* is the length of the original PDO index (1350 points). There is a strong relationship between the variance of the resulting time series and the α value used to create them, whereas the increase in amplitude of the climate indices we analyse is more modest. Hence we normalised each forcing series by its standard deviation to ensure results found in the ecosystem time series were independent of this. For each value of α, 1000 forcing time series were created. These were then applied to our two simple example models. We also forced each model with the PDO index.

The models were solved using the Euler method. Correlations between *f* and *y* were calculated using the Pearson product-moment correlation coefficient. When using the PDO to force the models, correlation and standard deviation were calculated in a moving window of 675 months (as in the preceding analysis). In these cases, the ecosystem time series are only 675 points long and each is associated with an α value of the PDO index (Fig. 5.8).

Chapter 6

Discussion

The performance of generic early warning signals

In the previous chapters, the idea of using critical slowing down as an early warning signal of an approach to a tipping point (Wissel, 1984) has been examined in three climate systems, two of which are thought to exhibit tipping point behaviour (Lenton *et al.*, 2008) and could potentially reach thresholds that would cause the tipping point to be passed within this century or the next (Kriegler *et al.*, 2009).

In Chapter 2, generic early warning signals were found in a general circulation model (GCM) when approaching the collapse of Atlantic meridional overturning circulation (AMOC) for which two states can be shown to exist theoretically (Stommel, 1961). For this specific climate system, this builds on previous work that tested for early warning signals of modelled AMOC collapse in low (Kleinen *et al.*, 2003) and intermediate (Held & Kleinen, 2004, Lenton *et al.*, 2009, Livina & Lenton, 2007) complexity models. The searching of critical slowing down and generic early warning systems work best when the tested system can be reduced to 1-dimension and the approach towards the tipping point is governed by the slow change in one parameter. For the AMOC, although in a complex system, the slowly increasing freshwater forcing is a single parameter that is causing the system to approach the collapse. That early warning signals of an approach towards a tipping point are found in a complex model shows some promise for using these methods on real world data.

The majority of literature on Amazon rainforest dieback is focused on the vulnerability and resilience to changes in the climate in GCMs due to modelling differences (Huntingford *et al.*, 2013, Sitch *et al.*, 2008) and differences in

parameter choices (Poulter *et al.*, 2010). However there is little research into the early warning of an approach towards a tipping point of Amazon rainforest dieback. In Chapter 3, generic early warning signals are tested on output from an ensemble of GCMs. Using observable variables that could be observed in the real world from remote sensing on a regional scale (broadleaf fraction) as well as variables that would be measured at a more local level, these indicators generally fail, with no robust suggestion of an approach towards a tipping point. This is due to the nature of the forcing, which is a more realistic 21st century forcing (A1B - Nakicenovic *et al.* (2000)) compared to the slow, linear forcing from the previous chapter. Of course the possibility that the Amazon rainforest model system is too complex for generic early warning signals to work cannot be fully ruled out. There is also not a single parameter that can be attributed to dieback, the closest being NPP which is dependent on other factors (such as temperature and precipitation). Indicators do work on a reduced version of the system when it is forced slowly and linearly, but are not found in this same reduced version when using a simplified version of the forcing used in the full model.

Results from chapters 2 and 3 combined suggest that the one main problem for generic early warning signals when predicting the approach of climate tipping points now and in the future is the nature of the forcing towards them, mainly the speed of the forcing compared to the speed of the internal dynamics of the system. Generic early warning signals based on critical slowing down are expected to fail when the timescales of the forcing and dynamics of the system cannot be clearly separated (Held & Kleinen, 2004). Over the 21st century, systems will be forced with a scenario similar to that used when testing the indicators on Amazon dieback rather than the slow linear forcing towards the

AMOC collapse. The complexity of the system being tested could also be a factor in the performance of the indicators. The collapse of the AMOC was dependent on a single parameter (the increasing rate of freshwater forcing), compared to the Amazon rainforest which was driven by both temperature and rainfall. Only by testing the generic early warning indicators on more systems, described by complex dynamics and under more realistic forcing scenarios will their limits begin to be understood.

The use of the system specific indicator for Amazon dieback gave better results than the generic early warning signals. However due to the nature of the experiment, which contained an ensemble of general circulation models where parameters were perturbed to sample a large range of possible responses (Lambert *et al.*, 2013), a threshold in the indicator could not be determined for the real world. An increase of the indicator was as much an early warning of the Amazon rainforest losing stability as could be found. However if the uncertainty in the perturbed parameters was reduced, a threshold may be determined. The basis of another system specific indicator is explored in Chapter 4, although the main message of this chapter concerns how a system can be slow to react to forcing (i.e. lag) and that the system may appear safe when the tipping point has already been passed (Jones *et al.*, 2009). As previously stated, this is related to the lag in the system due to the fast nature of the forcing compared to the time scale of the internal dynamics (Held & Kleinen, 2004). The combination of generic indicators as well as these more system specific indicators could provide the best prospect for early warning of tipping points.

The significance of early warning signals

Another aspect of early warning signals explored in the thesis was their significance, tested with the use of null models. This is something that has not had a large focus in literature (shown by Dakos *et al.* (2008) in supplementary information) but is important when determining if a strong early warning signal is being given. In the null models used in the chapters of this thesis (see Chapter 1 for an explanation), even when systems are not approaching a bifurcation, strong signals still appear by chance (i.e. false alarms). This means that it is hard to determine if a strong indicator found from a time series is due to the system actually approaching a bifurcation, or if this has occurred due to chance. The use of null models show that generally false alarms have what appear to be strong signals, but those indicators from time series approaching tipping points are stronger and thus more significant. It is important to note that the choice of window length and bandwidth whilst using early warning signals can affect the signals observed. Using the example system from Chapter 1, using a window length equal to half the time series gives an almost bimodal distribution in the Kendall's τ values in the null model, with modes at approximately -0.5 and 0.5 (Fig. 6.1a). Using a smaller window length, these modes begin to converge to a near normal distribution with a mean of 0 (Fig. 6.1b). This means that signals found using a smaller window length do not have to be as strong to be significant, as Fig. 6.1 shows. The increase in the AR(1) whilst using a window length of 100 (~10% of the time series length) is highly significant ($P < 0.001$) with a null model created using the same window length (solid line Fig. 6.1b). However, if the same signal was observed when using the larger window length

of 450 (50% of the length of the time series), then this would not be as significant (P = 0.053).

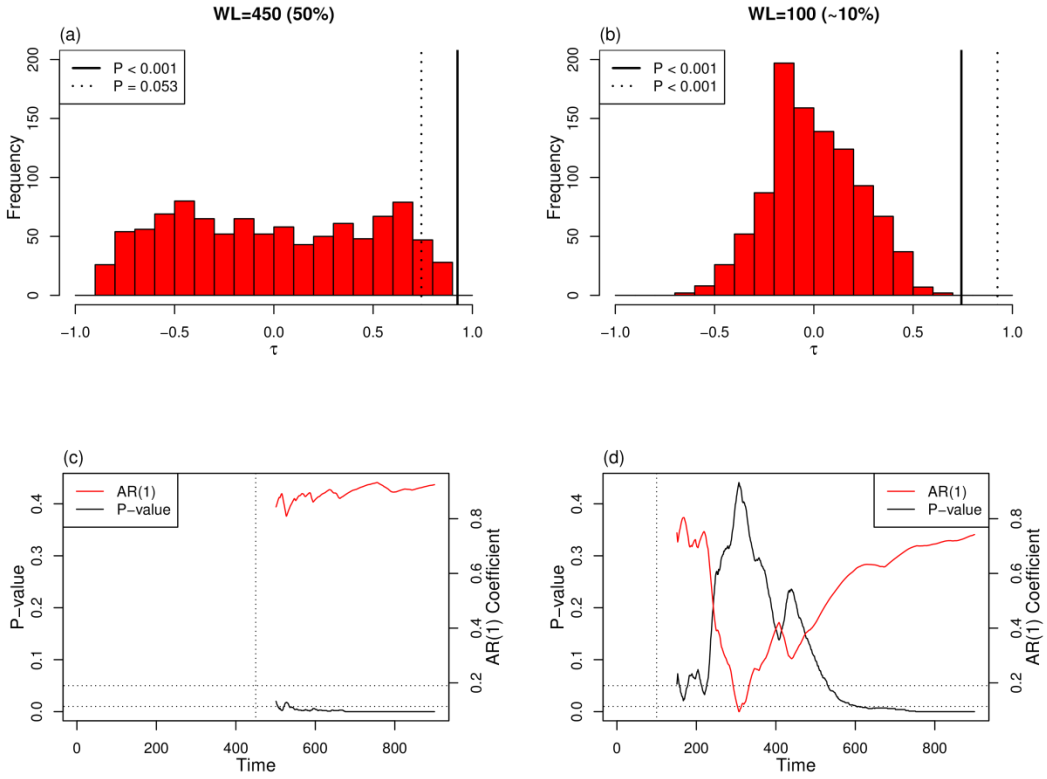

Figure 6.1: The effect of window length on null models used to test the significance of early warning signals. Null models (1000 members) calculated for the example system in Chapter 1 with window length set to (a) 450 (50% of the length of the time series) and (b) 100 (~10% of the length of the time series) and detrending bandwidth set to 100 in both cases (see Chapter 1). Black lines refer to the τ values of the signals observed in the time series approaching a tipping point in Chapter 1, with solid lines showing the signals found when using the same window length as the null model and dotted lines, when using the other. The time to significance (see Chapter 2) is calculated for the corresponding window lengths (c) 450 and (d) 100 (black) with the window lengths denoted by vertical dotted lines. Horizontal dotted lines refer to P-values of 0.05 and 0.01. AR(1) coefficient estimation is shown (red) alongside.

By including the value of the AR(1) coefficient estimation with the analysis on the time to significance (seen in Chapter 2) carried out on the example time series (Fig. 6.1 c,d), it is clearer that signals do not have to be as strong to be as significant when using a smaller window length. There is also some evidence, as in Chapter 2, that the indicator becomes significant at a similar time (after roughly 550-600) regardless of window length chosen. However this is difficult to determine in this case as the significance of the indictor when using the longer window length (Fig. 6.1c) is already high.

The difference in the distribution of null model τ values based on the choice of window length is something which warrants further investigation, allowing users to make more informed decisions on their choices of window length. During the testing of robust testing of early warning signals in Chapter 2 and Chapter 5, decreases in variance were observed whilst using a filtered bandwidth that filtered out everything but the highest frequency noise. When null models were used to assess the significance of these decreases (Chapter 5), there was no obvious change in the null model's distribution regarding the change in bandwidth. However, the effects of filtering bandwidth on significance could also be investigated.

In Chapter 2, this method of calculating significance (Dakos *et al.*, 2008) was used to calculate the length of time series needed to produce a significant result. This important step, which has not appeared in literature before, gives some indication of how far in advance tipping points can be predicted. This type of analysis is important when there is no prior knowledge that the system will tip (i.e. in the real world). Of course, how soon a significant signal is observed compared to the actual occurrence of the tipping point will vary for simulated

systems and their real world counterparts. However using this on the simulated systems gives some indication of how useful the early warning indicators could be. The forcing of the AMOC was slow and linear so testing the time to significance on systems that could potentially reach tipping points in the next century would give more insight into how useful this metric is.

Significance testing in Chapter 2 also examined signals from AMOC time series at different latitudes. As well as providing a more robust result by testing multiple time series of the same collapse, a locational dependence on the early warning signals was also discovered. Although these results were inconclusive (AR(1) and variance were most significant at different latitudes and results varied between the annual and decadal time series), it does suggest that monitoring systems in more than one location would be beneficial to minimise the chance of missing an early warning signal that could be location specific. With new arrays being set up to monitor the AMOC at different latitudes (SAMOC and OSNAP, see Chapter 2), this idea of testing early warning signals in more than one location can be even be done eventually for this specific system.

Why generic early warning signals might fail

In Chapters 2 and 3, lags in both systems are found. In the case of the AMOC, according to the equilibrium runs where the forcing is kept at constant values, the equilibrium loses stability before this would be realised in the transient run under the same forcing. The Amazon rainforest appears to be safe at the end of most of the simulated runs (at the end of the 21st century). However a prediction of its committed change (Jones *et al.*, 2009) shows that the tipping point may have already been passed in a lot of cases. This raises questions about how far away the 'political time horizon' of the systems, could be (Lenton *et al.*, 2008), especially as planning to sustain a system until 2100 does not necessarily mean it will remain 50 years later. Even under the heavy mitigation scenario RCP 2.6 (van Vuuren *et al.*, 2006, van Vuuren *et al.*, 2007) used in the analysis, a lot of simulated forests would still be predicted to dieback.

An important caveat concerning the lags in systems, particularly in the Amazon rainforest, is that they may not be simulated well in GCMs. As mentioned in Chapter 4, the response to recent droughts in the Amazon (such as in 2010) have been very quick, with loss observed within months (Phillips *et al.*, 2009). However drought experiments (Costa *et al.*, 2010) suggest that responses to slower, less harsh increases in temperature and water stress are more likely to happen on longer timescales. However determining if the lag timescales found by Jones *et al.* (2009) for example, are realistic, or to improve the modelling of these lags would be beneficial when testing how useful early warning signals could be in more realistic forcing scenarios.

Another reason generic early warning signals might fail is due to the rate at which the system is sampled, with a lower sampling rate potentially causing missed alarms. The sampling rate of the system has to be small enough to capture the fast time dynamics of the system caused by short term perturbations to the system (Held & Kleinen, 2004). Using the time series from Fig. 1.7 as an example, AR(1) and variance indicators are tested under different sampling rates (Fig. 6.2). The original time series cut prior to the tipping point (1600 points), is first detrended and then sampled at a rate of every 1, 2, 4, 5, 8, 10, 16 and 20 points (i.e. a sampling rate of 2 samples every other point for example). The indicators are then tested on the resulting time series using a window length equal to half the time series each time. This is the same method used in Thomas *et al.* (2015).

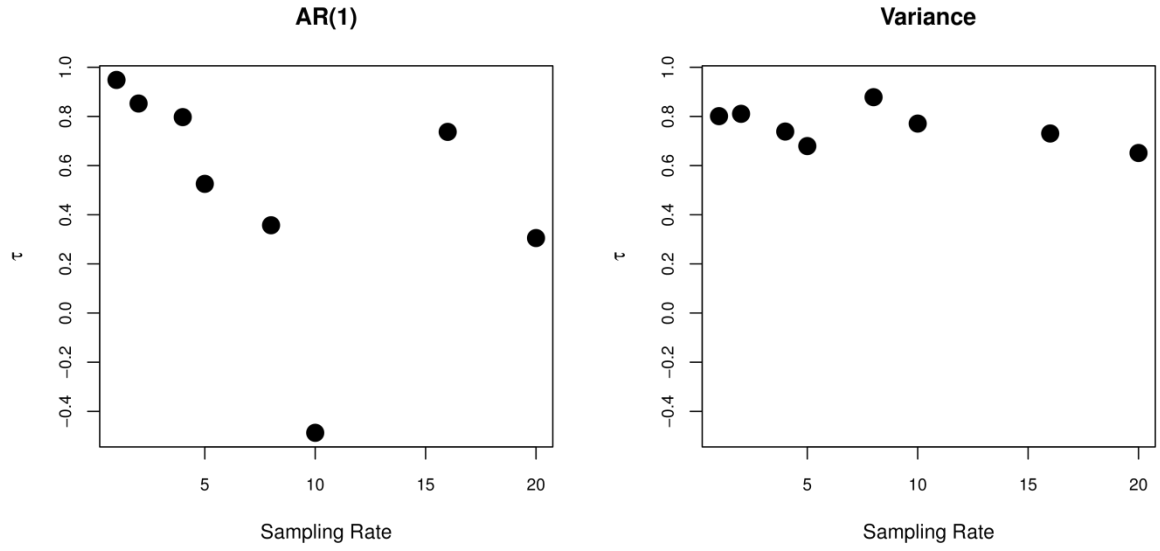

Figure 6.2: The effect of sampling rate (as described above) on (a) AR(1) and (b) variance early warning indicators. Kendall's τ is calculated on the time series of the indicators created when a window length equal to half the length of the time series is used.

It is clear that changing the sampling rate of a system alters the effectiveness of the indicators, especially AR(1) coefficient estimation (Fig. 6.2a). Comparatively, the variance indicator is less affected by less frequent sampling (Fig. 6.2b), but is still reduced. While some of the decline in τ could be attributed to the smaller length of time series used (similar to what is observed in the window length and significance testing above), a negative τ value is found when using the AR(1) indicator (Fig. 6.2a).

The decrease in effectiveness of AR(1) coefficient estimation as an early warning indicator as the sampling rate is decreased is intuitively understandable since the memory of the system is not properly represented in the system. Variance as an indicator however is not dependent on the memory of the system (such as the previous time point) which is why it is arguably unaffected by the sampling rate. A suggestion for future work could be to determine situations where variance becomes a more robust indicator due to sampling rate issues. This topic is especially important when analysing palaeoclimate time series, where the sampling rates are small and irregular, such as in Thomas *et al.* (2015). The effect of the amplitude of noise (strength of short term perturbations on the system), in combination with analysis on the sampling rate, could also be explored, as larger perturbations with an decreased sampling rate are likely to further affect the AR(1) indicator (Thomas *et al.*, 2015).

The potential effect of slowing down on ecosystems

These generic early warning signals show that systems approaching a bifurcation will slow down and become more sluggish in their response to perturbations. However there is little literature on what effect this will have on other systems that they drive. In Chapter 5, the Pacific Decadal Oscillation (PDO) (Mantua *et al.*, 1997) as well as sea surface temperatures in the North Pacific, are shown to be slowing down. Rather than being treated as a system that exhibits different stable states with this slowing down suggesting a regime shift is being approached, it was thought of as a red noise process (Rudnick & Davis, 2003). It is thought that ecosystems which are driven by red noise processes, created by integrating white noise (e.g. atmospheric forcing) such as the PDO, are themselves processes which integrate the red noise forcing (i.e. they are double integrated white noise) (Di Lorenzo & Ohman, 2013). However the effect of an increasing memory in the red noise process (an increasing AR(1) coefficient in the index) driving the ecosystem has not been explored elsewhere. In Chapter 5, observations of the PDO index are used to drive a simple approximation of ecosystems that have previously been linked to the phase of the PDO (Hare & Mantua, 2000, Litzow *et al.*, 2014, Mantua *et al.*, 1997), showing that there is the possibility for these ecosystem to undergo more pronounced and large amplitude regime shifts under a forcing with increased memory. Chapter 5 also shows that this is true of ecosystems that behave non-linearly with respect to the forcing (Hsieh *et al.*, 2005).

While the shifts the PDO appears to undergo are thought of as being part of a red noise process in this thesis, rather than true tipping point behaviour, there is

benefit in understanding how the ecosystems driven by a system with an increasing memory react i.e. this type of behaviour could also happen if the driving system is approaching a tipping point. In this case, the changes in behaviour of these dependent ecosystems (more pronounced shifts and higher variance over time) could act as another early warning system, especially when it is difficult to obtain data on the system approaching a tipping point due to problems for example sparse sampling such that the dynamics of the system are not fully captured at that time scale (Held & Kleinen, 2004). The idea that certain behaviour could be searched for in time series thought to be driven by the system in question has not been implemented as an early warning indicator before. However there is potential for the behaviour change in the affected system(s) to be used as part of an early warning signal for the system that is approaching the tipping point, either by providing information when there is not enough available for the system itself, or by complimenting generic early warning signals tested on it.

Preliminary analysis of global sea surface temperatures shows that slowing down of sea surface temperatures is not limited to just the North Pacific region (Fig. 6.3). Increases in AR(1) coefficient estimation are found on a grid point scale in the majority of the sea, especially in the Southern Hemisphere.

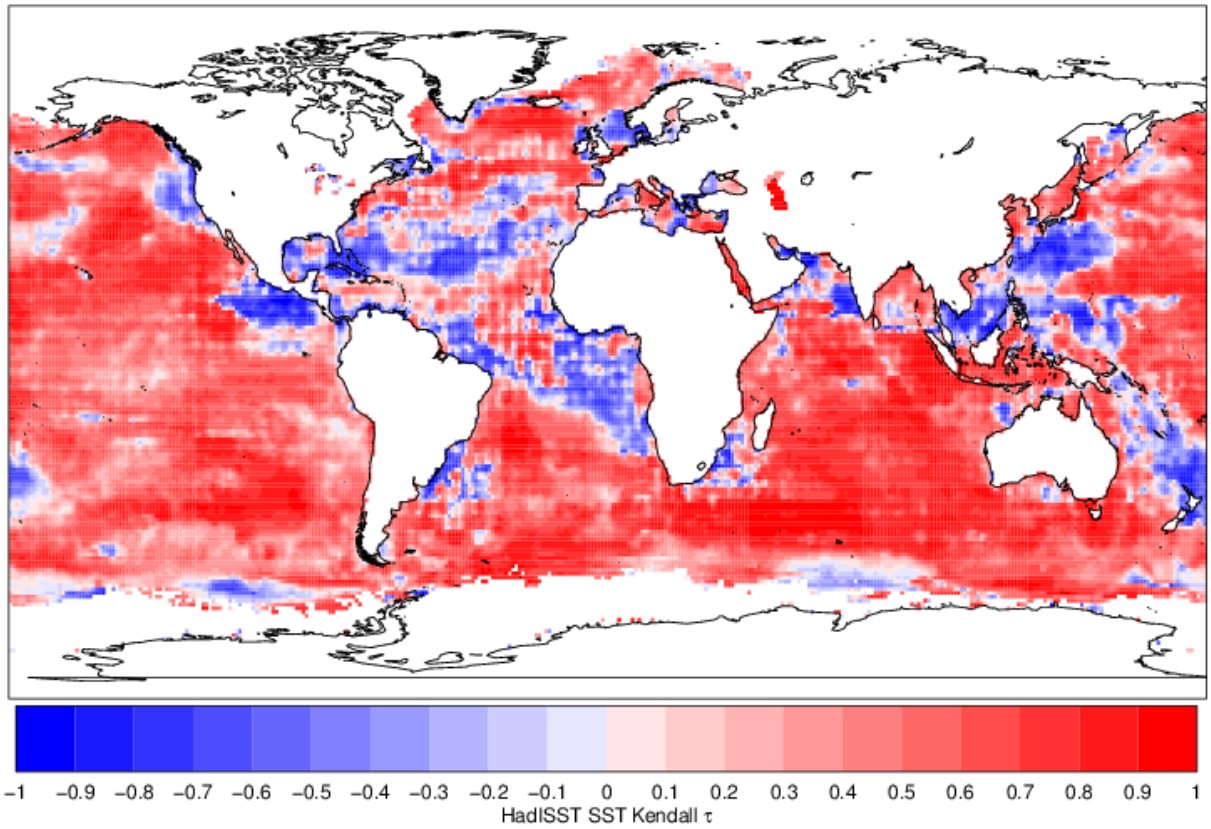

Figure 6.3: Grid point slowing down signals observed in HadISST. Kendall τ values are calculated for the time series of AR(1) estimation coefficient at each grid point which is complete over the dataset. The time series of SSTs have their annual cycle removed and a quadratic warming trend (as in Chapter 5). A window length of 850 months is used to calculate the indicator.

Figure 6.3 only shows the trend of the indicator rather than its change in value over time. Results from Chapter 5 show that the actual value of α is important in determining the effect reddening can have on ecosystems. In the North Atlantic for example, there is a mixture of slowing down and speeding up signals. However the decrease in α of those grid points which do not show slowing down may be small in comparison to the increases in α found in the rest of the region.

The Atlantic Multidecadal Oscillation (AMO; Kerr (2000)), calculated from North Atlantic SSTs, is the dominant mode of variability in the Atlantic, similar to the PDO in the North Pacific. The AMO is linked to changes in temperatures and rainfall over most of the Northern Hemisphere, especially in North America and Europe, with research suggesting that the warm phase of the AMO is linked to more frequent droughts in the US, whereas the cool phase is thought to bring more rainfall (Enfield *et al.*, 2001). The frequency of intense Atlantic hurricanes, as well as rainfall patterns in North East Brazilian and the Sahel are related to the phase of the AMO in modelling studies (Knight *et al.*, 2006, Zhang & Delworth, 2007).

Analysis of the AMO index derived from the Kaplan SST dataset (Enfield *et al.*, 2001) (Fig. 6.4a-c), as well as the raw SSTs in the North Atlantic (0°-60°N, 80°W-0°) from HadISST (Fig. 6.4d-f) show similar (although not as strong) slowing down signals (τ=0.709 and 0.815 respectively, Fig. 6.4b,e) to those observed in the PDO index (τ=0.941) and corresponding Pacific SSTs (τ=0.938). The change in α over the time series is also less than in the Pacific but the AR(1) values found suggest a change in the damping timescale of approximately 7 to 11 months in the Atlantic (using the HadISST SST time series, Fig 6.3).

There is also an increase in variance observed in the raw SSTs (τ=0.839, Fig. 6.4f), although not as strong a signal in found in the AMO index (τ=0.241, Fig. 6.4c). However the increasing AR(1) suggests that the slowing down observed in SSTs and their indices is not limited to the North Pacific and more research should be conducted into the reddening spectra of the oceans overall, as well as the impacts this might have on ecosystems. Unlike the PDO, the AMO is

more closely linked to changes in weather patterns rather than marine ecosystems. However, biological impacts related to the AMO are still observed (Drinkwater *et al.*, 2014), and so the reddening signal observed would still have an effect on Atlantic marine ecosystems according to the double integration hypothesis (Di Lorenzo & Ohman, 2013), which could lead to more abrupt and intense shifts in biological time series.

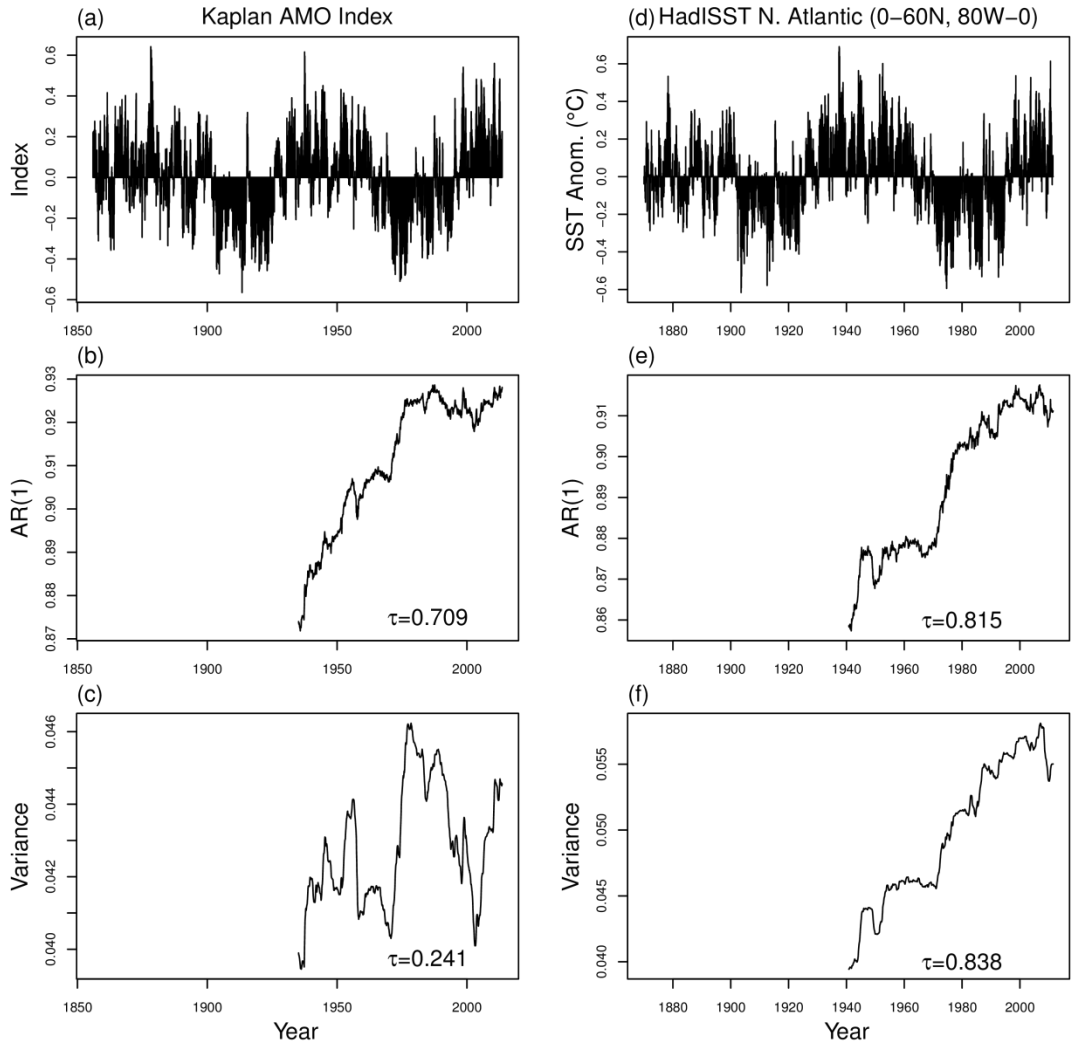

Figure 6.4: Changes in AR(1) coefficient estimation and variance for the AMO index and associated SSTs (0°-60°N, 80°W-0°) from HadISST. Signals from the Kaplan AMO index (a) are calculated on a moving window of 950 months (half the time series) for (b) AR(1) and (c) variance. The time series from HadISST uses the mean value of SSTs in the North Atlantic (0°-60°N, 80°W-0°), with an annual cycle and quadratic warning trend removed (d). A window length of 850 months (half the time series) was used to calculate (e) AR(1) and (f) variance.

Conclusion

The research in this thesis has shown that future tipping points of climate systems have the potential to be predicted in some cases, either with generic early warning signals, or system specific indicators. This is shown in complex GCMs that are forced with realistic 21st century emissions scenarios. However there are some obstacles preventing generic early warning signals from giving strong signals on the approach to a tipping point. These include the nature of the forcing of the system towards the tipping point, which if too fast and/or nonlinear could prevent strong signals from being found. The system could also be too complex for early warning signals to work. In some cases, combining results from generic early warning signals and more system specific indicators could provide a better prospect.

Something that should be considered is lags that occur in the system. In the systems tested in this thesis, there is a considerable time lag between where the tipping point is according to the equilibrium response and when it is actually realised in transient response (on which the early warning signals are tested on). This lag raises questions of the early warning signals, about whether they could still give significant early warning before the tipping point has been reached since they would no longer be near equilibrium, a condition that theory states is needed for these indicators to work effectively. Early warning may be given for a tipping point when it has already been passed but not yet realised.

Finally the critical slowing down observed in systems approaching tipping points could have adverse effects on other dependent systems, with changes in the redness of the forcing causing these systems to behave differently. There

should be considerations for both the health of the dependent system as well as the possibility that the changes in this system could also provide early warning of the tipping point in the forcing system.

References

Ashwin P, Wieczorek S, Vitolo R, Cox P (2012) 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.

Bailey RM (2010) Spatial and temporal signatures of fragility and threshold proximity in modelled semi-arid vegetation. *Proceedings of the Royal Society B: Biological Sciences,* **278**, 1064-1071.

Balan Sarojini B, Gregory JM, Tailleux R *et al.* (2011) High frequency variability of the Atlantic meridional overturning circulation. *Ocean Sci.,* **7**, 471-486.

Balmaseda MA, Smith GC, Haines K, Anderson D, Palmer TN, Vidard A (2007) Historical reconstruction of the Atlantic Meridional Overturning Circulation from the ECMWF operational ocean reanalysis. *Geophysical Research Letters,* **34**, L23615.

Bathiany S, Claussen M, Fraedrich K (2013) Detecting hotspots of atmosphere–vegetation interaction via slowing down – Part 2: Application to a global climate model. *Earth Syst. Dynam.,* **4**, 79-93.

Betts RA (1999) Self-beneficial effects of vegetation on climate in an ocean-atmosphere general circulation model. *Geophysical Research Letters,* **26**, 1457-1460.

Betts RA, Cox PM, Collins M, Harris PP, Huntingford C, Jones CD (2004) The role of ecosystem-atmosphere interactions in simulated Amazonian precipitation decrease and forest dieback under global climate warming. *Theoretical and Applied Climatology,* **78**, 157-175.

Biggs R, Carpenter SR, Brock WA (2009) Turning back from the brink: Detecting an impending regime shift in time to avert it. *Proceedings of the National Academy of Sciences,* **106**, 826-831.

Boettiger C, Hastings A (2012) Early warning signals and the prosecutor's fallacy. *Proceedings of the Royal Society B: Biological Sciences,* **279**, 4734-4739.

Booth BBB, Bernie D, Mcneall D *et al.* (2013) Scenario and modelling uncertainty in global mean temperature change derived from emission-driven global climate models. *Earth Syst. Dynam.,* **4**, 95-108.

Booth BBB, Jones CD, Collins M *et al.* (2012) High sensitivity of future global warming to land carbon cycle processes. *Environmental Research Letters,* **7**, 024002.

Boulton C, Good P, Lenton T (2013) Early warning signals of simulated Amazon rainforest dieback. *Theoretical Ecology,* **6**, 373-384.

Boulton CA, Allison LC, Lenton TM (2014) Early warning signals of Atlantic Meridional Overturning Circulation collapse in a fully coupled climate model. *Nat Commun,* **5**.

Brayshaw DJ, Woollings T, Vellinga M (2009) Tropical and Extratropical Responses of the North Atlantic Atmospheric Circulation to a Sustained Weakening of the MOC. *Journal of Climate,* **22**, 3146-3155.

Broecker WS, Peteet DM, Rind D (1985) Does the ocean-atmosphere system have more than one stable mode of operation? *Nature,* **315**, 21-26.

Brohan P, Kennedy JJ, Harris I, Tett SFB, Jones PD (2006) Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *Journal of Geophysical Research: Atmospheres,* **111**, D12106.

Carpenter SR, Brock WA (2006) Rising variance: a leading indicator of ecological transition. *Ecology Letters,* **9**, 311-318.

Carton JA, Grodsky SA, Lui H (2008) Variability of the oceanic mixed layer. *Journal of Climate,* **21**, 1029-1047.

Cheng W, Chiang JCH, Zhang D (2013) Atlantic Meridional Overturning Circulation (AMOC) in CMIP5 Models: RCP and Historical Simulations. *Journal of Climate,* **26**, 7187-7197.

Cimatoribus AA, Drijfhout SS, Livina V, Van Der Schrier G (2013) Dansgaard–Oeschger events: bifurcation points in the climate system. *Clim. Past,* **9**, 323-333.

Clark PU, Pisias NG, Stocker TF, Weaver AJ (2002) The role of the thermohaline circulation in abrupt climate change. *Nature,* **415**, 863-869.

Collins M, Booth BB, Bhaskaran B, Harris G, Murphy J, Sexton DH, Webb M (2011) Climate model errors, feedbacks and forcings: a comparison of perturbed physics and multi-model ensembles. *Climate Dynamics,* **36**, 1737-1766.

Collins M, Brierley CM, Macvean M, Booth BBB, Harris GR (2007) The Sensitivity of the Rate of Transient Climate Change to Ocean Physics Perturbations. *Journal of Climate,* **20**, 2315-2320.

Cook KH, Vizy EK (2008) Effects of Twenty-First-Century Climate Change on the Amazon Rain Forest. *Journal of Climate,* **21**, 542-560.

Costa ACLD, Galbraith D, Almeida S *et al.* (2010) Effect of 7 yr of experimental drought on vegetation dynamics and biomass storage of an eastern Amazonian rainforest. *New Phytologist,* **187**, 579-591.

Cox PM (2001) Description of the 'TRIFFID' Dynamical Global Vegetation Model. *Hadley Centre Technical Note,* **24**.

Cox PM, Betts RA, Collins M, Harris PP, Huntingford C, Jones CD (2004) Amazonian forest dieback under climate-carbon cycle projections for the 21st century. *Theoretical and Applied Climatology,* **78**, 137-156.

Cox PM, Betts RA, Jones CD, Spall SA, Totterdell IJ (2000) Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature,* **408**, 184-187.

Cox PM, Harris PP, Huntingford C *et al.* (2008) Increasing risk of Amazonian drought due to decreasing aerosol pollution. *Nature,* **453**, 212-215.

Cox PM, Pearson D, Booth BB, Friedlingstein P, Huntingford C, Jones CD, Luke CM (2013) Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature,* **494**, 341-344.

Dakos V, Kéfi S, Rietkerk M, Nes EHV, Scheffer M (2011) Slowing Down in Spatially Patterned Ecosystems at the Brink of Collapse. *The American Naturalist,* **177**, E153-E166.

Dakos V, Scheffer M, Van Nes EH, Brovkin V, Petoukhov V, Held H (2008) Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences,* **105**, 14308-14312.

Dakos V, Van Nes E, Donangelo R, Fort H, Scheffer M (2010) Spatial correlation as leading indicator of catastrophic shifts. *Theoretical Ecology,* **3**, 163-174.

Dakos V, Van Nes E, Scheffer M (2013) Flickering as an early warning signal. *Theoretical Ecology,* **6**, 309-317.

Dakos V, Van Nes EH, D'odorico P, Scheffer M (2012) Robustness of variance and autocorrelation as indicators of critical slowing down. *Ecology,* **93**, 264-271.

De Vries P, Weber SL (2005) The Atlantic freshwater budget as a diagnostic for the existence of a stable shut down of the meridional overturning circulation. *Geophysical Research Letters,* **32**, L09606.

Delworth T, Manabe S, Stouffer RJ (1993) Interdecadal Variations of the Thermohaline Circulation in a Coupled Ocean-Atmosphere Model. *Journal of Climate,* **6**, 1993-2011.

Delworth TL, Mann ME (2000) Observed and simulated multidecadal variability in the Northern Hemisphere. *Climate Dynamics,* **16**, 661-676.

Di Lorenzo E, Ohman MD (2013) A double-integration hypothesis to explain ocean ecosystem response to climate forcing. *Proceedings of the National Academy of Sciences,* **110**, 2496-2499.

Dirzo R, Raven PH (2003) Global state of bioversity and loss. *Annual Review of Environment and Resources,* **28**, 137-167.

Ditlevsen PD, Johnsen SJ (2010) Tipping points: Early warning and wishful thinking. *Geophysical Research Letters,* **37**, L19703.

Dong B, Sutton RT (2005) Mechanism of Interdecadal Thermohaline Circulation Variability in a Coupled Ocean–Atmosphere GCM. *Journal of Climate,* **18**, 1117-1135.

Drijfhout S, Gleeson E, Dijkstra HA, Livina V (2013) Spontaneous abrupt climate change due to an atmospheric blocking–sea-ice–ocean feedback in an unforced climate model simulation. *Proceedings of the National Academy of Sciences,* **110**, 19713–19718.

Drinkwater KF, Miles M, Medhaug I, Otterå OH, Kristiansen T, Sundby S, Gao Y (2014) The Atlantic Multidecadal Oscillation: Its manifestations and impacts with special emphasis on the Atlantic region north of 60°N. *Journal of Marine Systems,* **133**, 117-130.

Enfield DB, Mestas-Nuñez AM, Trimble PJ (2001) The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S. *Geophysical Research Letters,* **28**, 2077-2080.

Feng QY, Viebahn JP, Dijkstra HA (2014) Deep ocean early warning signals of an Atlantic MOC collapse. *Geophysical Research Letters,* **41**, 6009-6015.

Frankignoul C, Hasselmann K (1977) Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability. *Tellus,* **29**, 289-305.

Galbraith D, Levy PE, Sitch S, Huntingford C, Cox P, Williams M, Meir P (2010) Multiple mechanisms of Amazonian forest biomass losses in three dynamic global vegetation models under climate change. *New Phytologist,* **187**, 647-665.

Good P, Jones C, Lowe J, Betts R, Booth B, Huntingford C (2011) Quantifying Environmental Drivers of Future Tropical Forest Extent. *Journal of Climate,* **24**, 1337-1349.

Good P, Jones C, Lowe J, Betts R, Gedney N (2013) Comparing Tropical Forest Projections from Two Generations of Hadley Centre Earth System Models, HadGEM2-ES and HadCM3LC. *Journal of Climate,* **26**, 495-511.

Good P, Lowe JA, Collins M, Moufouma-Okia W (2008) An objective tropical Atlantic sea surface temperature gradient index for studies of south Amazon dry-season climate variability and change. *Philosophical Transactions of the Royal Society B: Biological Sciences,* **363**, 1761-1766.

Gray ST, Graumlich LJ, Betancourt JL, Pederson GT (2004) A tree-ring based reconstruction of the Atlantic Multidecadal Oscillation since 1567 A.D. *Geophysical Research Letters,* **31**, L12205.

Guttal V, Jayaprakash C (2008) Changing skewness: an early warning signal of regime shifts in ecosystems. *Ecology Letters,* **11**, 450-460.

Guttal V, Jayaprakash C (2009) Spatial variance and spatial skewness: leading indicators of regime shifts in spatial ecological systems. *Theoretical Ecology,* **2**, 3-12.

Hare SR, Mantua NJ (2000) Empirical evidence for North Pacific regime shifts in 1977 and 1989. *Progress in Oceanography,* **47**, 103-145.

Harris PP, Huntingford C, Cox PM (2008) Amazon Basin climate under global warming: the role of the sea surface temperature. *Philosophical Transactions of the Royal Society B: Biological Sciences,* **363**, 1753-1759.

Hasselmann K (1976) Stochastic climate models Part I. Theory. *Tellus,* **28**, 473-485.

Hawkins E, Smith RS, Allison LC, Gregory JM, Woollings TJ, Pohlmann H, De Cuevas B (2011) Bistability of the Atlantic overturning circulation in a global climate model and links to ocean freshwater transport. *Geophysical Research Letters,* **38**, L10605.

Hawkins E, Sutton R (2008) Variability of the Atlantic thermohaline circulation described by three-dimensional empirical orthogonal functions. *Climate Dynamics,* **30**, 439-439.

Held H, Kleinen T (2004) Detection of climate system bifurcations by degenerate fingerprinting. *Geophysical Research Letters,* **31**, L23207.

Hirota M, Holmgren M, Van Nes EH, Scheffer M (2011) Global Resilience of Tropical Forest and Savanna to Critical Transitions. *Science,* **334**, 232-235.

Hirschi JJM, Killworth PD, Blundell JR (2007) Subannual, Seasonal, and Interannual Variability of the North Atlantic Meridional Overturning Circulation. *Journal of Physical Oceanography,* **37**, 1246-1265.

Hodson DR, Sutton R (2012) The impact of resolution on the adjustment and decadal variability of the Atlantic meridional overturning circulation in a coupled climate model. *Climate Dynamics,* **39**, 3057-3073.

Hsieh C-H, Glaser SM, Lucas AJ, Sugihara G (2005) Distinguishing random environmental fluctuations from ecological catastrophes for the North Pacific Ocean. *Nature,* **435**, 336-340.

Hsieh C-H, Ohman MD (2006) Biological responses to environmental forcing: The linear trcking window hypothesis. *Ecology,* **87**, 1932-1938.

Huang CJ, Qiao F, Dai D (2014) Evaluating CMIP5 simulations of mixed layer depth during summer. *Journal of Geophysical Research: Oceans,* **119**, 2568-2582.

Huisman SE, Den Toom M, Dijkstra HA, Drijfhout S (2010) An Indicator of the Multiple Equilibria Regime of the Atlantic Meridional Overturning Circulation. *Journal of Physical Oceanography,* **40**, 551-567.

Huntingford C, Zelazowski P, Galbraith D *et al.* (2013) Simulated resilience of tropical rainforests to CO2-induced climate change. *Nature Geosci,* **6**, 268-273.

Jacob D, Goettel H, Jungclaus J, Muskulus M, Podzun R, Marotzke J (2005) Slowdown of the thermohaline circulation causes enhanced maritime

climate influence and snow cover over Europe. *Geophysical Research Letters,* **32**, L21711.

Jang CJ, Park J, Park T, Yoo S (2011) Response of the ocean mixed layer depth to global warming and its impact on primary production: a case for the North Pacific Ocean. *ICES Journal of Marine Science,* **68**, 996-1007.

Johns WE, Baringer MO, Beal LM *et al.* (2011) Continuous, Array-Based Estimates of Atlantic Ocean Heat Transport at 26.5°N. *Journal of Climate,* **24**, 2429-2449.

Johnson HL, Marshall DP (2002) A Theory for the Surface Atlantic Response to Thermohaline Variability. *Journal of Physical Oceanography,* **32**, 1121-1132.

Jones C, Lowe J, Liddicoat S, Betts R (2009) Committed terrestrial ecosystem changes due to climate change. *Nature Geosci,* **2**, 484-487.

Jungclaus JH, Haak H, Latif M, Mikolajewicz U (2005) Arctic–North Atlantic Interactions and Multidecadal Variability of the Meridional Overturning Circulation. *Journal of Climate,* **18**, 4013-4031.

Kalnay E, Kanamitsu M, Kistler R *et al.* (1996) The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society,* **77**, 437-471.

Kaplan A, Cane MA, Kushnir Y, Clement AC, Blumenthal MB, Rajagopalan B (1998) Analyses of global sea surface temperature 1856–1991. *Journal of Geophysical Research: Oceans,* **103**, 18567-18589.

Kaplan A, Kushnir Y, Cane MA (2000) Reduced space optimal interpolation of historical marine sea level pressure: 1854-1992. *Journal of Climate,* **13**, 2987-3002.

Kawase M (1987) Establishment of Deep Ocean Circulation Driven by Deep-Water Production. *Journal of Physical Oceanography,* **17**, 2294-2317.

Kéfi S, Dakos V, Scheffer M, Van Nes EH, Rietkerk M (2013) Early warning signals also precede non-catastrophic transitions. *Oikos,* **122**, 641-648.

Kennedy JJ, Rayner NA, Smith RO, Parker DE, Saunby M (2011) Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties. *Journal of Geophysical Research: Atmospheres,* **116**, D14103.

Kerr RA (2000) A North Atlantic Climate Pacemaker for the Centuries. *Science,* **288**, 1984-1985.

Kleinen T, Held H, Petschel-Held G (2003) The potential role of spectral properties in detecting thresholds in the Earth system: application to the thermohaline circulation. *Ocean Dynamics,* **53**, 53-63.

Knight JR, Folland CK, Scaife AA (2006) Climate impacts of the Atlantic Multidecadal Oscillation. *Geophysical Research Letters,* **33**.

Köhl A (2005) Anomalies of Meridional Overturning: Mechanisms in the North Atlantic. *Journal of Physical Oceanography,* **35**, 1455-1472.

Kriegler E, Hall JW, Held H, Dawson R, Schellnhuber HJ (2009) Imprecise probability assessment of tipping points in the climate system. *Proceedings of the National Academy of Sciences,* **106**, 5041–5046.

Krishnan R, Sugi M (2003) Pacific decadal oscillation and variability of the Indian summer monsoon rainfall. *Climate Dynamics,* **21**, 233-242.

Kuehn C (2011) A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics. *Physica D: Nonlinear Phenomena,* **240**, 1020-1035.

Lade SJ, Gross T (2012) Early Warning Signals for Critical Transitions: A Generalized Modeling Approach. *PLoS Comput Biol,* **8**, e1002360.

Lambert FH, Harris G, Collins M, Murphy J, Sexton DH, Booth BB (2013) Interactions between perturbations to different Earth system components simulated by a fully-coupled climate model. *Climate Dynamics,* **41**, 3055-3072.

Lee T, Marotzke J (1998) Seasonal Cycles of Meridional Overturning and Heat Transport of the Indian Ocean. *Journal of Physical Oceanography,* **28**, 923-943.

Lenton TM (2011) Early warning of climate tipping points. *Nature Clim. Change,* **1**, 201-209.

Lenton TM, Held H, Kriegler E, Hall J, Lucht W, Rahmstorf S, Schellnhuber HJ (2008) Tipping Elements in the Earth's Climate System. *PNAS,* **105**, 1786-1793.

Lenton TM, Livina VN, Dakos V, Van Nes EH, Scheffer M (2012) Early warning of climate tipping points from critical slowing down: comparing methods to improve robustness. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* **370**, 1185-1204.

Lenton TM, Myerscough RJ, Marsh R, Livina VN, Price AR, Cox SJ (2009) Using GENIE to study a tipping point in the climate system. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* **367**, 871-884.

Levermann A, Griesel A, Hofmann M, Montoya M, Rahmstorf S (2005) Dynamic sea level changes following changes in the thermohaline circulation. *Climate Dynamics,* **24**, 347-354.

Lewis SL, Brando PM, Phillips OL, Van Der Heijden GMF, Nepstad D (2011) The 2010 Amazon Drought. *Science,* **331**, 554.

Li W, Fu R, Dickinson RE (2006) Rainfall and its seasonality over the Amazon in the 21st century as assessed by the coupled models for the IPCC AR4. *Journal of Geophysical Research: Atmospheres,* **111**, D02111.

Litzow MA, Mueter FJ, Hobday AJ (2014) Reassessing regime shifts in the North Pacific: incremental climate change and commercial fishing are necessary for explaining decadal-scale biological variability. *Global Change Biology,* **20**, 38-50.

Litzow MA, Urban JD, Laurel BJ (2008) Increased spatial variance accompanies reorgranization of two continental shelf ecosystems. *Ecological Applications,* **18**, 1331-1337.

Livina VN, Kwasniok F, Lenton TM (2010) Potential analysis reveals changing number of climate states during the last 60 kyr. *Clim. Past,* **6**, 77-82.

Livina VN, Lenton TM (2007) A modified method for detecting incipient bifurcations in a dynamical system. *Geophysical Research Letters,* **34**, L03712.

Livina VN, Lenton TM (2013) A recent tipping point in the Arctic sea-ice cover: abrupt and persistent increase in the seasonal cycle since 2007. *The Cryosphere,* **7**, 275-286.

Malhi Y, Aragão LEOC, Galbraith D *et al.* (2009) Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. *Proceedings of the National Academy of Sciences,* **106**, 20610–20615.

Malhi Y, Roberts JT, Betts RA, Killeen TJ, Li W, Nobre CA (2008) Climate Change, Deforestation, and the Fate of the Amazon. *Science,* **319**, 169-172.

Manabe S, Stouffer RJ (1988) Two Stable Equilibria of a Coupled Ocean-Atmosphere Model. *Journal of Climate,* **1**, 841-866.

Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis RC (1997) A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production. *Bulletin of the American Meteorological Society,* **78**, 1069-1079.

Matthews HD, Eby M, Ewen T, Friedlingstein P, Hawkins BJ (2007) What determines the magnitude of carbon cycle-climate feedbacks? *Global Biogeochemical Cycles,* **21**, GB2012.

Mccabe GJ, Palecki MA, Betancourt JL (2004) Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States. *Proceedings of the National Academy of Sciences,* **101**, 4136-4141.

Meinshausen M, Raper SCB, Wigley TML (2008) Emulating IPCC AR4 atmosphere-ocean and carbon cycle models for projecting global-mean, hemispheric and land/ocean temperatures: MAGICC 6.0. *Atmospheric Chemistry and Physics Discussions,* **8**, 6153-6272.

Nakicenovic N, Alcamo J, Davis G, Other, Other, Other, Other (2000) *IPCC Special Report on Emission Scenarios*, Cambridge.

Oyama MD, Nobre CA (2003) A new climate-vegetation equilibrium state for Tropical South America. *Geophysical Research Letters,* **30**, 2199.

Parker DE, Folland CK, Jackson M (1995) Marine surface temperature: Observed variations and data requirements. *Climatic Change,* **31**, 559-600.

Phillips OL, Aragão LEOC, Lewis SL *et al.* (2009) Drought Sensitivity of the Amazon Rainforest. *Science,* **323**, 1344-1347.

Poulter B, Hattermann F, Hawkins ED *et al.* (2010) Robust dynamics of Amazon dieback to climate change with perturbed ecosystem model parameters. *Global Change Biology,* **16**, 2476-2495.

Pueyo S, De Alencastro Graça PML, Barbosa RI, Cots R, Cardona E, Fearnside PM (2010) Testing for criticality in ecosystem dynamics: the case of Amazonian rainforest and savanna fire. *Ecology Letters,* **13**, 793-802.

Rahmstorf S (1996) On the freshwater forcing and transport of the Atlantic thermohaline circulation. *Climate Dynamics,* **12**, 799-811.

Rahmstorf S, Crucifix M, Ganopolski A *et al.* (2005) Thermohaline circulation hysteresis: A model intercomparison. *Geophysical Research Letters,* **32**, L23605.

Rahmstorf S, Willebrand J (1995) The Role of Temperature Feedback in Stabilizing the Thermohaline Circulation. *Journal of Physical Oceanography,* **25**, 787-805.

Rayner D, Hirschi JJM, Kanzow T *et al.* (2011) Monitoring the Atlantic meridional overturning circulation. *Deep Sea Research Part II: Topical Studies in Oceanography,* **58**, 1744-1753.

Rayner NA, Parker DE, Horton EB *et al.* (2003) Global analysis of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research,* **108**, 4407.

Reynolds RW (1978) Sea surface temperature anomalies in the North Pacific Ocean. *Tellus,* **30**, 97-103.

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W (2002) An improved in situ and satellite SST analysis for climate. *Journal of Climate,* **15**, 1609-1625.

Reynolds RW, Smith TM (1994) Improved global sea surface temperature analyses using optimum interpolation. *Journal of Climate,* **7**, 929-948.

Riahi K, Grübler A, Nakicenovic N (2007) Scenarios of long-term socio-economic and environmental development under climate stabilization. *Technological Forecasting and Social Change,* **74**, 887-935.

Rietkerk M, Dekker SC, De Ruiter PC, Van De Koppel J (2004) Self-Organized Patchiness and Catastrophic Shifts in Ecosystems. *Science,* **305**, 1926-1929.

Rudnick DL, Davis RE (2003) Red noise and regime shifts. *Deep Sea Research Part I: Oceanographic Research Papers,* **50**, 691-699.

Salati E, Vose PB (1984) Amazon Basin: A System in Equilibrium. *Science,* **225**, 129-138.

Scheffer M, Bacompte J, Brock WA *et al.* (2009) Early warning signals for critical transitions. *Nature,* **461**, 53-59.

Scheffer M, Carpenter S, Foley JA, Folke C, Walker B (2001) Catastrophic shifts in ecosystems. *Nature,* **413**, 591-596.

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Ziese M, Rudolf B (2014) GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology,* **115**, 15-40.

Shiogama H, Emori S, Hanasaki N, Abe M, Masutomi Y, Takahashi K, Nozawa T (2011) Observational constraints indicate risk of drying in the Amazon basin. *Nat Commun,* **2**, 253.

Sime LC, Stevens DP, Heywood KJ, Oliver KIC (2006) A Decomposition of the Atlantic Meridional Overturning. *Journal of Physical Oceanography,* **36**, 2253-2270.

Sitch S, Huntingford C, Gedney N *et al.* (2008) Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). *Global Change Biology,* **14**, 2015-2039.

Smeed DA, Mccarthy GD, Cunningham SA *et al.* (2014) Observed decline of the Atlantic meridional overturning circulation 2004 to 2012. *Ocean Sci.,* **10**, 29-38.

Smith RS, Gregory JM, Osprey A (2008a) A description of the FAMOUS (version XDBUA) climate model and control run. *Geosci. Model Dev.,* **1**, 53-68.

Smith TM, Reynolds RW, Peterson TC, Lawrimore J (2008b) Improvements to NOAA's Historical Merged Land-Ocean Surface Temperature Analysis (1880-2006). *Journal of Climate,* **21**, 2283-2296.

Spracklen DV, Arnold SR, Taylor CM (2012) Observations of increased tropical rainfall preceded by air passage over forests. *Nature,* **489**, 282-285.

Staver AC, Archibald S, Levin SA (2011) The Global Extent and Determinants of Savanna and Forest as Alternative Biome States. *Science,* **334**, 230-232.

Steele JH, Henderson EW, Mangel M, Clark C (1994) Coupling between Physical and Biological Scales [and Discussion]. *Philosophical Transactions: Biological Sciences,* **343**, 5-9.

Stommel H (1961) Thermohaline Convection with Two Stable Regimes of Flow. *Tellus,* **13**, 224-230.

Sydeman WJ, Santora JA, Thompson SA, Marinovic B, Lorenzo ED (2013) Increasing variance in North Pacific climate relates to unprecedented ecosystem variability off California. *Global Change Biology,* **19**, 1662-1675.

Takimoto G (2009) Early warning signals of demographic regime shifts in invading populations. *Population Ecology,* **51**, 419-426.

Thomas ZA, Kwasniok F, Boulton CA, Cox PM, Jones RT, Lenton TM, Turney CSM (2015) Early warnings and missed alarms for abrupt monsoon transitions. *Clim. Past Discuss.,* **11**, 1313-1341.

Thompson JMT, Sieber J (2011) Climate tipping as a noisy bifurcation: a predictive technique. *IMA Journal of Applied Mathematics,* **76**, 27-46.

Timmermann A, Goosse H (2004) Is the wind stress forcing essential for the meridional overturning circulation? *Geophysical Research Letters,* **31**, L04303.

Timmermann A, Latif M, Voss R, Grötzner A (1998) Northern Hemispheric Interdecadal Variability: A Coupled Air–Sea Mode. *Journal of Climate,* **11**, 1906-1931.

Valdes P (2011) Built for stability. *Nature Geosci,* **4**, 414-416.

Van Der Mheen M, Dijkstra HA, Gozolchiani A, Den Toom M, Feng Q, Kurths J, Hernandez–Garcia E (2013) Interaction network based early warning indicators for the Atlantic MOC collapse. *Geophysical Research Letters,* **40**, 2714-2719.

Van Nes Egbert h, Scheffer M (2007) Slow Recovery from Perturbations as a Generic Indicator of a Nearby Catastrophic Shift. *The American Naturalist,* **169**, 738-747.

Van Vuuren DP, Eickhout B, Lucas PL, Den Elzen MGJ (2006) Long-Term Multi-Gas Scenarios to Stabilise Radiative Forcing - Exploring Costs and Benefits Within an Integrated Assessment Framework. *The Energy Journal,* **27**, 201-233.

Van Vuuren DP, Elzen MJ, Lucas P *et al.* (2007) Stabilizing greenhouse gas concentrations at low levels: an assessment of reduction strategies and costs. *Climatic Change,* **81**, 119-159.

Vecchi GA, Soden BJ, Wittenberg AT, Held IM, Leetmaa A, Harrison MJ (2006) Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. *Nature,* **441**, 73-76.

Vellinga M, Wood R (2002) Global Climatic Impacts of a Collapse of the Atlantic Thermohaline Circulation. *Climatic Change,* **54**, 251-267.

Vellinga M, Wood R (2008) Impacts of thermohaline circulation shutdown in the twenty-first century. *Climatic Change,* **91**, 43-63.

Vellinga M, Wu P (2004) Low-Latitude Freshwater Influence on Centennial Variability of the Atlantic Thermohaline Circulation. *Journal of Climate,* **17**, 4498-4511.

Veraart AJ, Faassen EJ, Dakos V, Van Nes EH, Lurling M, Scheffer M (2012) Recovery rates reflect distance to a tipping point in a living system. *Nature,* **481**, 357-359.

Wanamaker AD, Butler PG, Scourse JD, Heinemeier J, Eiríksson J, Knudsen KL, Richardson CA (2012) Surface changes in the North Atlantic meridional overturning circulation during the last millennium. *Nat Commun,* **3**, 899.

Wang C, Zhang L (2013) Multidecadal Ocean Temperature and Salinity Variability in the Tropical North Atlantic: Linking with the AMO, AMOC, and Subtropical Cell. *Journal of Climate,* **26**, 6137-6162.

White A, Cannell MGR, Friend AD (1999) Climate change impacts on ecosystems and the terrestrial carbon sink: a new assessment. *Global Environmental Change,* **9, Supplement 1**, S21-S30.

Williamson MS, Lenton TM (2015) Detection of bifurcations in noisy coupled systems from multiple time series. *Chaos,* **25**, 036407.

Wissel C (1984) A Universal Law of the Characteristic Return Time near Thresholds. *Oecologia,* **65**, 101-107.

Zhang R (2010) Latitudinal dependence of Atlantic meridional overturning circulation (AMOC) variations. *Geophysical Research Letters,* **37**, L16703.

Zhang R, Delworth TL (2007) Impact of the Atlantic Multidecadal Oscillation on North Pacific climate variability. *Geophysical Research Letters,* **34**, n/a-n/a.

Zickfeld K, Levermann A, Morgan MG, Kuhlbrodt T, Rahmstorf S, Keith D (2007) Expert judgements on the response of the Atlantic meridional overturning circulation to climate change. *Climatic Change,* **82**, 235-265.