Leading indicators in B- and R-tipping systems with focus on eigenvalue estimation

MASTER THESIS
submitted in fulfillment of the requirements of the degree
Master of Science

Submitted by
Martin Heßler
born in Bad Neuenahr-Ahrweiler

Research group of Prof. Dr. U. Thiele
Institute for theoretical physics
Physics department
2018

1st examiner : Dr. Oliver Kamps

2nd examiner : PD Dr. Svetlana Gurevich

Date of submission: 26th September 2018

Abstract

This Master's thesis deals with the numerical anticipation of critical transitions in complex dynamical systems from time series data. The necessary theory of B-, R- and N-tipping systems as well as the theoretical formalism for the description of complex dynamical systems and their stability is provided. Furthermore, basic tools of time series analysis as delay embedding, S-maps and autoregression schemes AR(p) are introduced and the theory of commonly used statistic leading indicators as well as a theoretical derivation to estimate the dominant eigenvalue of the Jacobian as a leading indicator is given. The phenomena of critical slowing down and flickering prior to a bifurcation are described and the numerical algorithms to estimate the statistical leading indicators and the dominant eigenvalue are presented. A detailed comparison of the predictive power, the robustness and the applicability of the various methods is performed on five synthetic B-tipping datasets. The dominant eigenvalue estimation with autoregression schemes is performed on empirical datasets of four different real world systems. It is shown that the estimation of the dominant eigenvalue enables to classify bifurcation types in the case of B-tipping systems and finally, the leading indicators' capacity to anticipate R-tipping points is falsified. The leading indicators fail to predict R-tipping in congruence with the theoretical prognosis.

Table of contents

List of abbreviations

adfd average distance from diagonal

AR(p) autoregression scheme of the order p

B-tipping bifurcation induced tipping

BV dataset Bonneville dataset

CD dataset Christopher M. Danforth dataset

CSD critical slowing down

DEV dominant eigenvalue

fnn false nearest neighbour

N-tipping noise induced tipping

R-tipping rate dependent tipping

ROC receiver operating characteristic

S-map sequential map

1 Introduction

Nonlinear phenomena are a fundamental part of nature. Their nonlinear features often cause real world systems to exhibit counterintuitive and rather complex behaviour. This can be notably dangerous, because of drastic changes in the system's solution space. In consequence, it is of great interest to understand and preferably predict such drastic system changes, called *critical transitions* or *tipping points* of *complex dynamical systems*.
Critical transitions can be excited due to various mechanisms. Recent publications distinguish between three generic types, the

1. *bifurcation induced* tipping, abbreviated B-tipping,
2. *rate dependent* tipping, abbreviated R-tipping,
3. and *noise induced* tipping, abbreviated N-tipping.

A bifurcation is defined as a qualitative change in the solution space of a complex system due to a parameter variation exceeding a critical threshold. Above the threshold previously stable states can become unstable or disappear and oscillations can rise up in a variety of possible bifurcation types. If a system undergoes a critical transition without a parameter exceeding the critical threshold it is called R- or N-tipping. N-tipping is generically unpredictable, because it pushes the system out of its equilibrium state towards an unstable one due to a large and improbable random noise event. More interesting concerning the question of predictability is the R-tipping mechanism: If a system's control parameter is changed rapidly, the system cannot follow the stable state in a quasi-static manner. The system relaxes too slow to follow the shifting stable state and is passed by an unstable state in state space. This leads to a critical transition that is caused by the rate of change of the parameter and is therefore called R-tipping. B-tipping is observed in many real world systems, but there is some evidence that also R- and N-tipping can be found in nature. One indication is given by a climate model that exhibits all types of tipping independently for a suitable parameter choice. [Ash+12; RS16; RS17]
The basic idea in order to predict critical transitions before they occur is trying to extract mathematical measures, so called *early warning signals* or *leading indicators*, directly from time series data of the system of interest. Commonly used measures are e.g. the autocorrelation and variance. They are based on *critical slowing down*, abbreviated CSD, the phenomenon that tends to occur before the system undergoes a B-tipping critical transition. CSD means a decreasing relaxation rate towards the equilibrium state after a perturbation if the system approaches the B-tipping point. This leads to a higher similarity of equally sampled past events, thus to a higher autocorrelation, and also to an increased variance over time because of cumulative effects of noise events.
Some systems fluctuate between two stable solutions before one of them becomes finally unstable at the tipping point. This is called *flickering* and leads to another possible leading

indicator.
The third central statistical moment, called *skewness* is a measure for the asymmetry of a given probability distribution. Clearly, the flickering process causes an asymmetry in the probability distribution of the two competing states. Apart from these often used leading indicators there is plenty of other approaches facing these predictive issues. Other approaches are for example drift-diffusion-jump models [CB11], which yield further inside in its critical dynamics, or determining the underlying potential [Liv+13; LKL10]. Another ansatz reconstructs the bifurcation diagram of a given system. The method is described in [LE11]. A good overview of the phenomena mentioned above and further approaches can be found in [Dak+12].
Complex dynamical systems play an important roll in almost every part of nature, society and life. For this reason investigations concerning the leading indicators proceed in a wide range of scientific communities and are definitely not limited on physics. The climate system is considered as a complex system. Climate changes are interpreted as critical transitions and leading indicator methods and Bayesian statistics are used to investigate past climate catastrophes, e.g. from geochemical data or atmospheric gas concentrations [Sch+09; DM04]. Ecological systems like the blossom onset of various plants in correspondence to weather phenomena, aquatic systems e.g. in a lake system and the oxygen concentration around the globe are investigated in [DM06; Gse+16; LMF15].
There is also recent work on predicting earthquakes. A changing probability distribution of the vertical seismic velocity or an increasing self similarity of the seismic data could be a marker for an imminent earthquake [Man+09; Tab+]. Also power systems and imminent power outages have direct impact to the life of thousands of people and are barely understood because of a high degree of the system's complexity [RW15].
In medicine it is tried to use the leading indicator approaches and nonlinear description of the organism for diagnostic purposes. For example the danger of an acute allograft rejection after a cardiothoracic transplantation could be anticipated in a non-invasive way by analysing the complexity of the heart rate variability. CSD could occur prior to attacks of chronical diseases as epilepsy, narcolepsy and other diseases [Izr+00; Rik+16]. Even the psychological states of consciousness are considered as complex system states. In this view, the onset and termination of depression are critical transitions and thus, they are tried to predict using common leading indicators [Lee+13].
Other areas in society could also take advantage of the complex system theory. In 2014 T. Chadefaux has presented a newspaper archive that can be used to anticipate outbreaks of wars and civil wars. In economy predicting critical transitions such as financial crisis is of great interest [Cha14; Jur+17].
It makes no difference in which area the analyses are performed, there are always the same difficulties. Dealing with simulated time series or even real world systems is challenging because of the rarely accessible information about the system. In real world applications the whole set of variables and parameters that influence the system are often unknown and just a few of them are measured and adequately sampled in a time series that could be used for an analysis. A reliable model or any fixed differential equations are missing for the systems of interest. For these reasons it is necessary to make some theoretical assumptions to deal with the topic of anticipating critical transitions at all. For example the time evolution of the leading indicators will be calculated using *rolling time windows*. In each window the system

that produces the time series data is assumed to be *stationary*. This means that the system state is fixed due to fixed external and internal parameters. A more precise definition will be given later. The rolling window approach allows to estimate the time evolution of the leading indicators, but also limits the applicability of the methods themselves: To estimate one value of the leading indicator of interest, such an amount of data is necessary that at least one rolling time window gives stable numerical results. Using the numerical approaches one is confronted with lots of other difficulties concerning the choice of a suitable window size and appropriate parameters for the numerical methods.

In the course of this Master's thesis, the necessary theory of complex systems and statistics in oder to deal with the numerical methods and analyses of the investigative part is provided in chapter 2. In subsection 2.3.1 of this chapter it is also derived the needed theoretical formalism for the development of a numerical method to estimate the *dominant eigenvalue*, abbreviated DEV.

Three numerical methods are developed to estimate the DEV from time series data with *autoregression* schemes and two variants of the *S-map algorithm*. They are described in chapter 3. The dominant eigenvalue can be seen as a leading indicator, comparable to the *recovery rate* described in [Dak+12]. The estimation of the DEV also allows to determine specific types of bifurcations. Therefore, in chapter 4 the DEV methods are applied to several simulated B-tipping test datasets with different types of bifurcations to investigate their predictive power and their ability to classify different types of bifurcations from time series data. The numerical *false nearest neighbour algorithm* and the *average distance from diagonal algorithm* are used to determine optimal numerical parameters as *embedding dimension* and *time lag*. It is validated that these algorithms provide suitable numerical parameters. The robustness against suboptimal numerical parameter choices and various rolling time window sizes is investigated using *receiver operating characteristic curves*, shortly ROC curves. Via ROC curves, the DEV methods are compared in detail to commonly used statistical leading indicators of autocorrelation, variance and skewness concerning their predictive power and their robustness, before the three DEV methods are compared one to another.

After an in-depth methodological research, the DEV method with autoregression is applied to four real world problems. They are chosen from different fields as ecology, physiology, power systems and climate. They undergo different types of bifurcations. In consequence, the DEV method's functionality is investigated for these specific real world examples.

In the second investigative part of chapter 5 the applicability of the previously considered leading indicators to two R-tipping systems is tested. The autocorrelation and the standard deviation are used as leading indicators for R-tipping systems in [RS16]. The predictive capacity of the measures has to be controlled in these cases, because of the theoretical absence of CSD prior to R-tipping points. Additionally, the dominant eigenvalue as a leading indicator is determined via an estimation of the *flow*, whereas the corresponding numerical method to estimate the *flow* is developed in chapter 3. The flow determines the rate of change for a variable over time. Finally, the results are summarized and an outlook is given in chapter 6.

2 Theory

A system is called a *complex dynamical system* if its behaviour obeys a nonlinear evolution. Typically, such a system is an open system that contains lots of different interacting parts. Their interaction cannot be modelled by superposition due to the nonlinearity. If such a system undergoes a drastic change in its behaviour, it is called a *critical transition*. [Str15; Sch+12]
This chapter deals with the basic mathematical formalism of complex systems and critical transitions. Furthermore, the statistical properties that can serve as early warning measures for critical transitions are summarized. Together with an introduction to important time series analysis tools at the end of this chapter the following sections provide a mathematical framework which is necessary to anticipate critical transitions.

2.1 Mathematical formalism of complex systems

In order to understand the mathematical nature of the stability of complex system states and possibly occurring bifurcations the two mathematical main types of dynamical systems are presented: With *differential equations* it is possible to model particular system dynamics continuous in time. A model, discrete in time, can be realized as iteration scheme that is called an *iterated map*. [Str15]

2.1.1 Iterated maps

As mentioned above, the main feature of an iterated map is the discrete time evolution modelling. The general form of an N dimensional iterated map can thereby be written as

$$\begin{aligned} x_{1,n+1} &= f_1(x_{1,n}, x_{2,n}, ..., x_{N,n}) \\ x_{2,n+1} &= f_2(x_{1,n}, x_{2,n}, ..., x_{N,n}) \\ &\vdots \\ x_{N,n+1} &= f_N(x_{1,n}, x_{2,n}, ..., x_{N,n}). \end{aligned} \tag{2.1}$$

This way of representing the underlying dynamics of a complex system is in particular a practical and a fast way for computer calculations, because of its inherent discretization. For some systems in the fields of digital electronics, economics, finance theory, certain impulsively driven mechanical systems and animal population without overlapping generations a time discrete model is even natural. Furthermore, they also serve as simple mathematical models for chaos and the investigation of dynamical properties of a system, e.g. the existence of a closed orbit via a so called *Poincaré map*.[Str15]

2.1.2 Ordinary differential equations

Representing the system dynamics continuously in time leads to a system of coupled ordinary differential equations

$$\begin{aligned} \dot{x}_1 &= f_1(x_{1,n}, x_{2,n}, ..., x_{N,n}) \\ \dot{x}_2 &= f_2(x_{1,n}, x_{2,n}, ..., x_{N,n}) \\ &\vdots \\ \dot{x}_N &= f_N(x_{1,n}, x_{2,n}, ..., x_{N,n}). \end{aligned} \tag{2.2}$$

The equations deal with the first derivatives $\dot{x}_i$ with $i = 1, 2, ..., N$, thus, the infinitesimal smooth evolution of the observables x_i in time. At this point it is convenient to introduce the famous *Lorenz system*

$$\begin{aligned} \dot{x} &= a(y - x) \\ \dot{y} &= x(b - z) - y \\ \dot{z} &= xy - cz \end{aligned} \tag{2.3}$$

as an example. This three dimensional first order ordinary differential equation system is originally a rough approximation of modelling convection phenomena in the atmosphere. This deterministic system exhibits very complex behaviour for various parameter sets a, b and c. Figure 2.1 illustrates this fact. [Str15]

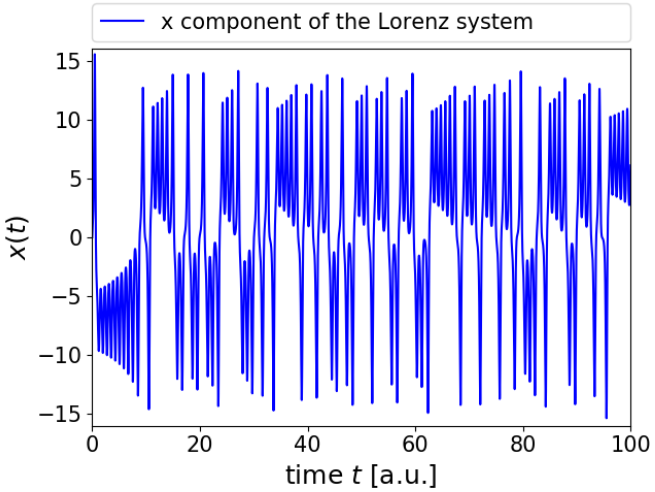


Figure 2.1: The x component of the Lorenz system seems to jump in an irregular manner between around $x = -15$ and $x = 15$ after several oscillations with increasing amplitudes prior to each jump.

2.1.3 The concept of state space

In general a single variable provides very little insight into a given system of interest. Instead of determining a system's features by the time evolution of some single observable x, as in

Figure 2.1, every system state is uniquely and completely defined by the combination of all variables influencing the system at time t. Thus, for describing the possible states of a complex system, the representation in a so called *state* or *phase space* is very useful. The state space is spanned by all variables determining the system. Remarkably, the variables follow their original and non-approximated nonlinear dynamics. Furthermore, every point in the state space could serve as initial condition. The time evolution in this space is given by *trajectories* that are continuous lines of adjacent states in time. Lots of qualitative information about the dynamics of a system, e.g. unstable and stable states or chaotic evolution, can be visualized and extracted from this representation. In the example of the Lorenz system the variables spanning the state space would be x, y and z. In Figure 2.2 the famous butterfly attractor is shown. An attractor in phase space is a minimal invariant set of points that attracts an open set of initial conditions. If this invariant set is moreover sensitive in its response to initial conditions, the attractor is called a *strange* or *chaotic attractor*. The attractor shown in Figure 2.2 is such a strange attractor. Originally, this name comes from the geometrical properties of this set of points. It has a fractal dimension between 2 and 3. In state space the largest set of points that is attracted by a fixed point is called *basin of attraction*. The basins of attraction of two competing fixed points are divided by the so called *separatrix*, a line in two dimensions or a hyperplane in higher dimensions of phase space. [Str15]

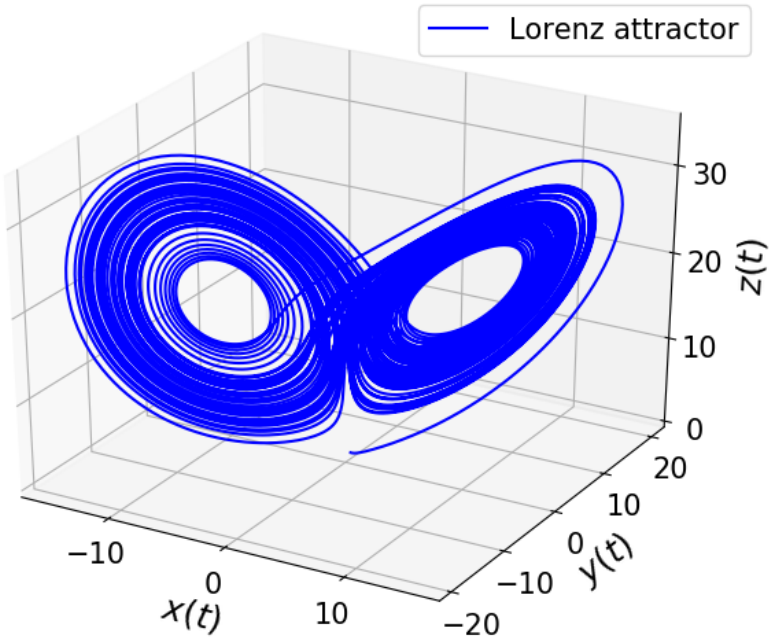


Figure 2.2: The Lorenz attractor is visualized in state space which is spanned by x,y and z of Equation 2.3. The graphic is drawn with the same parameters as Figure 2.1, but this representation visualizes the complete dynamical structure in state space.

2.2 Mechanisms of critical transitions

In this section three possible mechanisms are presented that can change the system's behaviour drastically by triggering a *critical transition* or *tipping point*.

2.2.1 Bifurcation induced tipping

One possible mechanism for a critical transition is an external system parameter that passes a certain critical threshold. A complex system exhibits a particular solution space for given system correlated parameters. The corresponding solution space changes just slightly if the control parameter does not exceed a certain threshold. If a given parameter passes this threshold, the solution space changes remarkably due to a so called bifurcation, where previously stable solutions develop unstable, disappear or vice versa. [Ash+12; RS16; RS17]
Figure 2.3 gives an example of the bifurcation induced tipping, also called B-tipping. The green landscape is the potential, i.e. $\dot{x} = -\frac{\mathrm{d}V}{\mathrm{d}x}$, of a given dynamical system. The system states are described by the position of the blue ball. It can be perturbed because of some naturally occurring noise, but remains always in the stable valley. If a control parameter of the complex system changes over time towards a critical value, the landscape begins to deform. In the middle case the valley is still there, the state is stable, but the restoring force of the potential is much less than before. The ball could be easily driven away from the state's basin of attraction. If the control parameter exceeds the critical value, a possible scenario is the exchange of stability: The valley becomes an unstable mountain, whereas the previously unstable state develops a stable valley. Obviously, in this analogon the ball is driven into the stable minimum.

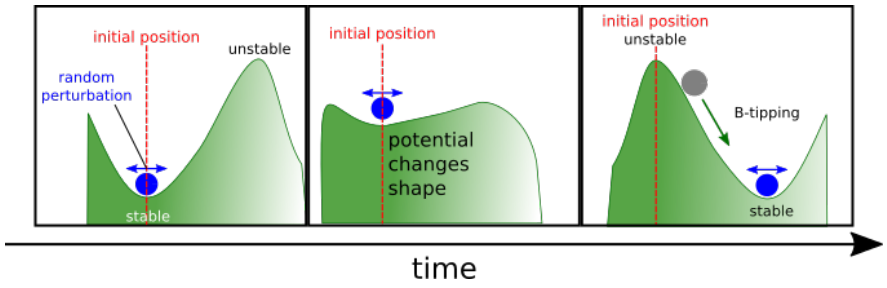


Figure 2.3: The mechanism of bifurcation induced tipping is visualized with a potential landscape, shown in green. If a control parameter changes in time across some critical value, the potential landscape is qualitatively deformed. The previously stable valley becomes an unstable mountain and the mountain becomes a valley. The change in the system's behaviour is shown by the blue ball, which corresponds to the current system state in this metaphor.

2.2.2 Rate induced tipping

A second mechanism of a critical transition is the so called rate induced tipping, also called R-tipping. In this case, the control parameter does not exceed a critical value and no bifurcation takes place. Instead of the B-tipping mechanism, the critical transition is stimulated due to a very rapid parameter change, more precisely, it depends on the rate of change of the control

parameter. In the R-tipping case the system is not able to follow the changing equilibrium state in a quasi-static way, before its current state develops unstable in consequence of the fast parameter change. [Ash+12; RS16; RS17]
Again it can be explained in the mechanical analogon, shown in Figure 2.4: a green potential landscape, i.e. $\dot{x} = -\frac{\mathrm{d}V}{\mathrm{d}x}$, with valleys and mountains, corresponding to stable and unstable states, and the system state marked by the blue ball. The blue arrow emphasizes that the ball is at rest. The R-tipping process could be imagined as if the potential landscape is pulled away under the blue ball due to a rapid parameter change. The current position of the stable valley and unstable mountain is widely shifted in a short time. The ball begins to roll down into the valley as indicated by the blue arrow in the middle picture, but relaxes much slower than the potential landscape shifts. Before the blue ball can change significantly its previous equilibrium position in order to follow the new equilibrium valley the potential landscape reaches the shape depicted in the right picture. The ball rolls down the mountain, the system state becomes unstable. This would not be the case if the blue arrow would be greater than the movement of the potential landscape, indicated by the green arrow in the middle picture. Then the ball would follow the moving valley. Therefore, a sufficiently great rate of change of the system's control parameter depending of the potential's curvature is necessary to observe R-tipping.

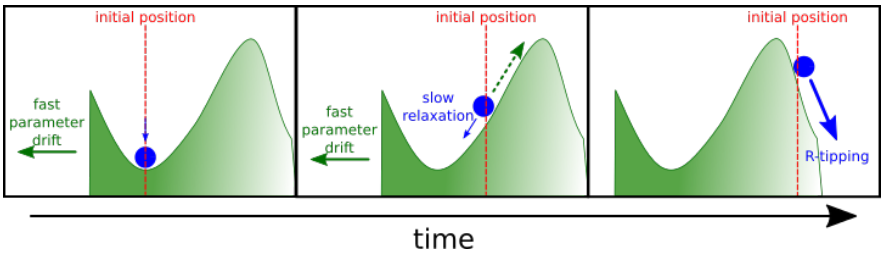


Figure 2.4: The mechanism of rate induced tipping is visualized with a potential landscape, shown in green. If a control parameter changes rapidly in time without reaching a critical value for B-tipping, the potential landscape could be imagined as rapidly shifted under the blue ball that indicates the current system state. In the central picture, the blue ball is more or less at its initial position due to the great rate of change of the control parameter. The ball begins to relax into the shifted stable valley, but the potential is shifted so fast, that the ball cannot roll down into the valley. At the R-tipping point the blue ball reaches the top of the unstable mountain.

2.2.3 Noise induced tipping

A third mechanism of a critical transition is the noise induced tipping or N-tipping. In the case of pure N-tipping there is no changing control parameter at all. The system is driven out of equilibrium into an unstable state just because of noise contributions. It could be possible that the intensity of these contributions change over time, but it is not a priori necessary to observe N-tipping. The only necessary condition is that there exists a small probability for a noise perturbation that pushes the system towards an unstable state. [Ash+12; RS16; RS17]
In Figure 2.5 a case of pure N-tipping is illustrated. The current system state is marked by the blue ball in a green potential landscape, i.e. $\dot{x} = -\frac{dV}{dx}$. The system state is exposed to noise perturbation processes, indicated by the blue double arrow. A probability distribution, shown above in transparent blue, visualizes the possible positions of the ball due to some noise perturbation. At the left part, the ball would always roll back into the stable valley, indicating a relaxation towards the stable state in the mechanical analogon. But at the right side exists a small probability, signed in red, to be pushed over the separatrix into the unstable state - respectively the unstable mountain top, that is indicated by the red dashed line. In this scenario it is possible to undergo a critical transition without parameter changes, neither rapid parameter drifts nor drifts passing any critical threshold.

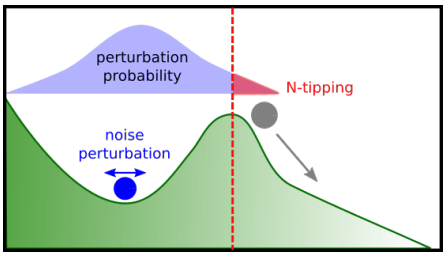


Figure 2.5: It is visualized the mechanism of noise induced tipping with a potential landscape, shown in green. The blue ball indicates the current system state. No control parameter has to be changed to observe N-tipping. As indicated by the transparent blue probability distribution of the possible positions of the blue ball due to noise contributions, the ball can be driven up and behind the unstable mountain top as marked by the red colour. The system is driven unstable just by a random noise effect.

2.3 Mathematical formalism of B-tipping

Bifurcations [Str15] are defined as qualitative changes in the solution space of a system. It may happen, that a previously stable state becomes unstable due to a changing parameter set. In consequence, the system changes its behaviour until it settles down to another stable state or *fixed point* in phase space. Bifurcations are one possible origin of *critical transitions* [Dak+12]. These transitions can exhibit remarkable and sudden changes in a system's behaviour. Often this is not wanted in practical applications and therefore, methods to anticipate these transitions are developed in order to avoid remarkable system modifications.

2.3.1 Linear stability of system states

The changing stability conditions for a fixed point are investigated for small perturbations around the fixed point and here treated in a linear manner.
Without loss of generality the theory of linear stability of iterated maps is presented in two

dimensions. It could be analogously expanded into higher dimensions. The linear stability of a fixed point (x^*, y^*) of a two dimensional iterated map

$$\begin{aligned} x_{n+1} &= f(x_n, y_n) \\ y_{n+1} &= g(x_n, y_n) \end{aligned} \tag{2.4}$$

is considered by applying a small perturbation $\left|\eta_n^i\right| \ll 1$ with $i \in \{x, y\}$ to the fixed point (x^*, y^*). Regarding the x component one can expand $f(x_n, y_n)$ in a Taylor series around the fixed point up to the linear order in x and y resulting in

$$x^* + \eta_{n+1}^x = f(x_n + \eta_n^x, y_n + \eta_n^y) \tag{2.5}$$

$$\approx \underbrace{f(x^*, y^*)}_{x^*} + \frac{\partial f}{\partial x}(x^*, y^*) \cdot \eta_n^x + \frac{\partial f}{\partial y}(x^*, y^*) \cdot \eta_n^y + O\left((\eta_n^x)^2, \left(\eta_n^y\right)^2, \eta_n^x \eta_n^y\right). \tag{2.6}$$

Subtracting x^* in Equation 2.6 and performing these steps analogously for the y component leads to

$$\vec{\eta}_{n+1} = \begin{pmatrix} \eta_{n+1}^x \\ \eta_{n+1}^y \end{pmatrix} = \left.\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix}\right|_{(x^*, y^*)} \cdot \begin{pmatrix} \eta_{n+1}^x \\ \eta_{n+1}^y \end{pmatrix} = J|_{(x^*, y^*)} \cdot \vec{\eta}_n. \tag{2.7}$$

The eigenvalues λ_i of the so called Jacobian matrix J define the development of the perturbation in the next time step. The Equation 2.7 can be written as

$$\vec{\eta}' = \underline{\underline{\Lambda}} \cdot \vec{\eta}' \tag{2.8}$$

with the diagonalized Jacobian $\underline{\underline{J}} = \underline{\underline{E \Lambda E^{-1}}}$ with the diagonal matrix $\underline{\underline{\Lambda}}$ of eigenvalues and the matrix $\underline{\underline{E}}$ of right eigenvectors and the scaled perturbations $\vec{\eta}' = \underline{\underline{E}}^{-1} \cdot \vec{\eta}$. [Str15]
Eigenvalues with $|\lambda_i| < 1$ cause the perturbation to decay, whereas the perturbation increases for eigenvalues with $|\lambda_i| > 1$. An analogous derivation for differential equations leads to

$$\dot{\eta} \approx \lambda \cdot \eta \tag{2.9}$$

and the criterion of the real part $\Re(\lambda) < 0$ for a stable state and $\Re(\lambda) > 0$ for an unstable state. [Str15]
A given time series with N randomly fluctuating points around a fixed point due to noise could be given by (x_n, y_n) with $n = 1, ..., N$. The permanent random noise could be seen as a perturbation $\vec{\eta}_n$ around a fixed point x^* resulting in $\eta_n^x = x_n - x^*$ and $\eta_n^y = y_n - y^*$. These equations applied to Equation 2.7 add an offset to the equation

$$\vec{x}_{n+1} - \vec{x}^* = J|_{(x^*, y^*)} \cdot (\vec{x}_n - \vec{x}^*) \tag{2.10}$$

$$\Leftrightarrow \vec{x}_{n+1} = J|_{(x^*, y^*)} \cdot \vec{x}_n \underbrace{- J|_{(x^*, y^*)} \vec{x}^* + \vec{x}^*}_{\vec{x}_{offset}} \tag{2.11}$$

$$\Leftrightarrow \vec{x}_{n+1} = J|_{(x^*, y^*)} \cdot \vec{x}_n + \vec{x}_{offset}. \tag{2.12}$$

Based on this form the eigenvalues and, in consequence, information about the linear stability of the investigated system can be gained, given that an approach is available that enables to estimate the coefficients of the Jacobian J from time series data.

2.3.2 Types of local bifurcations

Local bifurcations, different from *global bifurcations* that occur due to colliding structures in phase space, go along with the change of the eigenvalues of a Jacobian, derived from a local linear approximation around a fixed point [Str15]. The derivation can be found in subsection 2.3.1. The transition can be classified by the change in the dominant eigenvalues, abbreviated DEV, of the Jacobian as summarized in Table 2.1. Without any claim to completeness, the *tangent*, the *period doubling* and the *Neimark-Sacker bifurcation* [XCM10; KS08], the *Hopf bifurcation* for iterated maps, are considered for iterated maps. If a discrete system undergoes a tangent bifurcation, a stable and an unstable fixed point collide and annihilate or a fixed point appears out of nowhere. This is accompanied by the purely real eigenvalue λ reaching $|\lambda| = 1$. [Str15]
A period doubling bifurcation is a typical bifurcation for iterated maps. As its name implies, a regularly repeating series of points becomes doubled in its period if the bifurcation occurs. A cascade of such period doubling bifurcations can serve as a way to chaos, as described in detail for the logistic map in [Str15; Ger10]. The period doubling bifurcation is indicated by a pure real eigenvalue λ reaching $|\lambda| = -1$. [Str15]
If a system undergoes a Neimark-Sacker bifurcation, a closed invariant curve is born from a fixed point. This happens if a pair of complex conjugated eigenvalues $\lambda_{1,2} = r(\alpha)\exp(\pm i\theta(\alpha))$ reaches $|\lambda_{1,2}| = 1$. [XCM10; KS08]

Table 2.1: Local bifurcation types

Bifurcation type	System change	Eigenvalue λ trend
Tangent bifurcation	Rising fixed point or annihilation of stable and unstable fixed point.	$\Im(\lambda) = 0$, $\lvert\lambda\rvert \longrightarrow 1$
Period doubling bifurcation	Previous period is doubled.	$\Im(\lambda) = 0$, $\lvert\lambda\rvert \longrightarrow -1$
Neimark-Sacker bifurcation	An invariant closed curve rises from a fixed point.	$\Im(\lambda) \neq 0$, $\Im(\lambda_1) = -\Im(\lambda_2)$, $\lvert\lambda_{1,2}\rvert \longrightarrow 1$

2.4 Stochastic properties

For the investigation of complex systems it is necessary to define some basic concepts in mathematical statistics. The term *noise* [KLZ14] refers to the presence of stochastic components in a system or at least uncertainties in the measurement procedure. To state that each measured signal from a natural system contains noise is non-exaggerated. For the numerical analysis noise is even a fundamental feature that forces a system to relax from its noise perturbed state into its equilibrium state.

2.4.1 Probabilities

If one performs a random experiment with the set of events **A**, a certain result A will be observed n_A times for n experimental realisations with the so-called *frequency* n_A. The quotient

$$h_A = \frac{n_A}{n} \tag{2.13}$$

is called *relative frequency*.

Marginal probability $P(A)$ is called a function that fulfils

1.

$$0 \leq P(A) \leq 1, \;\; 0 \leq h_A \leq 1 \tag{2.14}$$

2.

$$P(O) = 0, \;\; P(I) = 1, \;\; h_O = 0, \;\; h_I = 1 \tag{2.15}$$

for the impossible event O and the safe event I,

3.

$$P(A \vee B) = P(A) + P(B), \;\; h_{AB} = h_A + h_B \text{ with } B \in \mathbf{A} \tag{2.16}$$

if the events A and B exclude one another with $A \wedge B = O$.

A probability defined this way yields a measure for the expected frequency of a certain event if an experiment is performed many times. For infinite experiments $n \longrightarrow \infty$ an event A with $P(A) = 0.5$ should be observed every second time in average. The real probability is normally not known and an estimation is needed. The relative frequency of an event A is commonly used to estimate the unknown real probability in an empirical way and to establish an axiomatic definition of the probability.[Bro+08; Koh05]
Often, two events are not independent of each other. Thus, if A and B are not independent of each other, once knowing that event A has occurred changes the probability to observe B. This is written in mathematical terms with the *conditional probability*

$$P(B|A) = \frac{P(A,B)}{P(A)}, \;\; P(A) \neq 0, \tag{2.17}$$

with the *joint probability* $P(A,B)$ that is the probability to find a certain pair of events (A,B). If two events are independent of each other, it is:

$$P(A|B) = P(A), \;\; P(B|A) = P(B) \text{ and } P(A,B) = P(A)P(B). \tag{2.18}$$

2.4.2 Stochastic processes

A *random variable*, denoted by X, is a variable that attains a specific real value $x_i \in R$, for $i = 1, 2, ..., N$, where R is a real set and N is the amount of events under random conditions. If R is finite or countable infinite, then X is called a *discrete* random variable that follows a distribution function

$$F(x) = \sum_{x_i \leqslant x} p_i \tag{2.19}$$

with the probabilities $P(X = x_i) = p_i$. The distribution function $F(x)$ measures the probability to find X attaining a value between $-\infty$ and x.
A random variable is *continuous* if the set R is infinite. The probability to find $X = x_i$ is zero. Therefore, a probability is defined for a specific interval $[a, b]$ using the probability density function $f(t)$:

$$P(a \leqslant X \leqslant b) = \int_a^b f(t)\mathrm{d}t. \tag{2.20}$$

The distribution function fulfils

$$\lim_{x \to -\infty} F(x) = 0 \wedge \lim_{x \to \infty} F(x) = 1, \tag{2.21}$$

and is a monotonic, right-continuous function which is defined by the integral

$$F(x) = P(X \leqslant x) = \int_{-\infty}^{x} f(t)\mathrm{d}t. \tag{2.22}$$

A set of random variables X that depend on a parameter, usually time t, are called a *stochastic process*. The random variable is then denoted by X_t and the whole stochastic process can be written as

$$\{X_t | t \in T\} \tag{2.23}$$

with the parameter space T. If both the parameter space T and the state space of the random variable X_t are discrete, the process is called a *stochastic chain*. [Bro+08; Koh05]

2.4.3 Statistical moments

In order to characterize a specific distribution of a random variable X in a stochastic process, the so-called *statistical moments* are used. The moment $E(X^n)$ of the order n is defined as

$$E(X^n) = \sum_{i=1}^{N} x_i^n p_i \tag{2.24}$$

$$E(X^n) = \int_{-\infty}^{\infty} x^n f(t)\mathrm{d}t \tag{2.25}$$

for discrete and continuous distributions. The first moment is called *expected value* or *mean* μ_X. The central moments $E((X-\mu_X)^n)$ of the order n are defined similarly, but they are evaluated for the fluctuations of X around the mean μ_X to the power of n. The second central moment

$$E((X-\mu_X)^2) = \begin{cases} \sum_{i=1}^{N}(x_i-\mu_X)^2 p_i \\ \int_{-\infty}^{\infty}(x_i-\mu_X)^2 f(t)\mathrm{d}t \end{cases} \tag{2.26}$$

is called *variance* and is a measure for the deviations of X from the mean μ_X. Its square root is called *standard deviation*. The third central moment

$$E((X-\mu_X)^3) = \begin{cases} \sum_{i=1}^{N}(x_i-\mu_X)^3 p_i \\ \int_{-\infty}^{\infty}(x_i-\mu_X)^3 f(t)\mathrm{d}t \end{cases} \tag{2.27}$$

is called *skewness* and is a measure for the asymmetry of a given probability distribution. [Bro+08; Koh05; Dak+12]

2.4.4 Autocorrelation

Based on the definitions in subsection 2.4.3 the *covariance* of a two dimensional random variable (X, Y) is given as

$$Cov(X,Y) = E((X-\mu_X)(Y-\mu_Y)). \tag{2.28}$$

The covariance is a comparative measure for the linear dependence or *correlation* of the two variables and thus, can attain arbitrary values. In order to gain absolutely comparable information of the correlation of two variables the covariance is normalized via the standard deviations of the two variables yielding the *correlation coefficient* ρ_{XY} as

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}. \tag{2.29}$$

The correlation coefficient is an absolute measure for the correlation of two random variables. The *autocorrelation* can be seen as a special case of the covariance definition and is explicitly written as

$$\rho_{X_t, X_{t-\tau}} = \frac{Cov(X_t, X_{t-\tau})}{\sigma_{X_t}\sigma_{X_{t-\tau}}}, \tag{2.30}$$

wherein X_t denotes the random variable at a certain reference time t and $X_{t-\tau}$ the random variable delayed by some time lag τ. Therefore, the autocorrelation is a measure for the interdependence of adjacent values of one specific random variable X for a given time lag τ. [Bro+08; KS04]

2.4.5 Stationary processes

A process is called *weakly stationary* if the first and the second statistical moment, i.e. the mean μ and the variance σ^2, are constant over time t and therefore the covariance is invariant under

arbitrary time shifts Δt, i.e. $Cov(X_{t_1}, X_{t_2}) = Cov(X_{t_1+\Delta t}, X_{t_2+\Delta t})$. For some linear stochastic processes this definition might be convenient. For nonlinear stochastic problems however, a more specific definition is needed. The nonlinear process of interest has to take place in a fixed parameter environment. In other words, all experimental or external parameters, but also the inherent parameters of the nonlinear process itself have to be constant over time. In a more formal way this means that the joint probabilities to find a system in an arbitrary state a at time t_a and to find it in the arbitrary state b at time t_b are independent of time. This definition of *strong stationarity* includes a permanent constant parameter environment and the condition of the nonlinearity to be sufficiently resolved in the stochastic nonlinear process X_t of interest. [KS04]

2.4.6 Gaussian white noise

A stream x_i of realisations of a random variable X with a flat power spectrum is termed *white noise*. In analogy to optics it carries the term "white", because every noise frequency contributes in equal manner to the noise signal similar to equally contributed light frequencies that create white light. In statistics this means

$$Cov(x_1, x_2) = 0, \tag{2.31}$$

i.e. the adjacent values are independent of each other. The second main feature is that white noise is generated by a time constant probability distribution $F(t)$, thus $F(t_1) = F(t_2) \; \forall \; t_1 \neq t_2$. The actual distribution is not determined by the term white noise, but the Gaussian probability density

$$\mathcal{N}(\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{2.32}$$

with $x \in \mathbb{R}$ the expected value μ and the standard deviation σ is commonly used. A process that obeys a Gaussian distribution is called a *normal distributed* process. The probability density $\mathcal{N}(\mu = 0, \sigma^2 = 1)$ is called the standard Gaussian distribution. [KLZ14; Bro+08; Hor07]

2.5 Phenomena in B-tipping systems

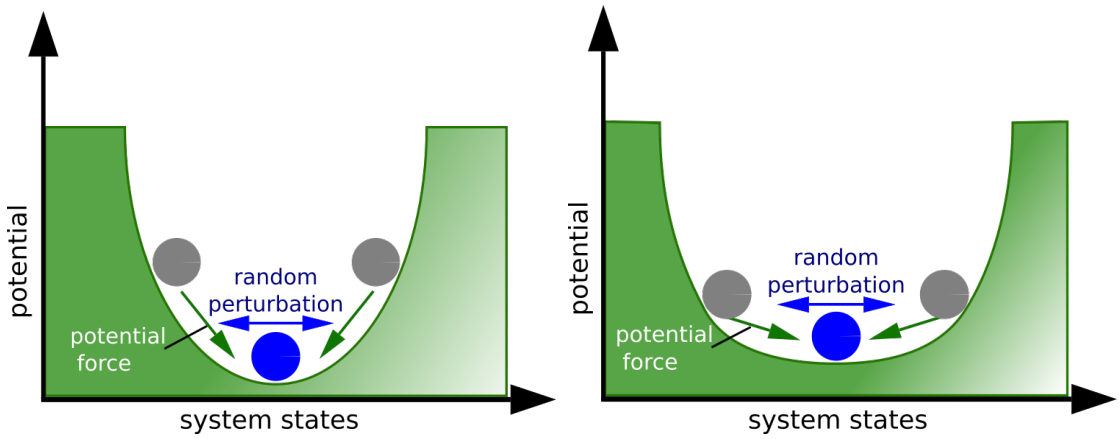
Mainly two phenomena are observed prior to a critical transition in a B-tipping system. These phenomena influence the stochastic properties of a system and thus, affect indirectly the time evolution of various leading indicators. This is discussed in the following subsections.

2.5.1 Critical slowing down

As shown in subsection 2.3.1, a small perturbation $\eta_n^{x'}$ decays more slowly if the corresponding eigenvalue λ approaches unity. That can be easily seen for a single component of Equation 2.8

$$\eta_{n+1}^{x'} = \lambda_x \cdot \eta_n^{x'}, \tag{2.33}$$

where the next perturbation is equal to the previous one scaled with λ_x. This phenomenon is called *critical slowing down*, abbreviated CSD, because of the lower recovery rate to the stable state near a bifurcation point. For $\lambda = 1$ the perturbation is persistent and does not decay or grow any further. Another helpful visualization of CSD is the underlying potential structure of a system in a stable state. The principle is visualized in Figure 2.6. A system in a stable state, marked by the big blue dot in Figure 2.6 is trapped in the strongly curved potential, i.e. $\dot{x} = -\frac{dV}{dx}$, in dark green of 2.6(a). It could be imagined as a ball in a mountain area. Every small perturbation, indicated by the blue double arrows, is strongly forced to return to the potential valley, indicated by the downwards pointing dark green arrows. Note, that the forcing is represented by the gradient of the green vectors corresponding to the potential curvature. [Sch+12] If the system approaches a bifurcation point, the underlying potential changes its shape, begins to flatten and develops a wider valley. A perturbation in 2.6(b) is not strongly forced by the potential walls like in 2.6(a). The blue dot therefore returns more slowly to its original state, illustrated by the small inwards pointing dark green vectors and their gradients. It is also possible that a perturbation is not yet entirely decayed, before the system is perturbed a second time. The phenomenon of CSD can have direct consequences for the statistical measures presented in section 2.4. Due to the lower recovery rates near a bifurcation, noise effects accumulate and the standard deviation σ tends to increase. At the same time adjacent measured observables are more and more similar to each other resulting in an increase of certain statistical measures as the autocorrelation $\rho_{X_t,X_{t-1}}$. The measures are defined in section 2.4.[Sch+12; Dak+12]



(a) Potential for fast decay

(b) Potential for slow decay

Figure 2.6: The green potential well in 2.6(a) could be imagined as a very narrow valley with steep flanks. Every perturbation of a system state, here illustrated as a blue ball with a double arrow, returns due to the strong forcing, indicated by the dark green arrows with steep gradient. If the system approaches a critical transition, the underlying potential can change its shape like shown in 2.6(b). Perturbations in the potential near a bifurcation point decay more slowly than before.

2.5.2 Flickering

The phenomenon of CSD is less visible for moderate and high noise perturbed data like the biological model of overexploitation of some biomass resource, shown in Figure 2.7. The noise drives the system relatively far away from its equilibrium state that is indicated by the green solid line. Due to a changing parameter environment a second stable and an unstable fixed point, indicated by the green dashed line, emerge in a saddle-node bifurcation around time step 3200. The unstable fixed point separates the two basins of attraction. In this interesting region the data is driven to the lower state if some noise perturbation is great enough to overcome the separatrix. For this reason the system jumps in an irregular manner between the two stable branches, casually said, flickers around the two stable states and in this way gives the name *flickering* to the phenomenon. Furthermore, near the separatrix the dynamics are typically slow and the system is bound for longer time near the separatrix. The probability distribution of the system changes over that time in an asymmetric manner, because the system may stay first of all preferably in the upper and later preferably in the lower state and is delayed near the separatrix in the transition region. This often leads to an intermittent rise of statistical measures as the skewness Γ in the transition region. Also the variance σ^2 of the system can be increased by this flickering effect. The measures are presented in section 2.4. Finally, the first stable and the unstable state annihilate in a second saddle-node bifurcation around time step 6000 and just one stable branch remains in the system. [Sch+12; Dak+12]

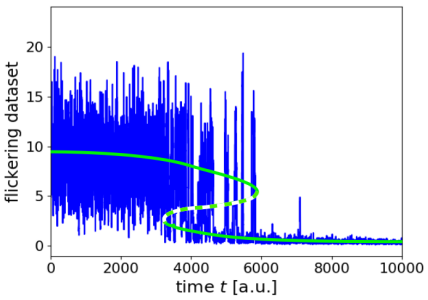


Figure 2.7: The dataset is based on a biological model for overexploitation of biomass resource. The data is highly stochastic and thus fluctuates strongly around some average value. The first stable state is indicated by the upper green solid line. By changing the system's control parameter, a second basin of attraction appears, divided by the green dashed separatrix. the data is flickering between the two stable states due to moderate or high noise influence and settles down to the lower branch after further parameter changes. This can give rise to a changing skewness of the system in the transition region.

2.6 Time series analysis

In most of the practical applications and for investigation of real world dynamical systems the governing mathematical laws, in general differential equations or iterated maps discussed in section 2.1 are unknown or in the best case partially known. Instead, the dynamical development of a certain measurable observable is given in sequential time steps Δt. Data sampled sequentially in time is called a *time series*. Examples for time series are the daily precipitation measured over a year for a specific place or the concentration of CO_2 in the atmosphere. Moreover, the investigated phenomena are not purely *deterministic*. Their time evolution is not fully governed by fixed mathematical laws, but it is influenced by some stochastic and noise contributions. In the following the considered phenomena are *stochastic*. Every time series is regarded to be one realisation of an infinite number of possible realisations due to the probability density function of a given stochastic process. An important feature of time series data

is that adjacent data points usually depend on each other. For simplicity all derivations in this chapter deal with equally sampled time series. [Box+15]

2.6.1 Takens' embedding theorem

Takens' theorem [Tak81] is widely used in time series analysis to gain information of complex systems and to reconstruct their state space dynamics out of time series data. Following the descriptions in [Sta+97] a short introduction to the basic mathematical ideas is given in this subsection.

A deterministic dynamical system's state is typically characterized by a point ξ lying on an m dimensional manifold M with $M \subseteq \mathbb{R}^{\mathrm{k}}$. The time evolution in state space is governed by a map $f^t_{\mathrm{state}} : M \times \mathbb{R}$, so that the new state at time $t + \tau$ of the state ξ at time t is given by $f^\tau_{\mathrm{state}}(\xi)$. This map $f^t(\xi)$ can typically be seen as the solution of a certain set of ordinary differential equations.

The original state ξ in state space is normally not accessible, but it is possible to measure some characteristical scalar quantity $\varphi(\xi)$ of the whole dynamical system depending on the system state ξ. The function $\varphi : M \to \mathbb{R}$ is called the *measurement function*. In this notation the time evolution of a measured scalar variable is defined by $\varphi(f^t_{\mathrm{state}}(\xi_0))$. Takens' theorem provides a mathematical approach to obtain information about the dynamics of the original system. If f^t_{state} and $\varphi(\xi)$ have certain properties[1], the state space dynamics can be reconstructed by scalar measurements $\varphi(\xi)$ except from an unknown, but smooth coordinate change.

Therefore, one introduces the *delay embedding map* $\Phi_{f_{\mathrm{state}},\varphi} : M \to \mathbb{R}^{\mathrm{d}}$ with the *embedding dimension d* in the form

$$\Phi_{f_{\mathrm{state}},\varphi}(\xi) = (\varphi(\xi), \varphi(f_{\mathrm{state}}(\xi)), \varphi(f^2_{\mathrm{state}}(\xi)), ..., \varphi(f^{(d-1)}_{\mathrm{state}}(\xi))). \tag{2.34}$$

In this case Takens' theorem states

> Let M be a compact m dimensional manifold. If $d \geq 2m + 1$, then the set of $(f_{\mathrm{state}}, \varphi)$ for which the map $\Phi_{f_{\mathrm{state}},\varphi}(\xi)$ is an embedding is open and dense in $\mathcal{D}^r(M) \times C^r(M, \mathbb{R})$ for $r \geq 1$. [Sta+97]

An *embedding* is called an arbitrary map $\Psi : M \to N$ if it is diffeomorphic with respect to the manifolds M and N. The set of C^r diffeomorphisms of M is denoted by $\mathcal{D}^r(M)$. $C^r(M, \mathbb{R})$ denotes the set of observation functions on M.

Many time series analysis methods are based on this theorem. Assuming that the necessary mathematical conditions for the system are fulfilled the mapped image $\Psi(M)$ is equivalent to M itself. Furthermore, it is possible to define a map $F_{\mathrm{obs}} = \Psi \circ f_{\mathrm{state}} \circ \Psi^{-1}$ for the mapped image $\Psi(M)$. Apart from the smooth coordinate change all coordinate invariant features of F_{obs} will be identical to f_{state}. In other words, all information gained from F_{obs} are information of the real state space dynamics f_{state} as well. The important point is that this invariant features of F_{obs} are accessible by the delay embedded observables

$$\vec{z}_{\mathrm{n}} = (\varphi(\xi_0), \varphi(f^{\mathrm{n+1}}_{\mathrm{state}}(\xi_0)), \varphi(f^{\mathrm{n+2}}_{\mathrm{state}}(\xi_0)), ..., \varphi(f^{\mathrm{n+d\text{-}1}}_{\mathrm{state}}(\xi_0))) \tag{2.35}$$

$$= (\varphi(\xi_0), \varphi(f^{1}_{\mathrm{state}}(\xi_{\mathrm{n}})), \varphi(f^{2}_{\mathrm{state}}(\xi_{\mathrm{n}})), ..., \varphi(f^{\mathrm{d\text{-}1}}_{\mathrm{state}}(\xi_{\mathrm{n}})) \tag{2.36}$$

[1] For the detailed mathematical properties see [Sta+97].

The map F_{obs} exhibits the same behaviour in the delay embedded observable space as its original counterpart $f_{\text{state}}(x_{\text{n}}) = x_{\text{n+1}}$ in state space:

$$F_{\text{obs}}(\vec{z}_{\text{n}}) = \Phi \circ f_{\text{state}} \circ \Phi^{-1}(\vec{z}_{\text{n}}) = \Phi \circ f_{\text{state}} \circ \Phi^{-1}(\Phi(\xi_{\text{n}}) \tag{2.37}$$

$$= \Phi \circ f_{\text{state}}(\xi_{\text{n}}) = \Phi(\xi_{\text{n+1}}) = \vec{z}_{\text{n+1}}. \tag{2.38}$$

The derivation of Takens' theorem in this form requires the system to be deterministic and both the dynamics and the obsvervations to be autonomous, thus to depend only on the state ξ. Most of the time series of realistic applications are not fully deterministic, contain at least in parts stochastic or forced features and noise. Therefore, some generalization has to be proven if it should be a useful tool for real applications. The shortly presented theorem is generalized to such noisy systems in [Sta+97]. Furthermore, the question how to choose the embedding dimension d and an appropriate sampling rate τ for a time series is investigated in [KBA92; RCL94].

2.6.2 Autoregressive models AR(p)

Under the assumption that successive values of the time series depend on each other, a linear regression model is used to calculate future values of a time series by the p past values using the *autoregressive model* AR(p)

$$x_{\text{n}} = \phi_1 x_{\text{n-1}} + \phi_2 x_{\text{n-2}} + ... + \phi_p x_{\text{n-p}} + a_{\text{n}} \tag{2.39}$$

with an independent random shock a_{n} for the n^{th} time step of a *Gaussian white noise process* defined by a Gaussian distribution given in Equation 2.32. It is convenient to define the *backshift operator* $\mathbf{B}$ applied to a time series point x_{n} as follows

$$\mathbf{B}x_{\text{n}} = x_{n-1}. \tag{2.40}$$

Further, by defining the *autoregressive operator* $\phi_{\text{p}}(\mathbf{B})$ as

$$\phi(\mathbf{B}) = 1 - \phi_1 \mathbf{B}^1 - \phi_2 \mathbf{B}^2 - ... - \phi_{\text{p}} \mathbf{B}^{\text{p}}, \tag{2.41}$$

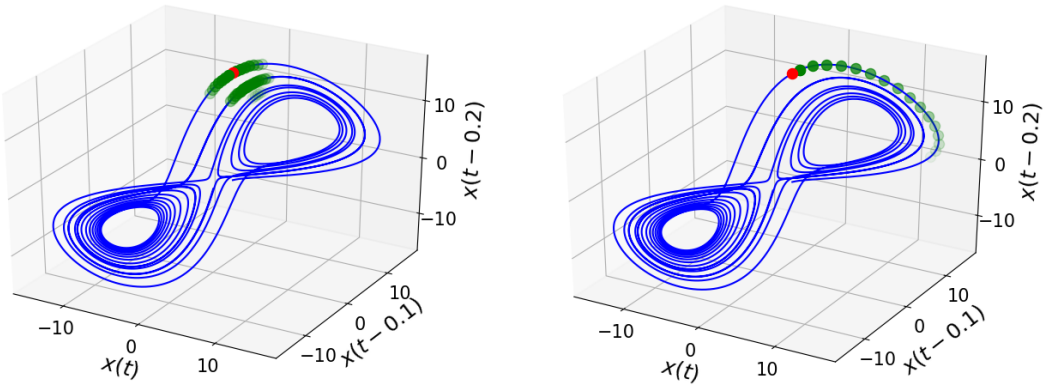
any autoregressive process of the order p can be efficiently written as

$$\phi_{\text{p}}(\mathbf{B}) x_{\text{n}} = a_{\text{n}}. \tag{2.42}$$

There are $p + 2$ unknown parameters ϕ_{n}, μ and σ that have to be estimated by a least squares fit of the time series data. AR(p) models are widely used for stationary and nonstationary processes. [Box+15]

2.6.3 S-map estimation

Dynamic linear models like AR(p), presented in subsection 2.6.2, predict future states by assuming a dependency between subsequent time series data, i.e. system states nearby in the time domain. The dynamics of such systems typically also depend on the dynamically similar states that are system states nearby in state space. Furthermore the AR(p) coefficients $\phi_{1,2,...,\text{p}}$



(a) S-map in state space

(b) AR(p) in state space

Figure 2.8: Both three dimensional figures show a state space reconstruction of the Lorenz attractor, originally shown in Figure 2.2. The reconstruction is done with Takens' embedding theorem. The S-map estimation and the AR(p) model use two different basic concepts in order to predict the following state from the current one. The current state is visualized as a red dot in 2.8(a) and 2.8(b). The S-map method uses a fixed number of next neighbours of the current state and weights the neighbouring states depending on their vicinity to the current state. The next neighbour is weighted the most. In the illustration the fifty first next neighbours of the current state are marked in green with a shading depending on their distance to the red point, in order to give an idea of the weights in the S-map method.
The AR(p) approach uses states on the attractor equally distanced in the time domain. In 2.8(b) every fifth of the first fifty past states of the red actual state are shown as green dots. The shading illustrates that the future state depends more on the recently past states than on states that are past for a long time.

are constant over the full range of the prediction interval. [Box+15]
The sequential map, abbreviated S-map, approach focuses on the state dependence of subsequent system states and is designed especially for processes that are not in static equilibrium by sequentially recalculating the interaction coefficients for each state. For a system, originally described by d variables, a corresponding state space vector $\vec{x} \in \mathbb{R}^d$ consists of entries of all d variables for a given time t^*. The basic model equation

$$\hat{x}_n(t^* + 1) = S_0 + \sum_{j=1}^{d} S_{nj} \cdot x_j(t^*) \tag{2.43}$$

has a similar form as the AR(p) model in Equation 2.39. But in contrast to the AR(p) the set of coefficients S_0, S_{nj} is updated every time step in order to predict the next time series value $\hat{x}_n(t^* + 1)$ and thus, the S-map estimation results in a locally approximation in state space. The indices of the coefficients S_{nj} with $j = 1, \ldots d$ denote the d coefficients needed for the estimation of the n$^{\text{th}}$ spanvector component $\hat{x}_n(t^* + 1)$ at time $t^* + 1$. Furthermore, the coefficients are estimated based on dynamically similar states instead of using nearby states in the time domain. Formally speaking, the Z next neighbours of a state space vector $\vec{x}(t^*)$ are used to estimate the coefficients for the S-map in order to obtain $\vec{\hat{x}}(t^* + 1)$, whereby every neighbour is weighted with respect to its distance to $\vec{x}(t^*)$. The weights are realized in an exponential

manner. [Dey+16]
In order to determine the S-map coefficients for one time t^*, the system

$$\vec{V}^n_{t_k+1} = \underline{\underline{W}}_{t_k} \vec{S}^n \tag{2.44}$$

is solved for each component, here denoted by n. The vector $\vec{S}$ contains the coefficients $S_0, ..., S_d$ that have to be estimated for Equation 2.43. The $Z \times (d+1)$ matrix $\underline{\underline{W}}_{t_k}$ consists of a first column filled with ones for the constant offset S_0 and the k weighted next neighbouring state space vectors at time t_k in each row:

$$W_{kj} = w_k \cdot x_j(t_k). \tag{2.45}$$

The vector $\vec{V}_{t_k+1}$ consists of the weighted n^{th} component of the k next neighbouring vectors at time t_{k+1}:

$$V_k = w_k \cdot x_n(t_k + 1). \tag{2.46}$$

The next neighbours are obtained from a state space reconstruction based on Takens' theorem that is described in subsection 2.6.1. Typically, the system is overdetermined because of $Z \gg (d+1)$ and thus, the solution is given by a singular value decomposition of the system of equations that is a least squares solution for the coefficients $\vec{S}$. [Dey+16; Sug94] The difference between the S-map approach and the AR(p) process are outlined in Figure 2.8.

3 Numerical methods

In order to analyse properties of a complex system with time series data some numerical tools and algorithms are necessary. They are based on the theoretical descriptions of chapter 2 and will be discussed in a practical and instructive manner in the following sections.

3.1 Takens' embedding theorem in numerics

The basic implementation of Takens' theorem in numerics, theoretically explained in subsection 2.6.1, is rather straight forward. A nonlinear stochastic process is investigated by sampling some observable of the system of interest over time. The resulting time series X_t is shifted $d-1$ times about some predefined time lag τ to create the time lagged time series. The embedding dimension is denoted by d. Together with the non-shifted time series they are used to construct the delay embedding variables $\vec{z}$. A more difficult and controvert discussed question is how to find a suitable embedding dimension d and a sufficient time lag τ. The following subsections describe two of the numerous possible numerical methods to solve this problem. [KS04]

3.1.1 Defining the embedding dimension

The mathematical Takens' theorem states that a system with an original dimension m can be definitely delay embedded in every dimension $d \geq 2m+1$. For any algorithm that needs a delay embedding space it is convenient to find the lowest practical embedding dimension, often $d \leq 2m+1$, in order to save computation time and avoid redundancy in the delay embedded vectors that can lead to numerical artefacts. The *false nearest neighbour*, abbreviated fnn, *algorithm* is an efficient and robust possibility to find a sufficient embedding dimension d. [KS04; KBA92]

Imagine a system of dimension m. If an embedding dimension $d < m$ is chosen some axes of the original state space are eliminated. Points that are very different in their dynamics, thus, that do not lie next to each other in the original state space, can be forced to come close together by the missing axes in the delay embedding space. This points in the delay embedding space are so-called *false next neighbours* because they would not be neighbours in a sufficient high delay embedding. Based on this idea it is necessary to evaluate

$$X_{\text{fnn}} = \frac{\sum_{n=1}^{N-m-1} \Theta\left(\frac{|s_n^{(m+1)} - s_{k(n)}^{(m+1)}|}{|s_n^{(m)} - s_{k(n)}^{(m)}|} - R_{\text{threshold}}\right) \Theta\left(\frac{\sigma}{R_{\text{threshold}}} - |s_n^{(m)} - s_{k(n)}^{(m)}|\right)}{\sum_{n=1}^{N-m-1} \Theta\left(\frac{\sigma}{R_{\text{threshold}}} - |s_n^{(m)} - s_{k(n)}^{(m)}|\right)} \tag{3.1}$$

with the standard deviation σ, a distance threshold $R_{\text{threshold}}$, the next neighbour $s_{k(n)}^{(m)}$ of $s_n^{(m)}$ in m dimensions, whereas $k(n)$ denotes the index k with $|s_n - s_k| = \min$ and the Heaviside function $\Theta(\cdot)$. The threshold has to be chosen large in order to take into account exponential divergence due to deterministic chaos. The first Heaviside function in the nominator yields unity if the

neighbours in m dimensions are false next neighbours. In other words, their distance increases over the critical threshold $R_{\text{threshold}}$. The second Heaviside function filters all pairs that cannot be false next neighbours due to an initial distance greater than $\frac{\sigma}{R_{\text{threshold}}}$. The term defines the maximum possible distance on average. Depending on the threshold $R_{\text{threshold}}$ the ratio X_{fnn} of false next neighbours should decrease with the embedding dimension m. If a minimal value is reached, this is used as the right embedding dimension d and the corresponding best time lag is calculated with the method of the following subsection. [KS04; KBA92]

3.1.2 Defining the time lag

The mathematical Takens' theorem does not state anything about a convenient time lag τ. A delay embedding in m dimensions with two different time lags τ would be equivalent from the mathematical point of view [KS04]. However, in practise it is necessary to take into account the relation between the internal dynamics' time scale and the sampling rate of a time series and in consequence the time lag τ. If the time lag is small compared to the internal dynamics' time scale, the delay embedded values are very similar to each other. This strong correlation will lead to points located along the main diagonal of the delay embedding space. If the time lag is large compared to the internal dynamics' time scale, the correlation approaches zero and the points fill a large cloud in the $\mathbb{R}^m$ whereby the deterministic part is confined to very small scales of the structure. [RCL94]
This geometrical reasoning motivates the method of *average distance from diagonal*, abbreviated adfd. The measure for the average distance $\langle S_m \rangle$ from diagonal for a given delay embedding is defined as

$$\langle S_m(\tau) \rangle = \frac{1}{N} \sum_{i=1}^{N} |\vec{X}_i^{\tau} - \vec{X}_i^{0}| \tag{3.2}$$

with the Euclidean distance measure $|\cdot|$ of the delay vector $\vec{X}_i^{\tau}$ from the non-delayed vector $\vec{X}_i^{0}$. In practise, the average distance from diagonal can be calculated with the scalar time series X_t and the formula

$$\langle S_m(\tau) \rangle = \frac{1}{N} \sum_{i=1}^{N} \sqrt{\sum_{j=1}^{m-1} (X_{i+j\tau} - X_i)^2}. \tag{3.3}$$

An optimal time lag τ is found for a maximum average distance $\langle S_m(\tau) \rangle$ from diagonal, because the attractor is well unfolded and the *redundancy error* is inverse proportional to the average distance to diagonal and thus, minimized. The redundancy error is typically great for small delay times because successive coordinates are almost strongly correlated and similar. In this case additive noise has a relatively strong effect. The redundancy error decreases with bigger time lags τ and reaches a minimum because an attractor is bound to a certain volume in state space. The *irrelevance error* is the error due to dynamical error of exponential growth. The irrelevance error increases with greater time lag because of the extended time for this exponential divergence with greater time lag τ. In reality it would be necessary to find the best

trade-off between redundancy error and irrelevance error, but quantifying the latter is not possible. For this reason the effects of irrelevance error are assumed to be small and negligible relative to the redundancy error. [RCL94]

3.2 Methods to anticipate critical transitions

Predicting critical transitions from time series data is challenging by the fact that one has almost no information about the underlying dynamical laws, e.g. an n-dimensional iterated map or a set of differential equations, described in subsection 2.1.1 and subsection 2.1.2. Therefore, practical numerical applications in the field are forced to rely on formal assumptions in order to estimate predictive measures for an uprising critical transition, so-called *leading indicators* or *early warning signals*. The following subsections give further insight into the basic assumptions and the methods how to estimate various early warning signals from time series data.

3.2.1 The dominant eigenvalue estimated with AR(p)

Beginning with Equation 2.7 one can apply Takens' embedding theorem in order to estimate the eigenvalues of the Jacobian and to get some insight into the stability of a given state. First, recapitulate the significance of the formula

$$\vec{x}_{\mathrm{n+1}} = J|_{(x^*,y^*)} \cdot \vec{x}_{\mathrm{n}} + \vec{x}_{\mathrm{offset}}. \tag{3.4}$$

Originally, the formula starts with modelling the linear stability of a fixed point. In the form given here in Equation 3.4 the linear time evolution of a state vector $\vec{x}_{\mathrm{n}}$ is modelled by first applying the Jacobian $J|_{(x^*,y^*)}$ to that state vector and second by adding some offset $\vec{x}_{\mathrm{offset}}$. In real applications, the full set of observables that determine the system is typically unknown. In other words, just a few or even no entries of the state vector $\vec{x}_{\mathrm{n}}$ are known and, in consequence, the original state space dynamics are not accessible by the given equation. Nevertheless, it is possible to get insight into the original dynamics by applying Takens' embedding theorem. Without loss of generality, assume just one quantity x is measured in a nonlinear stochastic time series X_t. Following the notation in subsection 2.6.1 it is an observable $x_{\mathrm{n}} = \varphi(\xi_{\mathrm{n}})$ of the state space. Equation 3.4 can be rewritten in terms of the time delayed embedding variables $\vec{z}_{\mathrm{n}}$ by constructing vectors with time delayed data points $\vec{z}_{\mathrm{n}} = (x_{\mathrm{n}}, x_{\mathrm{n\text{-}1}}, ..., x_{\mathrm{n\text{-}d+1}})$ from X_t:

$$\vec{z}_{\mathrm{n+1}} = \begin{pmatrix} x_{\mathrm{n+1}} \\ x_{\mathrm{n}} \end{pmatrix} \tag{3.5}$$

$$= \begin{pmatrix} j_1 & j_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{\mathrm{n}} \\ x_{\mathrm{n\text{-}1}} \end{pmatrix} + \begin{pmatrix} z_{\mathrm{offset}} \\ 0 \end{pmatrix} \tag{3.6}$$

$$= J_{\mathrm{embed}} \cdot \vec{z}_{\mathrm{n}} + \vec{z}_{\mathrm{offset}}. \tag{3.7}$$

In consequence of Equation 3.7 the Jacobian J_{embed} should contain the linear part of the observable states' dynamics F_{obs} that is equivalent to the linear part of the original dynamics f_{state}. Thus, the eigenvalues of the Jacobian J_{embed} and the corresponding stability of the system state

can be calculated by estimating the coefficients j_1, j_2 and z_{offset}. A simple way to estimate the coefficients of Equation 3.6 is the autoregression AR(p) process, explained in subsection 2.6.2. The order p of the process is assumed to be given by the optimal delay embedding dimension of the time series X_t, determined by the false nearest neighbour algorithm. For further details of the algorithm, see subsection 3.1.1. Basically, the general equation

$$x_{n+1} = \sum_{i=0}^{p-1} j_{i+1} \cdot x_{n-i\tau} + z_{\text{offset}} \tag{3.8}$$

needs to be solved. The equation can be economically rewritten as

$$\phi_p(\mathbf{B}) x_{n+1} = z_{\text{offset}} \tag{3.9}$$

with the AR(p) formalism of subsection 2.6.2 with the coefficients $j_{i+1} = \phi_{i+1}$ and a sampling $\Delta t = \tau$ of the time series. The Equation 3.8 is the generalized form of the first row of Equation 3.6 written for a delay embedding space of dimension $d = p$. The available data can be used to train the AR(p) model and to fit the desired parameters by a pre-implemented function in the python package *statsmodels*. The function fits the coefficients j_{i+1} via an unconditional maximum likelihood for a given order p and a constant time lag $\tau = \Delta t$ with a constant offset z_{offset}. The coefficients are inserted in the Jacobian and the eigenvalues of the Jacobian are determined.

3.2.2 The dominant eigenvalue based on flow estimation

The *flow* [Str15] is called the infinitesimal change of a system variable over time. It is often signed by $\dot{x}$ as the left hand side of an ordinary differential equation that describes the variable's time evolution in the complex dynamical system of interest. It is usually part of a system of ordinary differential equations as described in subsection 2.1.2. If the system is slightly driven out of its stable equilibrium position, it relaxes over time towards the stable fixed point x^*. This change of a system variable x in time, respectively the flow $\dot{x}$, denotes the time evolution of the perturbation η around the fixed point in this special case. Expressed in mathematical terms, the flow is

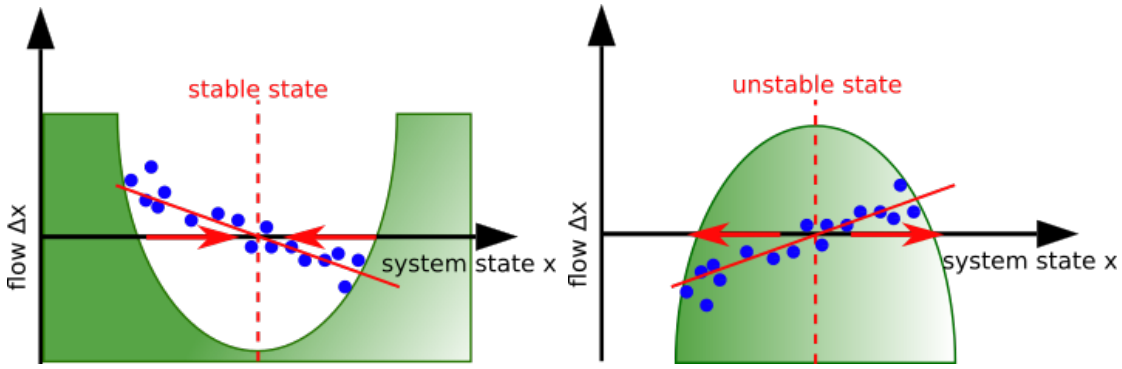
$$\dot{x} = \dot{x^*} + \dot{\eta} = \dot{\eta}, \tag{3.10}$$

and in consequence nothing else than the time evolution $\dot{\eta}$ of the perturbations. That can be numerically estimated.
Therefore, the differences

$$\dot{x} \approx \Delta x = x_{t+1} - x_t = x^* + \eta_{t+1} - (x^* + \eta_t) = \eta_{t+1} - \eta_t = \Delta\eta \tag{3.11}$$

of subsequent time series values are calculated. In Equation 3.11 the sampling of the time series data is assumed to be small such that it is an approximation for the infinitesimal time step dt of $\dot{x} = \frac{\mathrm{d}x}{\mathrm{d}t}$. In consequence, the differences Δx are scaled to that time sampling interval. In a second step the function values $\Delta x(x)$ are plotted. They are schematically shown as blue dots in Figure 3.1. The system in 3.1(a) fluctuates around a stable equilibrium. The projection

of the flow to the x axis is marked by the red arrows. It points to the right for x left from the stable state and vice versa for x right from the stable state, indicating a positive flow at the left and a negative flow at the right side of the stable state. This behaviour is exactly the opposite for the unstable case, shown in 3.1(b). If one fits the function $\Delta x(x)$ for example with a linear model[1], as schematically suggested by the red solid line in Figure 3.1, one can estimate its slope and thus, regarding the theory of linear stability for differential equations in subsection 2.3.1, obtain an estimate of the eigenvalues in the system of differential equations. In the case of more complicated fit functions one has to evaluate the first derivative of the fitted function to obtain an estimate of the DEV.



(a) Flow estimation (stable case)

(b) Flow estimation (unstable case)

Figure 3.1: In 3.1(a) the procedure to estimate the eigenvalue by flow reconstruction in numerics from time series data is schematically shown. Assuming a parabolic potential, shown in green, around a stable fixed point, the blue points are values of the function $\Delta x(x)$ that assigns the difference $\Delta x = x_{t+1} - x_t = \Delta\eta$ to a time series point x_t. The red arrows are projections of the relaxation flow towards the stable fixed point. By fitting the function $\Delta x(x)$, here suggested by the red solid linear line, one can estimate the eigenvalue of a system of differential equations as the slope of the fitted model. Higher order polynomials or other mathematical expressions could be used as well. The DEV is obtained from the first derivative of the fitted model. In 3.1(b) it is presented the same idea for an unstable fixed point.

3.2.3 Statistical measures as early warning signals

In section 2.4 the statistical measures of standard deviation σ, skewness Γ and the autocorrelation $\rho_{X_t,X_{t-\tau}}$ are described. They can serve as indirect measures of an approaching critical transition due to CSD and flickering, described in detail in subsection 2.5.1 and subsection 2.5.2. The estimation of these quantities is computed by statistical formulas [Dak+12], i.a. designed for time series data X_t of length n. To calculate the standard deviation σ for a given dataset, the

[1]The easiest fit function is a linear one, but the procedure can also be applied to higher order polynomials or other mathematical expressions as fit functions.

variance

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \mu_{X_t})^2 \tag{3.12}$$

with the mean μ_{X_t} of the data is used. The standard deviation is the root of the variance.
The calculation of the autocorrelation $\rho_{X_t,X_{t-\tau}}$ for the time lag τ from a finite dataset is done via

$$\rho_{X_t,X_{t-\tau}} = \frac{\sum_{i=\tau}^{n}(X_i - \mu_{X_t})\cdot(X_{i-\tau} - \mu_{X_t})}{(n-\tau-1)\cdot\sigma^2} \tag{3.13}$$

with the mean μ_{X_t} and the standard deviation σ of the data X_t.
The skewness Γ as a leading indicator for a critical transition with antecedent flickering is the third statistical moment. Therefore, it is calculated in practise similar to the second statistical moment σ^2 as

$$\Gamma = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_{X_t})^3}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_{X_t})^2}}. \tag{3.14}$$

3.2.4 Rolling window method

If a dynamical system undergoes a parameter variation over time, the process is non-stationary. In this case a recorded time series of the system is characterized by time varying statistical measures, as variance and autocorrelation, and contains information about the varying complex dynamics of the system. An estimation of such statistical measures or dynamical variables, e.g. the dominant eigenvalue, one time over the whole time series is consequently useless and results in wrong characterization of the underlying process. The numerical *rolling window method* provides the possibility to estimate the variation of such measures approximately in time. The method states that for sufficient small *time segments* or *windows* $[t_{\text{start}}, t_{\text{start}} + t_{\text{window}}]$ of a non-stationary time series the dynamics are approximately stationary. Thus, a reasonable value for the statistical measures and some dynamical variables can be calculated for each single time window. The time window is successively shifted in time by Δt and the calculations are repeated for each shifted time window $[t_{\text{start}} + \Delta t, t_{\text{start}} + t_{\text{window}} + \Delta t]$. Following this method one can gain information about a particular measure's evolution in time. The approach is exemplarily used in Figure 3.2 for

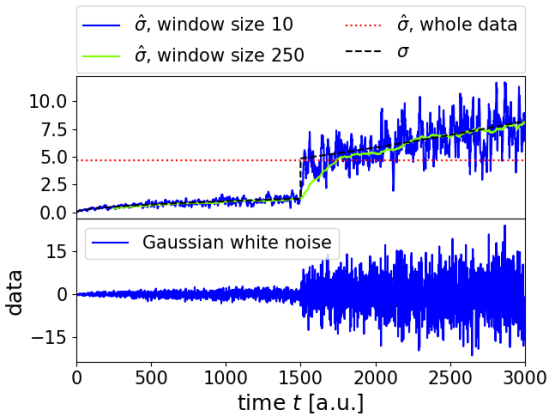


Figure 3.2: The upper graph shows the time evolution of the standard deviation of a Gaussian white noise process by the black dashed line. The resulting dataset, shown in the lower plot, is used to estimate the standard deviation via the rolling window method. A smaller window size leads to faster reactions, but higher volatility and vice versa.

Gaussian white noise, introduced in subsection 2.4.6, with a standard deviation σ that changes in time, indicated by the black dashed line. If the standard deviation would be estimated once over the whole time series, the erroneously assumed stationarity leads to a wrong standard deviation σ, given by the red dotted line. The solid blue curve is the standard deviation estimate $\hat{\sigma}$ for a window size of 10 with step size of unity. The curve responds fast to the changing standard deviation, but is rather volatile. The opposite is given by the standard deviation estimate $\hat{\sigma}$ for a window size of 250 with step size of unity, given by the green solid curve. It is much less volatile, but reacts rather slow to changes in the standard deviation as illustrated by the discrete step at time 1500. Therefore, one has to find some trade-off for the window size in real applications. [ZW06; Ehe17]

3.2.5 The dominant eigenvalue estimated with S-maps

The S-map method, theoretically introduced in subsection 2.6.3, is similar to the AR(p) an autoregressive estimation process of the form

$$\hat{x}_n(t^* + 1) = S_0 + \sum_{j=1}^{d} S_{nj} \cdot x_j(t^*) \tag{3.15}$$

with the regression coefficients S_0 and S_{nj} for the estimation of $\hat{x}_n(t^* + 1)$. The coefficients S_{nj} are the desired solutions of Equation 3.8 and the derivation of the method is equivalent to that in subsection 3.2.1 up to that equation. To solve the latter equation and to implement the S-map estimation method from a single times series, it is necessary to find a suitable embedding dimension d and a time lag τ. This can be done with the methods of subsection 3.1.1 and subsection 3.1.2. Thus, the general formulated equations of subsection 2.6.3 with original state space entries or multivariate time series entries are implemented in the delay embedding space from a single time series.
There are two main differences to AR(p) processes. First, the autoregression is done sequentially from each t^* to $t^* + 1$ with a new set of coefficients. For the AR(p) the coefficients are estimated once from all t_n to t_{n+1} within the time series. Second, the data is not auto-regressed to data that is past in time, but to a specified number Z of next neighbours in state space. These neighbours are weighted in an exponential manner by their absolute distance to the state vector at time t^*, implemented by the weights

$$w_k = \exp\left(\frac{-\theta \cdot |\vec{x}(t_k) - \vec{x}(t^*)|}{\bar{r}}\right) \tag{3.16}$$

with the Euclidean norm $|\cdot|$, a nonlinearity factor θ and the mean distance $\bar{r} = \frac{1}{k}\sum_{k=1}^{Z} |\vec{x}(t_k) - \vec{x}(t^*)|$. A nonlinearity factor $\theta = 0$ results in linear weighting, larger values for θ give stronger weight to the most proximal next neighbours of $\vec{x}(t^*)$ in state space. The general equation that needs to be solved in order to obtain the desired coefficients S_{nj} is given in Equation 2.44. For clearness it is written again as

$$\vec{V}^n_{t_k+1} = \underline{\underline{W}}_{t_k} \vec{S}^n \tag{3.17}$$

with the weighted vector $\vec{V}^n_{t_k+1}$ of the n^{th} component of the Z next neighbours, the weighted matrix $\underline{\underline{W}}_{t_k}$ with the first column of unity for the constant offset S_0 and the entries of the Z next neighbours at time t_k in each row. The vector $\vec{S}^n$ contains the desired coefficients.
Due to the delay embedding of a single time series the general Equation 2.44 can be explicitly implemented as

$$\underbrace{\begin{pmatrix} w_{\text{NN}} & 0 & 0 & \dots & 0 \\ 0 & w_{\text{NN-1}} & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & w_{\text{NN-Z+1}} & 0 \\ 0 & 0 & \cdots & 0 & w_{\text{NN-Z}} \end{pmatrix}}_{=:\underline{\underline{W}}_{t_k}} \cdot \begin{pmatrix} x_{\text{NN}}(t+\tau) \\ x_{\text{NN-1}}(t+\tau) \\ x_{\text{NN-2}}(t+\tau) \\ \vdots \\ x_{\text{NN-Z}}(t+\tau) \end{pmatrix} \tag{3.18}$$

$$= \underline{\underline{W}}_{t_k} \cdot \begin{pmatrix} 1 & x_{\text{NN}}(t) & x_{\text{NN}}(t-\tau) & \dots & x_{\text{NN}}(t-(d-1)\cdot\tau) \\ 1 & x_{\text{NN-1}}(t) & x_{\text{NN-1}}(t-\tau) & \dots & x_{\text{NN-1}}(t-(d-1)\cdot\tau) \\ 1 & x_{\text{NN-2}}(t) & x_{\text{NN-2}}(t-\tau) & \dots & x_{\text{NN-2}}(t-(d-1)\cdot\tau) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{\text{NN-Z}}(t) & x_{\text{NN-Z}}(t-\tau) & \dots & x_{\text{NN-Z}}(t-(d-1)\cdot\tau) \end{pmatrix} \cdot \begin{pmatrix} S_0 \\ S_{(1,1)} \\ S_{(1,2)} \\ \vdots \\ S_{(1,d)} \end{pmatrix}$$

with the weights w and the time series entries x of the next neighbours, whereby the next neighbour is denoted by NN and the most distanced neighbour by NN $- Z$, where Z is the number of next neighbours taken into account. The system consists typically of more rows than columns because of $Z \gg (d+1)$ and its solution is obtained with a singular value decomposition, implemented by a predefined python function of the *numpy.linalg* package. [Dey+16; Dao07]

Note, that there are two practical issues with respect to the rolling window method. First, the method states that for each window the dynamics are approximately stationary. Thus, in this case a delay embedding with the aim of attractor reconstruction of that stationary dynamics is, strictly speaking, forced to be applied to each single time window again. In other words, the transformation into the delay embedding space should be applied separately for every time window. In this way one loses a few delay embedding vectors for every time window because of the necessary time shift for the embedding. For an adequate amount of time series data this should not result in a serious problem. In an alternative approach the delay embedding is just applied one time for the whole series. The time windows are shifted in the delay embedding space. Thus, one loses just once a few delay embedding vectors and creates the same delay embedding states as in the first variant. The criterion of stationarity is fulfilled for every single window within the delay embedding space.
Second, the S-map algorithm is designed to compute the coefficients of each delay embedding state separately in a given window. This results in a set of coefficients for one window and in consequence, in a set of dominant eigenvalues. This issue is solved in two ways:

1. For each window the dominant eigenvalues for each time step from t^* to $t^* + 1$ are calculated. In a second step the average is taken from all dominant eigenvalues calculated

before. This approach requires high computational and implementing effort and in consequence, the calculations are rather time consuming.

2. An average state vector for each rolling time window is defined and the calculations are performed just once for that vector. One does not need any further averaging. This approach could be seen as estimation of the average state evolution in time. The computational costs are relatively low.

3.3 Prediction quality measures

The implementation of the considered early warning indicators is described in section 3.2. In the following, two measures to investigate the quality of the introduced early warning signals will be presented. First, they give information about the robustness of the leading indicator methods depending on the numerical parameter choice. Second, they can serve as a measure of the predictive power of the leading indicators depending on the difference in time between the raised early warning and the actual critical transition time.

3.3.1 Bifurcation type classification quality of AR(p)

A successful estimation of the dominant eigenvalues in the Gaussian complex plane allows to classify various local bifurcations based on their characteristic eigenvalue time evolution in vicinity of a critical transition, summarized in subsection 2.3.2. Remember, e.g. a fold bifurcation should have purely real eigenvalues that approach unity just before the critical transition occurs. If the eigenvalue estimation process is not stable for a given window size or the chosen order p of the AR(p) process, it could produce scattering points, wrong eigenvalues with imaginary parts or even wrong time evolution trends in the complex plane. These numerical instability artefacts can result in wrong bifurcation type classification. In order to deal with this problem a quality measure is designed that should give an overview of the classification results in a range of window sizes and orders p. In the example of the fold bifurcation, the eigenvalues should stay on the real axis and move towards unity with time. This means that the distance of each single estimated eigenvalue to unity is the more important the closer the system is to the fold bifurcation in the time domain. This concept, expressed in mathematics, results in the quality measure

$$Q_{\text{fold}} = \frac{1}{N} \sum_{t=1}^{N} |1 - x_t| \cdot \exp\left(\frac{\gamma \cdot t}{N}\right) \tag{3.19}$$

for the classification of a fold bifurcation in a time series with N entries. $|\cdot|$ denotes the Euclidean norm. The distance of every eigenvalue to unity is weighted exponentially in time in order to account for the importance of the eigenvalue trend towards unity just before the bifurcation point. The weighting parameter γ can be set to reach a convenient measure resolution for a visualisation. Generally speaking, the smaller the quality measure Q_{fold} the better the type classification.
The measure is almost identical for a flip bifurcation due to the fact that the eigenvalues are

purely real again. The only difference consists in the time evolution trend towards minus unity, resulting in

$$Q_{\text{flip}} = \frac{1}{N}\sum_{t=1}^{N} |-1-x_t| \cdot \exp\left(\frac{\gamma \cdot t}{N}\right) \tag{3.20}$$

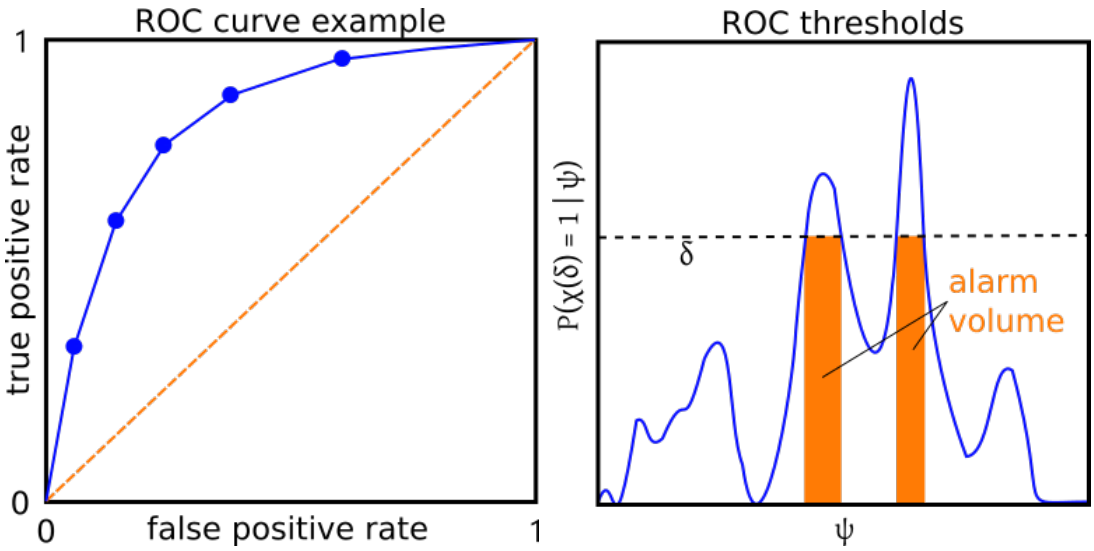
for the classification of a period doubling bifurcation in a time series with N entries.
For a Hopf bifurcation the situation is a little bit more complicated. The eigenvalues are complex conjugated and their absolute value approaches unity if a critical transition occurs. This is expressed as

$$Q_{\text{Hopf}} = \frac{1}{N}\sum_{t=1}^{N} ||x_t|-1| \cdot \exp\left(\frac{\gamma \cdot t}{N}\right) \tag{3.21}$$

to take into account the dependency of the absolute value for a Hopf bifurcation. This quality measure is implemented by controlling whether complex conjugated eigenvalues exist. If this is the case, the measure is evaluated for one set of these eigenvalues.

3.3.2 Receiver operating characteristic as prediction quality measure

In general, the *receiver operating characteristic*, abbreviated ROC, curve is a tool to visualize the classification quality of a stochastic classifier based on a certain measure. In the binary case one addresses the problem of distinguishing state A and state B by significant varying features of these two states. An ideal classifier should detect all states A and B at some threshold δ of the classifier without false declaration of any state. This cannot be achieved in real world applications. A real classifier is the better the more right classifications are done with the lowest possible miss classifications at the same time for a given threshold δ. This can be visualized in ROC curves, schematically shown in Figure 3.3. Different classification thresholds result in certain combinations of true positive and false positive classifications. For each threshold that changes the true and false positive rates significantly, a blue dot in the ROC plane is plotted. If the classifier would act in a random manner, the thresholds would lie on the orange dashed bisectrix. Each ROC curve, marked in blue, that exceeds the bisectrix is the result of a significant classifier. Each ROC curve under the bisectrix is the result of a miss classifier that will classify more states wrong than right. The thresholds could be written next to the blue dots if they are of interest. Summarized a classifier is the better the nearer it is to the upper left corner of the ROC plane. [Cen85]
In order to design a measure of the prediction quality of leading indicators ψ for critical transitions, happening in the future, the ROC curve is used following the instructions in [ZKH15]: It is introduced a binary tracer variable χ. In each time step t the tracer variable signs whether a bifurcation occurs at $t + k$ with the so-called *lead time* k. The tracer is set to unity at t for an uprising bifurcation at $t + k$ and zero otherwise.
The conditional probability $P(\chi = 1|\psi)$, theoretically described in subsection 2.4.1, can be used as the naive Bayesian classifier in order to construct the ROC curves. If the indicator ψ is known at time t, the conditional probability $P(\chi = 1|\psi)$ denotes the probability of a warning



(a) ROC curve scheme

(b) ROC discrimination threshold

Figure 3.3: In 3.3(a) the basic ROC scheme is illustrated: It serves as a measure whether a binary classifier is significant or not. The false positive rate is plotted on the x axis, the true positive rate on the y axis. Every blue dot corresponds to a chosen threshold of the classification task. The classifying power is the better the nearer the blue curve is to the upper left corner. A ROC curve congruent to the orange dashed bisectrix would be a random classifier. A ROC curve under the bisectrix indicates a miss classifier. The chosen thresholds δ of the blue dots could be written into the ROC plane if they are of interest.
The thresholds δ are determined by the procedure shown in 3.3(b): For a given leading indicator ψ the conditional probability $P(\chi = 1|\psi)$ is evaluated as a binary classifier. The binary tracer χ represents positive and negative events, here the occurrence or absence of a critical transition in the future with lead time k. If the conditional probability for a specific value of the measure ψ exceeds the threshold δ, the corresponding event is classified as a positive one, otherwise as a negative one. The classification is compared to the real distribution of positive and negative events to examine the true and false positive rate of the ROC curve, shown in 3.3(a), and to plot the corresponding blue dot. The threshold δ is varied continuously and the true and false positive rates are evaluated analogously for every variation of δ. Significant changes in the true and false positive rates lead to a new blue dot in the corresponding ROC curve.

at time t for a critical transition happening in future at $t + k$. A ROC curve constructed in this way is a measure for the prediction quality of the indicator ψ. The probability conditioned on the known indicator ψ at time t provides the possibility to raise an early warning based on the given indicator ψ. The concept of the lead time k accounts for the predictive character of the problem in task. The conditional probability can be calculated as

$$P(\chi = 1|\psi) = \frac{P(\psi,\chi)}{P(\psi)} \tag{3.22}$$

with the joint probability $P(\psi,\chi)$ of the indicator and the tracer and the marginal probability $P(\psi)$, both defined in subsection 2.4.1. The marginal probability $P(\psi)$ of an indicator ψ, e.g.

the standard deviation, the autocorrelation or the dominant eigenvalue, is estimated by making a histogram of the indicator values. Therefore, one divides the range of the indicator variable ψ in a fixed number of intervals, the bins of the histogram, respectively. The estimated indicator values are counted corresponding to their bin interval for all time steps t. The histogram bins are divided by the total number of events to obtain the marginal probability $P(\psi)$. This is the unconditioned probability to find ψ in a specific bin interval.
The joint probability of the pairs (ψ, χ) of the indicator and the tracer at each time t is estimated analogously by making a two dimensional histogram. The second dimension for the tracer has just two bins due to its binary definition. For practical reasons the number of bins for ψ is chosen equally for both, the marginal and the joint probability $P(\psi)$ and$P(\psi, \chi)$. Each of the bins is again divided by the total number of pairs (ψ, χ) in order to obtain the joint probability $P(\psi, \chi)$. The histograms are constructed using the python functions *numpy.histogram* and *numpy.histogram2d*.
With the conditional probability $P(\chi = 1|\psi)$ one can assign a probability for a critical transition warning to each time t if the corresponding indicator ψ is known. The decision whether an alarm is raised or not depends on the decision threshold δ. The thresholds in the ROC curve indicate significant variations in both, the false positive rate and the true positive rate. The significance of the threshold δ is shown in 3.3(b). The conditional probability $P(\chi = 1|\psi)$ is given for known indicator values ψ. If this conditional probability exceeds the continuously varied threshold δ, an early warning is raised as illustrated by the orange bins. Afterwards, one has to compute the true positive and false positive rate r_{tpr} and r_{fpr} of the raised early warnings. In stochastic tests these names refer to events, that are rightly classified positive or rather erroneously classified positive. They are defined by

$$r_{\text{tpr}} = \frac{n_{\text{tp}}}{n_{\text{tp}} + n_{\text{fn}}} \tag{3.23}$$

$$r_{\text{fpr}} = \frac{n_{\text{fp}}}{n_{\text{fp}} + n_{\text{tn}}}, \tag{3.24}$$

with the number of true positive, false negative, false positive and true negative events n_{tp}, n_{fn}, n_{fp} and n_{tn}, using the stochastic nomenclature [Loh12]. The true positive rate, the false positive rate and the corresponding thresholds are calculated using the pre-implemented python function *metrics.roc_ curve*.

3.4 Data simulation

The following section deals with the numerical methods that are used to generate time series from iterated maps and from differential equations that are influenced by noise. The description aims for a short conclusion of the practical basics of the procedures.

3.4.1 Euler-Maruyama method for differential equations

The stochastic one-step *Euler-Maruyama procedure* [KP92; Ehe17] is a simple approach to solve scalar Itô differential equations

$$\mathrm{d}x(t) = \alpha(t, x(t))\mathrm{d}t + \beta(t, x(t))\mathrm{d}W, \tag{3.25}$$

that consist of an ordinary differential equation part $\alpha(t, x(t))\mathrm{d}t$ and a stochastic non-deterministic part $\beta(t, x(t))\mathrm{d}W$ given by a *Wiener process* $\mathrm{d}W$ [Has07]. A Wiener process is a stochastic process that is continuous in time with normal distributed increments. The simple Euler scheme for ordinary differential equations is expanded for the stochastic part. The considered time interval is divided into equidistant time steps Δt. ΔW_j defines

$$\Delta W_j = W_{j+1}(t_j + \Delta t) - W_j(t_j) \tag{3.26}$$

and can be written as

$$\Delta W_j \sim \mathcal{N}(0, \Delta t) \sim \sqrt{\Delta t}\mathcal{N}(\mu = 0, \sigma^2). \tag{3.27}$$

In conclusion, the time series for a given stochastic differential Equation 3.25 is computed with the Euler-Maruyama procedure by iterating over

$$x(t + \Delta t) = x(t) + \alpha(t, x(t))\Delta t + \beta(t, x(t))\sqrt{\Delta t}\mathcal{N}(\mu = 0, \sigma^2). \tag{3.28}$$

3.4.2 Classical Runge-Kutta method

The commonly used one-step *Runge-Kutta 4 method* [Bro+08] for numerical integration is shortly introduced in this section. It is a method of stage four: It contains four supporting points to calculate the next estimated point of a given ordinary differential equation $\dot{y} = f(y, t)$. The integration is often performed over time t with a time step h. The next function value y_{t+1} is computed via

$$y_{t+1} = y_t + \frac{1}{6}(k_1 + 2 \cdot k_2 + 2 \cdot k_3 + k_4) \tag{3.29}$$

with the coefficients

$$k_1 = f(y_t, t) \tag{3.30}$$

$$k_2 = f\left(y_t + \frac{h}{2} \cdot k_1, t + \frac{h}{2}\right) \tag{3.31}$$

$$k_3 = f\left(y_t + \frac{h}{2} \cdot k_2, t + \frac{h}{2}\right) \tag{3.32}$$

$$k_4 = f(y_t + h \cdot k_3, t + h). \tag{3.33}$$

In Equation 3.29 an additional noise term is used to introduce noise in a system solved with the Runge-Kutta 4 method.

3.4.3 Iteration scheme for stochastic iterated maps

Iterated maps are designed to be discrete in time. This fact makes it very easy to calculate time series from iteration rules for specified parameters [Str15]. An N dimensional iterated map consists of a system of iteration rules as described in subsection 2.1.1. Partial stochastic behaviour of the calculated time series is modelled by adding some random white noise influence

ω_{noise} in every iteration step. For practical reasons one can scale the additive noise influence with a parameter ζ compared to the previous time series value x_n. In a formal way, the general iteration rule

$$\begin{aligned} x_{1,n+1} &= f_1(x_{1,n}, x_{2,n}, ..., x_{n,n}) + \omega_{\text{noise},1} \cdot \zeta \cdot x_{1,n} \\ x_{2,n+1} &= f_2(x_{1,n}, x_{2,n}, ..., x_{n,n}) + \omega_{\text{noise},2} \cdot \zeta \cdot x_{2,n} \\ &\vdots \\ x_{n,n+1} &= f_n(x_{1,n}, x_{n,n}, ..., x_{n,n}) + \omega_{\text{noise},n} \cdot \zeta \cdot x_{n,n} \end{aligned} \tag{3.34}$$

has to be evaluated for successive time steps. Incorporating noise in at least one component x_n of a coupled system is typically sufficient to account for stochastic influences.

4 Methodological investigations in B-tipping systems

In the first part of this chapter a set of simulated data with varying bifurcation types is presented. The numerical methods for anticipating critical transitions in complex systems, introduced in chapter 2, are applied to these datasets. In a last step, the methods are used to predict critical transitions from real time series data, thus real experimental applications.

4.1 Synthetic B-tipping data

The simulated datasets are chosen to generate a portfolio of different local bifurcations. The Ricker-type map provides a fold bifurcation, the Hénon map undergoes a period doubling bifurcation and the Hopf map passes a Hopf bicurcation. Apart from this simple maps with additive noise, a differential equation is used to simulate a process that exhibits the phenomenon of critical slowing down and flickering. The fold and the Hopf bifurcation have a real experimental data counterpart.

4.1.1 The Ricker-type model

The discrete Ricker-type model is developed to describe the population dynamics of a wide range of organisms like fish, insects and birds. The model is taken from [Dak+17] and slightly modified. It is written as

$$x_{n+1} = x_n \exp\left(r - \varpi \cdot x_n\right) - F\frac{x_n^2}{x_n^2 + \nu^2} + \omega_{\text{noise}} \cdot \zeta \cdot x_n, \tag{4.1}$$

where ω_{noise} defines a Gaussian white noise process with the Gaussian distribution $\mathcal{N}(\mu = 0, \sigma^2 = 0.0625)$. The model is simulated over 10000 time steps using the simple iteration scheme, described in subsection 3.4.3, with the initial condition $x_0 = 7.5$ and the parameters $\varpi = 0.1$, $\nu = 0.75$, $r = 0.75$ and $\zeta = 0.05$. The control parameter F is linearly

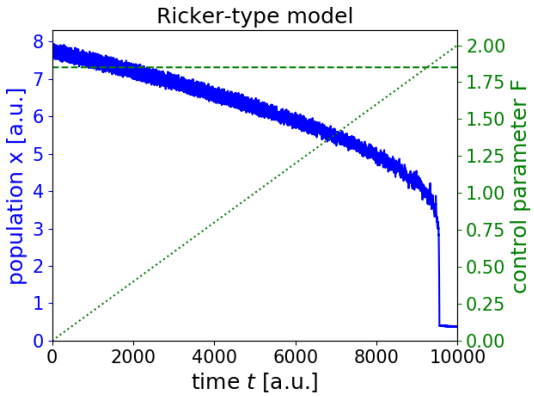


Figure 4.1: The time series generated by the Ricker-type model Equation 4.1 over 10^4 time steps with $\varpi = 0.1$, $\nu = 0.75$, $r = 0.75$ and $\zeta = 0.05$, Gaussian white noise ω_{noise} with $\mathcal{N}(\mu = 0, \sigma^2 = 0.0625)$ and the initial condition $x_0 = 7.5$ is shown by the blue solid line. At approximately time 9100 a fold bifurcation occurs. The corresponding threshold value of the control parameter $F \approx 1.85$, drawn as green dashed horizontal line, is reached around this time as shown by the crossing green dotted line.

increased from 0 to 2 over the whole time range. The critical threshold is around $F \approx 1.85$ [Dak+17]. The resulting time series is shown in blue in Figure 4.1.

4.1.2 The Hénon map

The investigated time series that exhibits a period doubling bifurcation is generated by the two dimensional Hénon map with additive Gaussian white noise. Originally, it was developed by Hénon in order to investigate the Lorenz equations by simplifying them and conserving their typical features at the same time. The mathematical model can be written as

$$\begin{aligned} x_{n+1} &= 1 - a \cdot x_n^2 + y_n + \omega_{\text{noise}} \cdot \zeta \cdot x_n \\ y_{n+1} &= b \cdot x_n \end{aligned} \tag{4.2}$$

with $b = 0.3$, the control parameter a and a Gaussian white noise process ω_{noise} with the Gaussian distribution $\mathcal{N}(\mu = 0, \sigma^2 = 0.0625)$. The white noise process scales with $\zeta = 0.01$ and the time series value x_n. The time series, shown in blue in Figure 4.2, is calculated by the simple iteration scheme of subsection 3.4.3 starting with $(x_0,\ y_0) = (0.3,\ 0.3)$. During successive iterations the control parameter is increased linearly from 0.1 to 0.4 over a range of 10^4 time steps, shown in green in Figure 4.2. The control parameter's threshold value is $a = 0.3675$, depicted as green dashed horizontal line [AAH13].

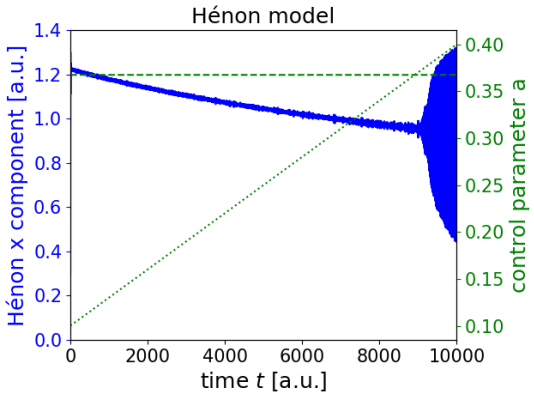


Figure 4.2: The time series generated by the Hénon map Equation 4.2 over 10^4 time steps with $b = 0.3$, $\zeta = 0.01$, Gaussian white noise ω_{noise} with $\mathcal{N}(\mu = 0, \sigma^2 = 0.0625)$ and with the initial values $(x_0,\ y_0) = (0.3,\ 0.3)$ is shown by the blue solid line. At around 9100 time steps a period doubling bifurcation occurs. The corresponding threshold value of the control parameter $a = 0.3675$, drawn as green dashed horizontal line, is reached around this time, shown by the crossing green dotted line.

4.1.3 The Hopf model

An iterated map that exhibits a Hopf bifurcation is presented in [KS04]. It is chosen in a slightly modified variant because of its simple computation costs and programming effort. The model can be written in the form

$$\begin{aligned} x_{n+1} &= \frac{1}{\sqrt{2}}\left(u(x_n - 1) + (y_n - 1) + (x_n - 1)^2 - 2(y_n - 1)^2\right) + 1 + \omega_{\text{noise}} \cdot \zeta \cdot x_n \\ y_{n+1} &= \frac{1}{\sqrt{2}}\left(-(x_n - 1) + u(y_n - 1) + (x_n - 1)^2 - (x_n - 1)^3\right) + 1 - 2(y_n - 1)^2 \end{aligned} \tag{4.3}$$

with a Gaussian white noise process ω_{noise} with the Gaussian distribution $\mathcal{N}(\mu = 0, \sigma^2 = 0.0625)$. In the modified model the Hopf bifurcation is shifted to the critical control parameter $u = 1$ instead of $u = 0$ as in the original model in [KS04] and the equilibrium is determined by $(1,\ 1)$ which is also chosen as initial condition. The simulation is done with the method of subsection 3.4.3 for $\zeta = 0.01$ and the control parameter u is linearly increased in the interval $[0.25,\ 1.05]$ over 10000 time steps. The resulting dataset is shown in Figure 4.3.

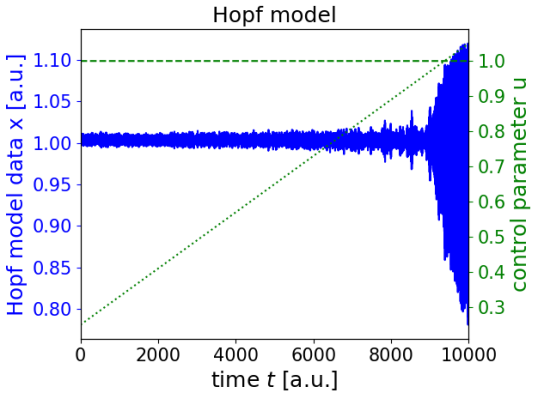


Figure 4.3: The time series generated by the Hopf model Equation 4.3 over 10^4 time steps with $\zeta = 0.01$, Gaussian white noise ω_{noise} with $\mathcal{N}(\mu = 0, \sigma^2 = 0.0625)$ and with the initial values $(x_0,\ y_0) = (1,\ 1)$ is shown by the blue solid line. At around 9300 time steps a Hopf bifurcation occurs. The corresponding threshold value of the control parameter $u = 1$, drawn as green dashed horizontal line, is reached around this time, shown by the crossing green dotted line.

4.1.4 Ecological resource model with critical slowing down

The stochastic differential equation

$$\mathrm{d}x = \left(rx \cdot \left(1 - \frac{x}{K}\right) - c\frac{x^2}{x^2 + h^2}\right)\mathrm{d}t + \sigma x \mathrm{d}W \tag{4.4}$$

describes the time evolution of some vegetation resource x that is harvested by herbivores. The biomass x grows logistically, represented by the first term of Equation 4.4, with the growth rate r, the population's carrying-capacity K, the half-saturation constant h and the grazing rate c.

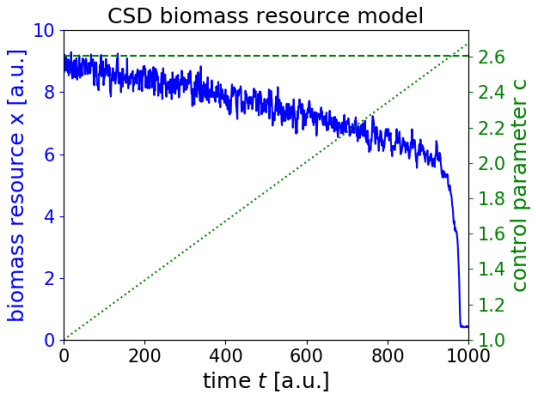


Figure 4.4: The time series generated by the differential Equation 4.4 over 1000 time steps with $r = 1$, $h = 1$, $K = 10$ and $\varsigma = 0.03$ and a standard normal distributed white noise process $\varsigma x \mathrm{d}W$ with the initial condition $x_0 = 9$ is shown by the blue solid line. After approximately 970 time steps the fold bifurcation occurs. That is the origin of overexploitation. The corresponding threshold value of the control parameter $c \approx 2.604$, drawn as green dashed horizontal line, is reached around the same time, illustrated by the crossing of the green dotted line.

The term dW is a Gaussian white noise process with intensity $(\sigma x)^2/\mathrm{d}t$. Under certain conditions the stable state of just slightly fluctuating biomass resource undergoes a fold bifurcation to a state of overexploitation. The critical transition takes place for $c \approx 2.604$ and $r = 1$, $h = 1$, $K = 10$ and $\varsigma = 0.03$. Following the descriptions in [Dak+12; CB11] the parameters are not chosen to investigate a specific configuration, but to simulate the fold bifurcation in conjunction with CSD for an exemplary investigation of the numerical methods on this dataset.
The blue curve of biomass resource in Figure 4.4 is computed for the initial condition $x_0 = 9$ by applying the Euler-Maruyama method, explained in subsection 3.4.1. The control parameter c is linearly increased from 1 to 2.6771 over a range of 1000 time steps. The green curve illustrates the linear increase of c, the green dashed horizontal line is given by the critical threshold $c \approx 2.604$ that leads to overexploitation.

4.1.5 Ecological resource model with flickering

For the simulation of the flickering dataset the Equation 4.4 is modified and coupled to a time-correlated noise process. This process can be modelled as

$$j_{n+1} = \left(1 - \frac{1}{T}\right) j_n + \beta\eta \tag{4.5}$$

and it is an expression for time-correlated variations of biomass described by red noise with the random numbers η of a Gaussian distribution with variance $\sigma^2 = \beta^2$. The simulation is done with the Euler-Maruyama method, described in subsection 3.4.1, in a time interval of $t \in [0, 10000]$ with a time step of $\mathrm{d}t = 0.05$. The red noise term is added to the computed system variable x in each time step and also scaled with the latter.

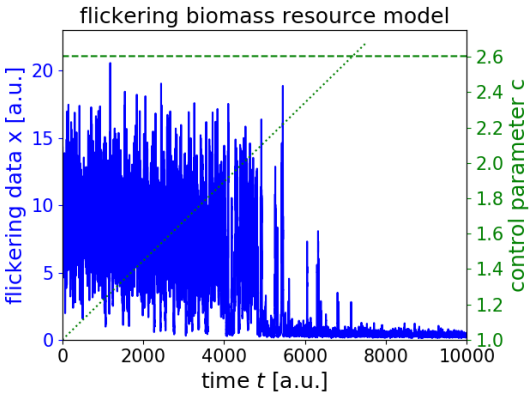


Figure 4.5: The flickering time series that is generated by the differential Equation 4.4 coupled to the red noise process of Equation 4.5, is shown by the blue solid line. It is simulated in the time interval $[0, 10000]$ with the time step $\mathrm{d}t = 0.05$ and $r = 1$, $h = 1$, $K = 10$, $\beta = 0.07$ and $\eta \in \mathcal{N}(\mu = 0, \sigma^2 = \beta^2)$ for the initial condition $x_0 = 9$ and $j_0 = 0.1$. The coexistence of two stable states leads to flickering in the interval of around 4000 to 7300 time steps, before the fold bifurcation to overexploitation occurs. The corresponding value of the control parameter $a \approx 2.604$, drawn as green dashed horizontal line, is reached around 7300, illustrated by the crossing of the green dotted line.

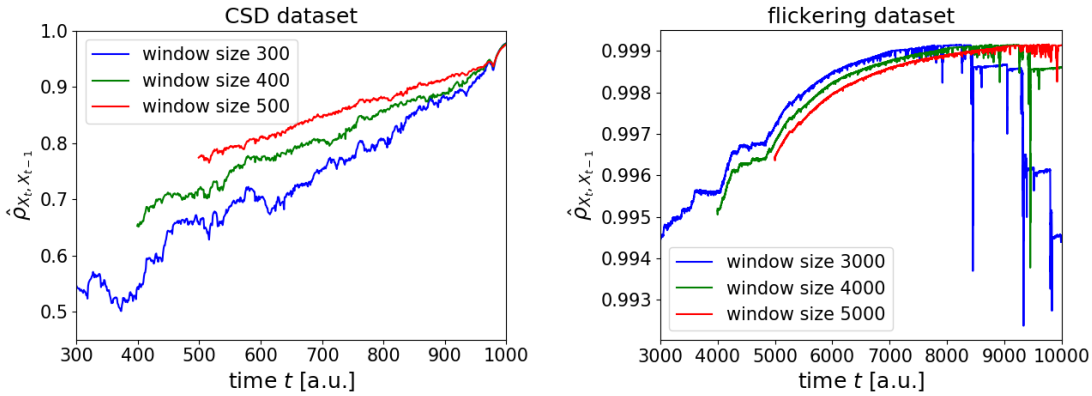
With the initial condition $x_0 = 9$ and $j_0 = 0.1$ and the parameters $T = 20$ and $\beta = 0.07$ the flickering dataset, visible in Figure 4.5, is simulated. Therefore, the unmentioned parameters are chosen analogously to the simulation in subsection 4.1.4 and the control parameter c is linearly increased in the interval $[1,\ 2.6771]$ over the time interval $t \in [0, 7500]$.

4.2 Statistical measures as leading indicators

The statistical measures and their time evolution are calculated for the simple CSD dataset and for the corresponding flickering version. The procedure is implemented using the rolling window method and the empirical estimation methods for the statistical measures, introduced in subsection 3.2.4 and subsection 3.2.3. The time windows are always shifted about one time sampling interval of the time series. The estimated leading indicators are synchronized to the last time window point. The functionality and the stability of the measures against varying reasonable window sizes is shown exemplarily for the CSD and the flickering dataset, before the methods are applied to the Ricker-type model, the Hénon model and the Hopf model in each subsection.

4.2.1 Autocorrelation

Based on the idea of critical slowing down the autocorrelation $\rho_{X_t,X_{t-1}}$ should increase over time for the simulated CSD dataset and the flickering dataset of the ecological resource model, described in subsection 4.1.4 and subsection 4.1.5. The expected positive trend is observable for both datasets, presented in Figure 4.6. The measures are evaluated for each



(a) Autocorrelation trend of the CSD dataset

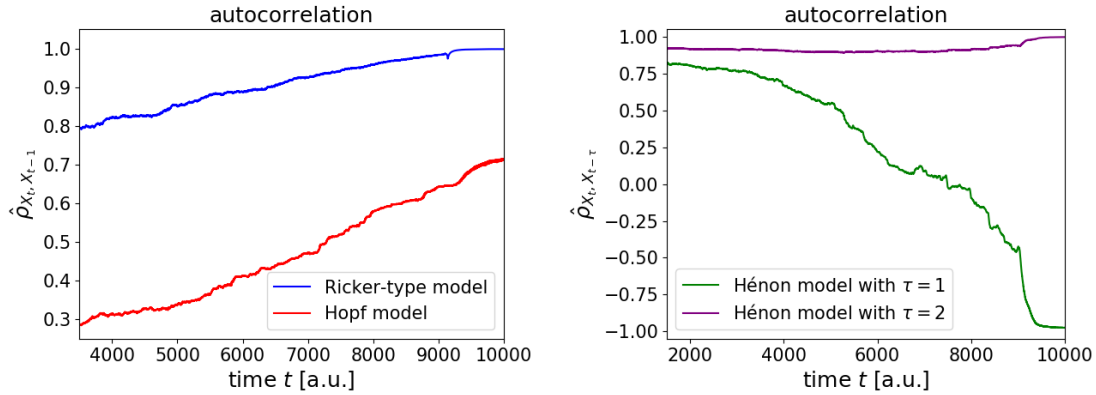
(b) Autocorrelation trend of the flickering dataset

Figure 4.6: The graphs show the estimated time evolution of the autocorrelation $\rho_{X_t,X_{t-1}}$ for the CSD dataset in 4.6(a) and for the flickering dataset in 4.6(b). The calculations are performed for various rolling time window sizes. The positive trend of the autocorrelation is observed for various window sizes in both systems. Bigger rolling time window sizes cause a decreasing interval of the observed trends, but the data is less volatile. The results for the flickering dataset are in general less volatile due to the higher time series sampling rate. The autocorrelation that is calculated with smaller time window sizes reacts faster to system variations as observable in 4.6(b). The dips in 4.6(b) are caused by widely scattered time series data that passes the time window boundaries.

dataset with three different rolling time window sizes. The lowest window size represents the amount of data that is at least needed to obtain a reasonable and stable estimation of the time evolution of the autocorrelation. Lower window sizes are highly volatile and cannot reproduce the time evolution or at least not clearly. The first suitable window size is often somewhat

around one third of the whole time series length. The highest chosen window size represents one half of the investigated time intervals. Higher values are less reasonable because of the implicit stationarity condition of subsection 3.2.4. If the control parameter and thus, the system's dynamics change over the whole time interval, as given in the simulated datasets, the approximation of stationary dynamics is better fulfilled for smaller windows. Furthermore, for practical reasons it is not desirable to use a method with bigger window sizes because of the needed amount of data and the highly delayed first estimation of the statistical measures here at time 500 for the CSD dataset and 5000 for the flickering dataset. The chosen rolling time window sizes of the two datasets are also comparable in their ratio to the whole time interval. In 4.6(a) the positive trend of the autocorrelation is clearly visible for the chosen rolling window sizes. With increasing window size the trend flattens a bit and is less volatile. The range of the time evolution for a window size of 300 is roughly 0.4, for a window size of 500 it is only 0.1. Nonetheless, the trend is clearly visible for the chosen window sizes.

4.6(b) shows slightly different results for the flickering dataset. A positive trend clearly



(a) Autocorrelation trend for the Ricker-type and the Hopf model

(b) Autocorrelation trend for the Hénon model

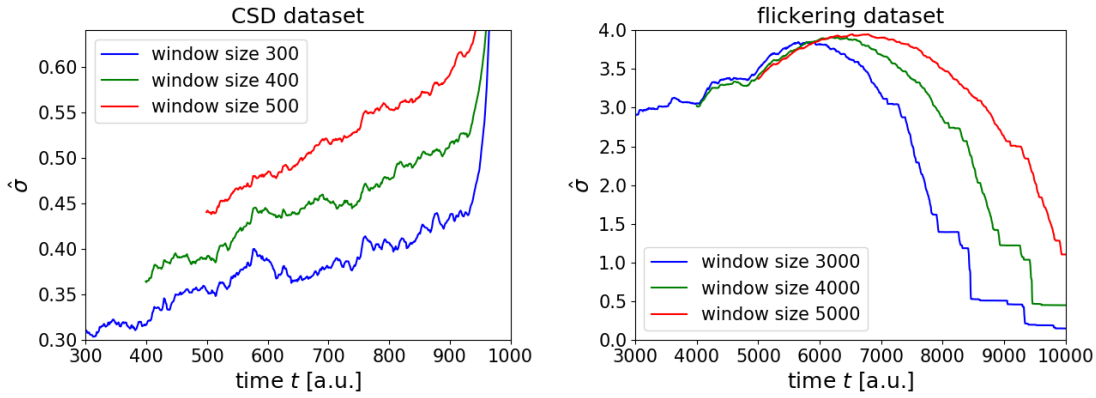
Figure 4.7: The graphs show the estimated time evolution of the autocorrelation $\rho_{X_t,X_{t-\tau}}$ for a set of three models, each of which undergoes a different type of local bifurcation. The rolling time window size is chosen to be 1500 time series sampling steps. There is a positive trend in the autocorrelation for the Ricker-type and the Hopf model in 4.7(a) that indicates an uprising bifurcation. The Hénon model exhibits a negative trend in autocorrelation for the time lag $\tau = 1$ and remains at unity for $\tau = 2$, visible in 4.7(b), because of the period doubling bifurcation. The autocorrelation trend does not suggest an uprising bifurcation in the case of the Hénon model.

exists, but the spanned range lies in the 10^{-3} scale. The graphs are less volatile than the CSD counterparts. This is probably a consequence of the much higher sampling rate of the flickering dataset. For the different window sizes there is no great difference in the prime development of the curves, but in the end the autocorrelation tends to decrease for smaller window sizes. The smaller window size leads to a faster reaction to the varying conditions as discussed in subsection 3.2.4. Therefore, the later negative trend is a result of the new unique stable state that is primarily competing with another one. The dips are artefacts that emerge

when a widely scattered value is shifted into or out of the rolling time windows.
The numerical procedure is exemplarily applied to a set of three systems using a suitable rolling time window size of 1500 time steps: the Ricker-type model, the Hénon model and the Hopf model, described in subsection 4.1.1, subsection 4.1.2 and subsection 4.1.3. The results are shown in Figure 4.7. The autocorrelation increases about a range of 0.3 for the Ricker-type model and 0.4 for the Hopf model indicating an uprising critical transition. It does not indicate a clear trend for the Hénon model, although it undergoes a local period doubling bifurcation as well. Note, that the autocorrelation for the Hénon map decreases for a lag of unity because of the period doubling bifurcation. After the period doubling the Hénon map without noise jumps between two values and the autocorrelation decreases to minus unity because one bigger value is followed by a smaller value in alternating manner of period two. For a time lag $\tau = 2$ the autocorrelation remains consequently at more or less unity because every even time lag correlates the bigger or smaller values of the period with themselves.

4.2.2 Standard deviation

The results for the estimated standard deviation $\hat{\sigma}$ are presented in Figure 4.8. For the CSD dataset the positive trends are clearly visible for all window sizes. The diverging part in the end corresponds to the occurring bifurcation at around time $t = 970$. For increasing window size the standard deviation is estimated generally higher and the evolution is less volatile.
For the flickering dataset the trend for a window of size 3000 is positive up to time $t = 6000$ and for the window of size 5000 up to time $t = 7000$ and over a range of unity. Afterwards a negative trend is visible due to the new unique stable state of the flickering dataset, similar to the effects observed for the autocorrelation estimation in subsection 4.2.1.



(a) Standard deviation trend for the CSD dataset

(b) Standard deviation trend for the flickering dataset

Figure 4.8: The graphs show the estimated time evolution of the standard deviation $\hat{\sigma}$ for the CSD dataset in 4.8(a) and for the flickering dataset in 4.8(b). The calculations are performed for various window sizes. The estimation process is generally less volatile for bigger rolling time windows, but the smaller rolling time windows react faster to the system changes in 4.8(b).

The standard deviation is analysed for the Ricker-type, the Hénon and the Hopf model for a window size of 3500 time steps. It is presented in Figure 4.9. The standard deviation exhibits a positive trend in a range of around 0.3 for the Ricker-type model, but shows no trend for the Hénon and the Hopf model, although they undergo a bifurcation as well.

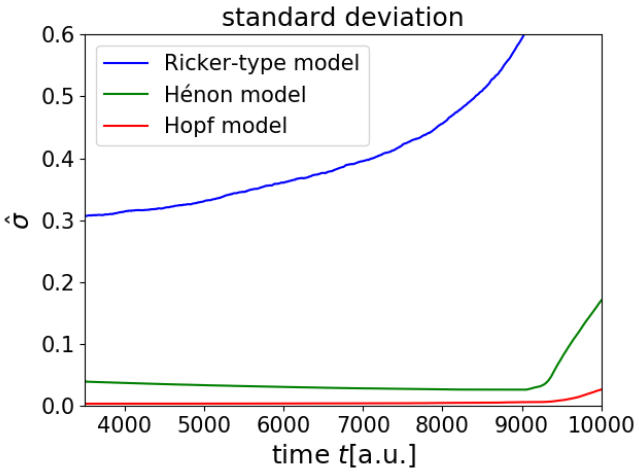
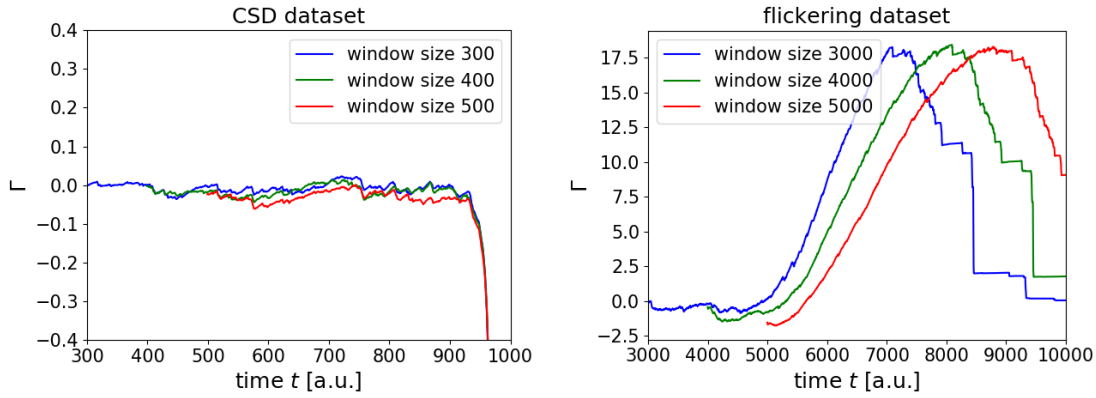


Figure 4.9: The figure shows the estimated time evolution of the standard deviation $\hat{\sigma}$ for a set of three models, each of which undergoes a different type of local bifurcation. The rolling time window size is chosen to be 3500 time series sampling steps. The standard deviation suggests an uprising critical transition detected by a positive trend only for the Ricker-type model .

4.2.3 Skewness

The skewness Γ is a statistical measure, especially useful for the detection of flickering, described in subsection 2.5.2. It is shown in Figure 4.10. The CSD dataset is not affected by flickering and thus, its probability distribution is symmetric. This leads to a skewness value of around zero. The negative asymmetry in the end results from the bifurcation at time $t = 970$ that gives rise to a left side asymmetry in the probability distribution for rolling time windows which include parts of the pre- and parts of the past-bifurcation regime.
The skewness is well detected for the flickering dataset in the region of the two competing fixed points. The probability distribution is asymmetric to the right side indicated by the positive skewness trend. It decreases just at the time when the previouly stable state disappears in a fold bifurcation. The slower reaction to dynamic changes for greater time windows is also well observed in Figure 4.10.

For the set of three models with different bifurcations the skewness remains almost zero as shown in Figure 4.11 for a time window size of 3500 time series sampling steps. The result suggests that there is no flickering occurring for the investigated models and thus, gives true information. However, a critical transition is also not indicated by the skewness Γ in these cases. Note, that the small downward trend for the Ricker-type model is in a scale of 10^{-2} and thus, does not provide clear information. The later strong downward trend results from the time windows that contain partly data of the pre- and post-bifuraction regime.

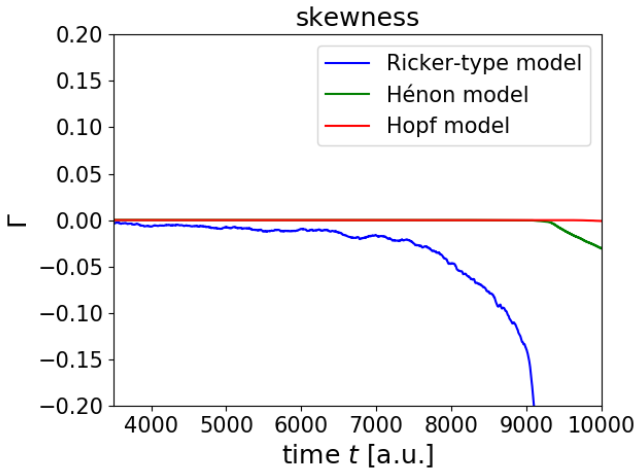


(a) Missing skewness trend for the CSD dataset

(b) Skewness trend for the flickering dataset

Figure 4.10: The graphs show the estimated time evolution of the skewness Γ for the CSD dataset in 4.10(a) and for the flickering dataset in 4.10(b). The calculations are performed for various window sizes. No skewness trend is observable for the CSD dataset because of the absence of flickering. The skewness trend caused by flickering is well observed for the flickering dataset. The slower reaction to dynamical variations of the estimated skewness with bigger time window size is also visible.

Figure 4.11: The figure shows the estimated time evolution of the skewness Γ for a set of three models, each of which undergoes a different type of local bifurcation. The rolling time window size is chosen to be 3500 time series sampling steps. The skewness does not suggest an uprising critical transition, although every model undergoes a critical transition. The skewness Γ gives valuable information about the absence of flickering in the investigated models.



4.3 The dominant eigenvalue as leading indicator

Apart from the commonly used leading indicators that are investigated in section 4.2, the following section deals with the properties of the dominant eigenvalue of a given complex system as leading indicator. The optimal choice of the numerical parameters, the robustness of the methods against numerical parameter errors and the predictive power of the dominant eigenvalue in comparison to the standard statistical measures are considered in detail.

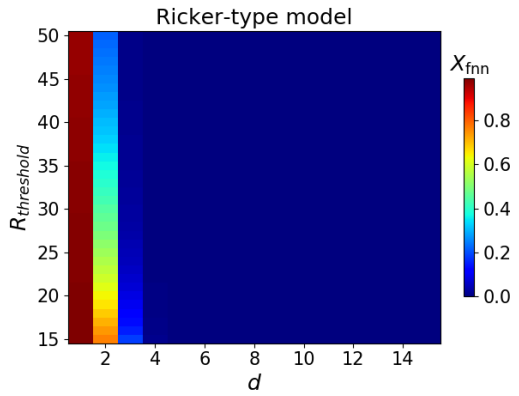
4.3.1 Optimal embedding parameters

To obtain a numerical estimation of the dominant eigenvalue, abbreviated DEV, of a given time series it is necessary to use Takens' embedding theorem in its practical version described in section 3.1. Reasonable methods in order to find an appropriate embedding dimension d and a suitable time lag τ are the false nearest neighbour, abbreviated fnn, algorithm for the embedding dimension d and the average distance from diagonal, abbreviated adfd, algorithm for the time lag τ. These methods are based on geometrical arguments and a detailed description can be found in subsection 3.1.1 and subsection 3.1.2. Both algorithms are applied to the set of three models, each of which undergoes a different kind of local bifurcation: the Ricker-type model, the Hénon model and the Hopf model.
In a first step the fnn algorithm is applied for a range of threshold distances $R_{\text{threshold}}$ from 15 to 50. The results underline the stability of the algorithm to declare false next neighbours over a wide range of critical distances. The calculations are performed for a range of embedding dimensions from 1 to 15. The results are presented in the colormaps 4.12(a), 4.12(c) and 4.12(e). For a sufficient high embedding dimension d no candidates for false nearest neighbours remain and thus, the second Heaviside function in Equation 3.1 remains zero. This function is also the denominator of Equation 3.1. For this reason, the mathematical error of division by zero is numerically caught by setting the false nearest neighbour statistics $X_{\text{fnn}} = 0$. For the Hénon model this can be observed from embedding dimension 7 to 15 in 4.12(c). For the Hopf model in 4.12(e) the false nearest neighbour ratio X_{fnn} of the orders 5 to 15 is set to zero.
For each system the lowest embedding dimension is chosen with a ratio X_{fnn} of false next neighbours of $X_{\text{fnn}} \lesssim 0.2$ with respect to most of the distance thresholds $R_{\text{threshold}}$. Following this approach an embedding dimension of $d_{\text{Ricker}} = 3$ is chosen for the Ricker-type model and an embedding dimension $d_{\text{Hopf}} = 4$ is chosen for the Hopf model. The embedding dimension[1] of the Hénon model is chosen to be $d_{\text{Hénon}} = 2$.
The time lag is calculated for the chosen embedding dimensions using the adfd algorithm with time lags τ beginning at 1 and ending at 25. The results are shown in the graphs 4.12(b), 4.12(d) and 4.12(f). For the Ricker-type model the algorithm reaches fast a plateau. Therefore, a time lag of $\tau_{\text{Ricker}} = 1$ is chosen. The Hopf model exhibits some kind of periodically seeming behaviour of the average displacement from diagonal depending on the time lag and shows its maximum value for the time lag $\tau_{\text{Hopf}} = 2$. This periodicity is stronger emphasized if the calculations are only performed in the post bifurcation regime, respectively for the time series data of times greater than approximately 9300. For the Hénon map the adfd approach suggests two possibilities: even or odd lags, respectively lag 1 or 2.[2] Note, that the average distance difference for these time lags is in the order of 10^{-2}, repeating periodically, and thus, there is no clear hint what is an appropriate time lag. Choosing $\tau_{\text{Hénon}} = 2$ for the S-map approach in subsection 4.3.7 gives reasonable results. The periodical trend of the time lag reflects the period doubling bifurcation as suggested in 4.12(d). The lower values correspond to the pre-

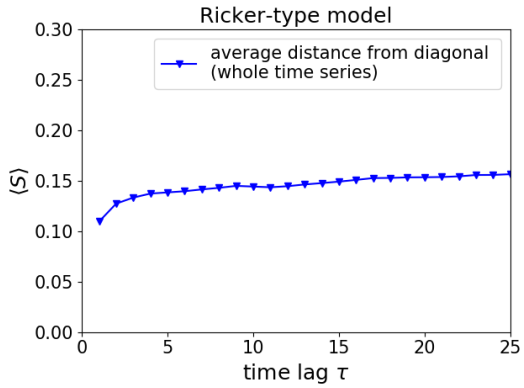
[1] The first analyses for the Hénon model are performed with a dimension of $d = 3$ chosen for the arguments mentioned above regarding the fnn calculations in 4.12(c). The results are even better for an embedding dimension of $d_{\text{Hénon}} = 2$, which is why the presented calculations in the following chapters are done with $d_{\text{Hénon}} = 2$. This is also the real dimension of the model and the fnn ratio X_{fnn} is still $X_{\text{fnn}} \lesssim 0.5$.

[2] The results for $d_{\text{Hénon}} = 2$ are similar to the initially chosen embedding dimension $d = 3$.

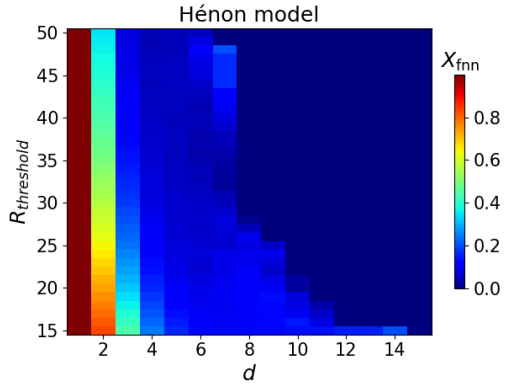
bifurcation regime. After the period doubling bifurcation the even time lags produce higher average distances from diagonal due to the two alternating values of the Hénon model, because the odd lagged vector components result in higher differences with respect to the non-lagged components. Although, a time lag of two exhibits reasonable results using the S-map approach for DEV estimation. Due to the fact that adjacent vectors with an even time lag are composed of only one of the two alternating fixed points in the post-bifurcation regime, the vector differences in this regime are higher. The changing dynamics could be better resolved in an embedding with $\tau_{\text{Hénon}} = 2$ because of similar adjacent vectors in the pre-bifurcation regime and stronger differing vectors short before and in the post-bifurcation regime. The time lags are used for the DEV analysis with S-maps as described in subsection 3.2.5. The DEV estimation method with autoregression AR(p) is performed with time lags of one time series sampling step $\tau = \Delta t$, because it yields reasonable results with less implementation effort.



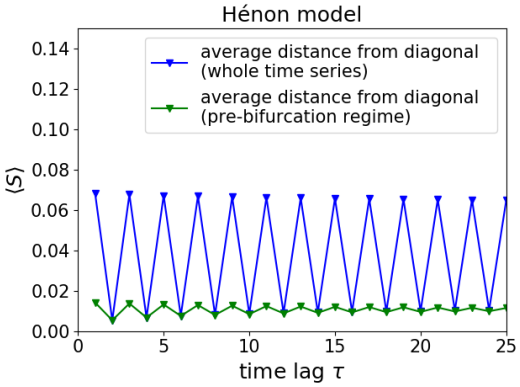
(a) False nearest neighbours for the Ricker-type model



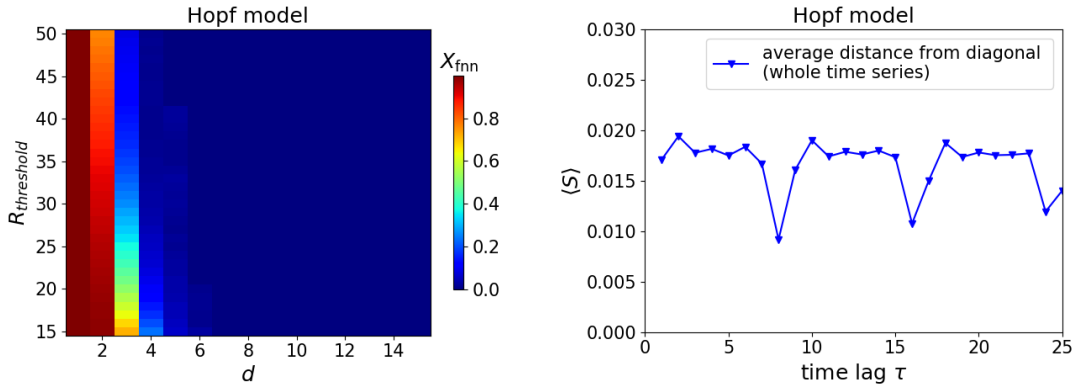
(b) Average distance from diagonal for the Ricker-type model



(c) False nearest neighbours for the Hénon model



(d) Average distance from diagonal for the Hénon model



(e) False nearest neighbours for the Hopf model

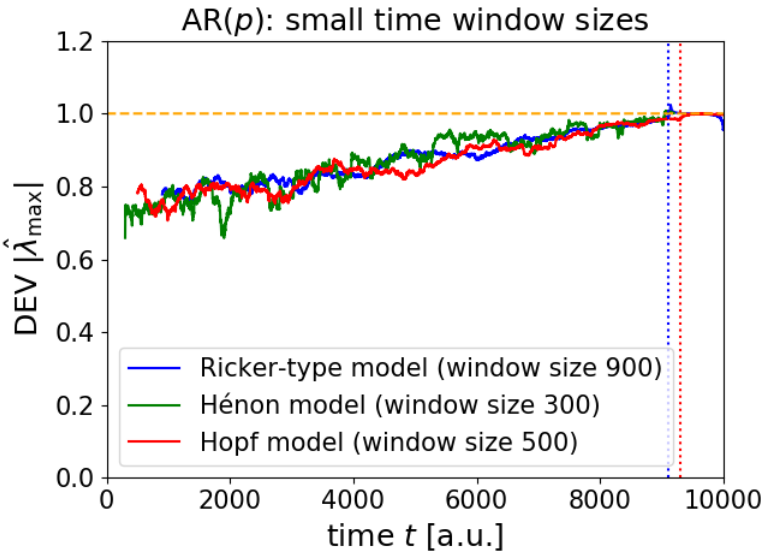
(f) Average distance from diagonal for the Hopf model

Figure 4.12: The colormaps in 4.12(a), 4.12(c) and 4.12(e) show the evolution of the false nearest neighbour ratio X_{fnn} in dependence of the embedding dimension d for a wide range of distance thresholds $R_{\text{threshold}}$. The first used embedding dimensions for the Ricker-type, the Hénon and the Hopf model are $d_{\text{Ricker}} = 3$, $d_{\text{Hénon}} = 3$ and $d_{\text{Hopf}} = 4$, whereas for even better results the dimension of the Hénon model embedding is corrected afterwards to $d_{\text{Hénon}} = 2$. The average distance from diagonal algorithm, evaluated in 4.12(b), 4.12(d) and 4.12(f), suggests time lags of $\tau_{\text{Ricker}} = 1$ and $\tau_{\text{Hopf}} = 2$. For a detailed discussion of the results and the chosen time lag $\tau_{\text{Hénon}} = 2$ of the Hénon model see in the running text. The determined time lags are used for the DEV S-map approach, described in subsection 3.2.5. The results of the DEV AR(p) approach, described in subsection 3.2.1, are reasonable with a time lag of one time series sampling step $\tau = \Delta t$.

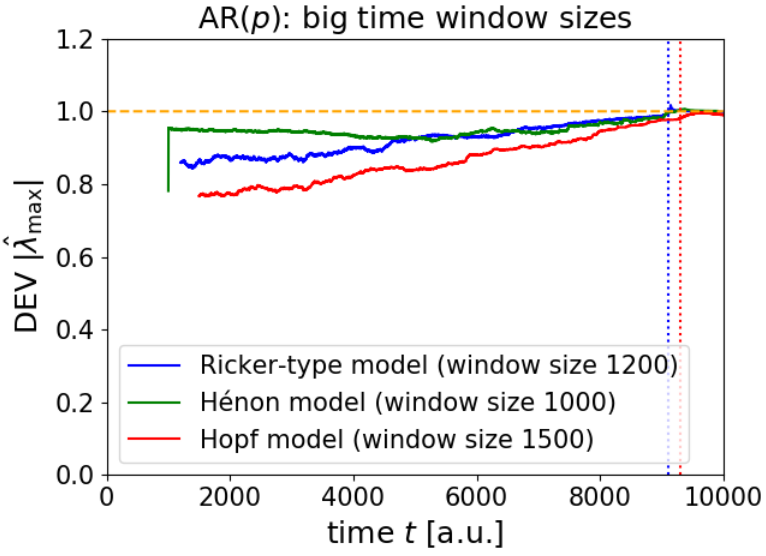
4.3.2 Absolute eigenvalue trend for various bifurcation types

The DEV is calculated for the Ricker-type model, the Hénon model and the Hopf model with the method, introduced in subsection 3.2.1. The order p of the AR(p) process is chosen equal to the embedding dimension d of the three models, specified in subsection 4.3.1. The autoregression is calculated with the pre-implemented python routine *ar_ model* of the *statsmodels.tsa* package with a fixed time lag of one time series sampling step for the regression scheme. The results are shown in Figure 4.13 for a respectively small and big window size. In both cases the positive trend of the eigenvalues indicates an uprising critical transition. It is clearly observable. The theory of linear stability for iterated maps, described in detail in subsection 2.3.1, states that the absolute value of the eigenvalue indicates instability by approaching a value of unity prior to a bifurcation. Around time 9100 for the Ricker-type and the Hénon model and around 9300 for the Hopf model a critical transition takes place, signed by the blue and red dotted vertical lines, and the DEV trend reaches approximately unity, marked by the orange dashed horizontal line. Thus, for a simple bifurcation induced by continuous parameter variation, the time of the critical transition can additionally be roughly estimated. A bigger window size provides similar results. The curves though, are less volatile. For a bigger window size the bifurcation time for the Hopf and the Hénon model is slightly later predicted and the Hopf model shows a little lower DEVs than for the smaller window size around the bifurcation point. In general, a bigger window size leads to a positive lift of the absolute values for the datasets.

The correlation between window size, volatility and reactivity to system changes, here observable in the shift of the bifurcation time for larger rolling time windows, is also observed for the simple statistical measures.
The window size in order to obtain reasonable results is generally much smaller for the Hénon model than for the other two models. For the Hénon model higher window sizes begin to lift



(a) Absolute value of the DEV for window size 900 , 300 and 500 for the Ricker-type, the Hopf model and the Hénon model



(b) Absolute value of the DEV for window size 1200, 1000 and 1500 for the Ricker-type, the Hopf model and the Hénon model

Figure 4.13: The two figures present the positive trend of the absolute dominant eigenvalues in time for three different models, each of which undergoes a different type of bifurcation. The evaluations are performed with a comparatively small rolling time window size in 4.13(a) and a bigger window size in 4.13(b). The dominant eigenvalue should reach approximately unity, signed by the orange dashed horizontal line, when the complex system exhibits a critical transition. The bifurcation occurs approximately at time step 9100 for the Ricker-type and the Hénon model, marked by the blue dotted vertical line, and around time step 9300 for the Hopf model, marked by the red dotted vertical line. The dip in the DEV curve of the Hénon model is a consequence of its initial transient.

the first half of the DEV curve and make the trend just visible in the other half as visible in 4.13(b). The dip in the DEV curve of the Hénon model is a consequence of its initial transient. A window, chosen too big, results in slightly decreasing eigenvalues at first, probably due to the decreasing shape of the Hénon model and its transient values in the beginning.

4.3.3 Robustness of the leading indicator

To investigate the robustness of the dominant eigenvalue $\hat{\lambda}_{\max}$ with respect to its ability to predict critical transitions, ROC curves for different time window sizes and different AR(p) orders p are evaluated. The ROC curves are constructed following the methodological instructions in subsection 3.3.2.
For reliable results it is convenient to investigate a time series with several bifurcations in order to improve the relative frequency of positive critical transition events. Therefore, similar to the data simulations in [ZKH15], the CSD dataset of biomass resource overexploitation, introduced in subsection 4.1.4, is simulated eleven times over 1000 time steps. After each 1000 time steps the control parameter is reset to its initial value and starts to change over time again. The dataset in the upper part of Figure 4.14 is computed with this simulation procedure. Every bifurcation from a finite biomass resource x to its vanishing is signed by the red dotted vertical lines.
The necessary binary tracer variable χ is shown in the middle graph of Figure 4.14. It is unity at times t at which should raise an early warning of a critical transition that will take place a lead time k later and zero otherwise. In Figure 4.14 the tracer variable χ for a lead time $k = 100$ is shown.
The lower graph in Figure 4.14 shows an indicator, here the dominant eigenvalue $\hat{\lambda}_{\max}$, calculated with a time window size of 350 time steps and the order $p = 3$ for the presented CSD time series. The calculations include results for time windows that contain data from the previous and the subsequent CSD dataset at the same time. These pointless dominant eigenvalues are marked by the shaded orange rectangles. They occur obviously after every CSD dataset in a range of one time window size. The indicator datasets are prepared for the ROC curve application by deleting these overlapping windows that would distort the statistics otherwise. In consequence, the tracer is also adjusted to the prepared indicators.
In this subsection the focus is on the robustness of the dominant eigenvalue estimation rather than its performance in comparison to the standard statistical measures. A detailed comparison can be found in subsection 4.3.4. The blue ROC curves in 4.15(a) to 4.15(c) show the prediction quality of a critical transition for the dominant eigenvalue as a leading indicator with time window sizes of 5, 10 and 20 and a fixed order $p = 3$. The lead time is fixed to $k = 5$ just before the bifurcation occurs. For very small time windows the classifier is not much better than a random classifier as seen in 4.15(a). This is due to missing data that would stabilize the dominant eigenvalue estimation. More data tends to stabilize the numerical procedure and allows for higher prediction quality as indicated by the ROC curves in 4.15(b) and 4.15(c) with the blue ROC curves approaching the upper left corner of the ROC plane. The trend of improved classifying quality with increasing window size is an expected result due to the need of a sufficient amount of data to stabilize every numerical estimation method for leading indicators [ZKH15].

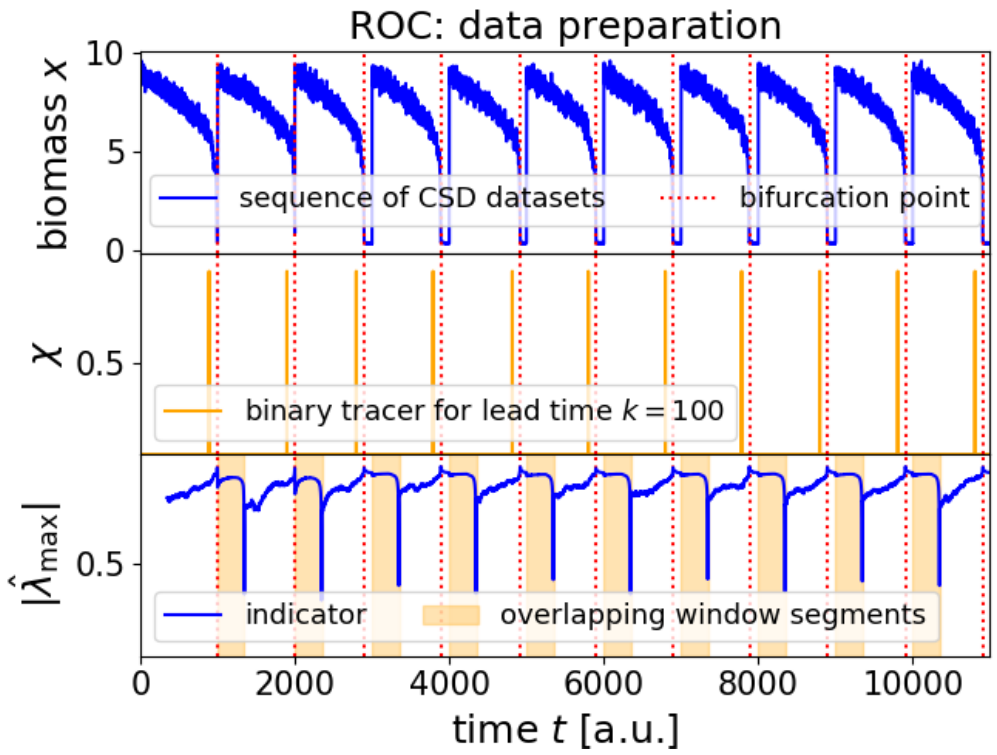
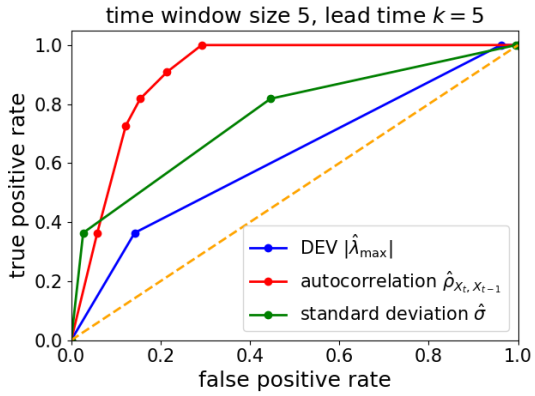


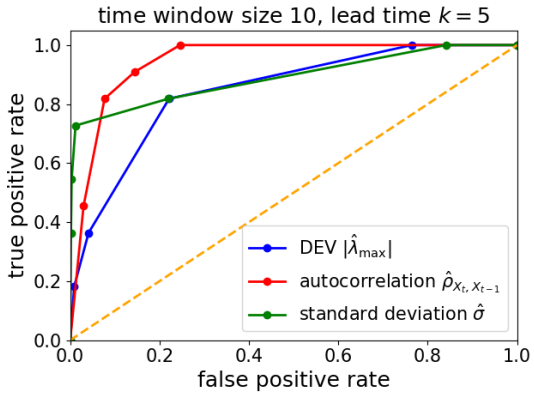
Figure 4.14: In the upper graph the investigated time series is shown. It is a sequence of eleven CSD datasets, simulated as described in subsection 4.1.4. The red dotted vertical lines indicate the time of bifurcation. The binary tracer variable is shown in the graph in the middle for a lead time of $k = 100$. It classifies the times t at which should be raised an early warning of a critical transition at $t+k$. The lower graph shows an indicator example: the dominant eigenvalue $\hat{\lambda}_{\max}$. After each single CSD dataset a time interval with the length of one rolling time window contains useless eigenvalues that are estimated from overlapping data of two subsequent CSD datasets. They are signed by the shaded orange rectangles. This overlapping segments are deleted and the tracer variables are adjusted to the prepared indicators before the ROC analysis is performed.

In 4.15(d) the robustness of the prediction quality against suboptimal AR(p) orders is shown. The time window size of 350 time steps is chosen in order to avoid instability due to a small amount of data per window. An optimal order[3] is around $p = 3$ for the CSD dataset. The prediction quality tends to decrease for higher orders. Although note, that this trend is just significantly visible for suboptimal orders p that are more than two times the optimal one. Even for order $p = 8$ the red ROC curve decreases just slightly compared to the blue optimal one. A significant change is just observed for very high orders, respectively $p = 15$ and $p = 50$. The results are visible for lead times $k \gtrsim 100$. Smaller lead times are less affected by a suboptimal order p.

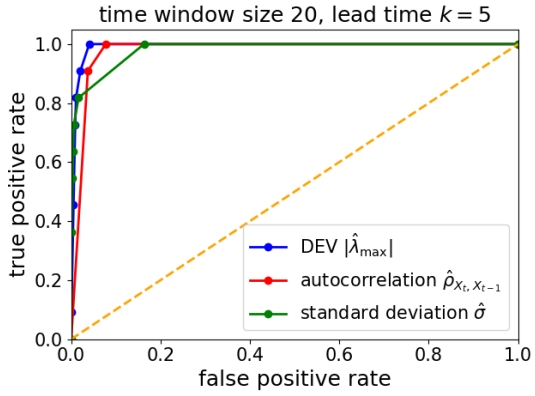
[3]The results of the fnn analysis to determine the optimal order p for the CSD dataset can be found in subsection A.2.1.



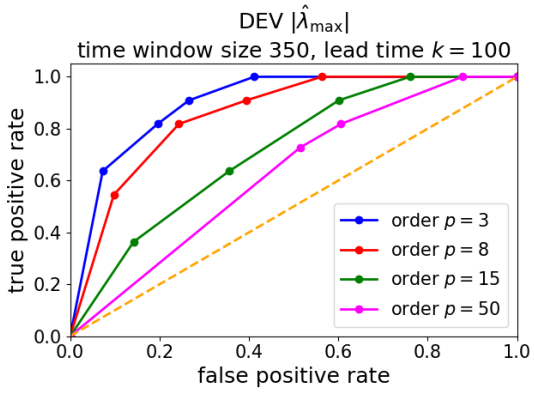
(a) Leading indicator stability for a window size of 5



(b) Leading indicator stability for a window size of 10



(c) Leading indicator stability for a window size of 20



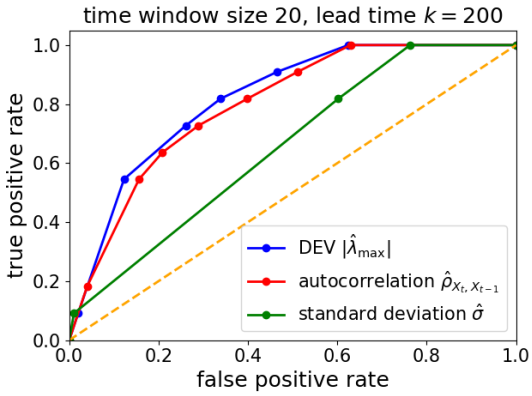
(d) DEV method stability with respect to order p of the AR(p)

Figure 4.15: In the graphs 4.15(a) to 4.15(c) the prediction quality is shown depending on different time window sizes with a fixed lead time $k = 5$ just before the bifurcation point. The analysis is performed for the dominant eigenvalues and the standard statistical measures of standard deviation and autocorrelation at a lag of 1, which are significant for the CSD dataset. In 4.15(d) the prediction quality in dependency of the AR(p) order p is presented. The trend is visible for lead times $k \gtrsim 100$. Smaller lead times are less affected by suboptimal orders p. The window size is chosen big with 350 time steps to avoid instabilities due to a small amount of data per time window.

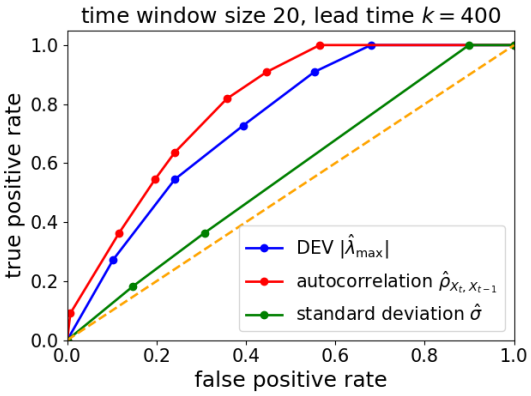
4.3.4 Comparison with the statistical measures

The dominant eigenvalue $\hat{\lambda}_{max}$ as leading indicator tends to need more data per rolling time window as the autocorrelation $\hat{\rho}_{X_t,X_{t-1}}$ and the standard deviation $\hat{\sigma}$ as visible in 4.15(a) to 4.15(c). For a time window size of 5 the dominant eigenvalue $\hat{\lambda}_{max}$ is just slightly better than a random classifier, whereas the standard deviation $\hat{\sigma}$ gives a little better results. The autocorrelation $\hat{\rho}_{X_t,X_{t-1}}$ is already for a very small window size a significant indicator for small lead times k. With increasing time window sizes the ROC curves of the three investigated

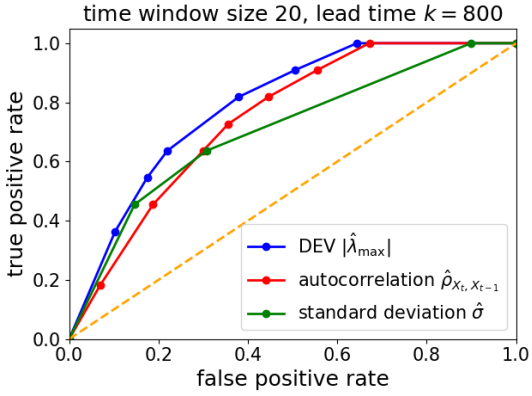
measures improve. For a window size of 10 the dominant eigenvalue $\hat{\lambda}_{\max}$ is still slightly worse than the standard deviation $\hat{\sigma}$. The autocorrelation ROC curve improves also a bit.
The ROC curve quality stabilizes for time window sizes equal or greater than 20 for small lead times k. The phenomenon that the estimation of the dominant eigenvalues $\hat{\lambda}_{\max}$ needs more data than the other two leading indicators, standard deviation $\hat{\sigma}$ and autocorrelation $\hat{\rho}_{X_t,X_{t-1}}$ respectively, should be explainable by the estimation methods: For the DEV estimation a number of values equal to the chosen order p has to be estimated from the time series data. Moreover, these estimated values are used for further calculations that increase the numerical uncertainties and errors. In contrast, the statistical measures are estimated directly.
The stable time window size of 20 time steps is used to investigate the prediction quality for various lead times k. The greater the lead time the earlier a warning of a critical transition in future can be raised. The calculations are performed for the three measures that are mentioned above. The results are presented in Figure 4.16. The standard deviation can be a very good classifier for very short lead times as observed in 4.15(c), but falls off in quality for higher lead times. The autocorrelation and the dominant eigenvalue are relatively good classifiers for an uprising critical transition also for higher lead times. As observable by comparison of 4.16(b) and 4.16(c) for higher lead times there is some fluctuation in the goodness of the ROC curves



(a) ROC lead time $k = 200$



(b) ROC lead time $k = 400$



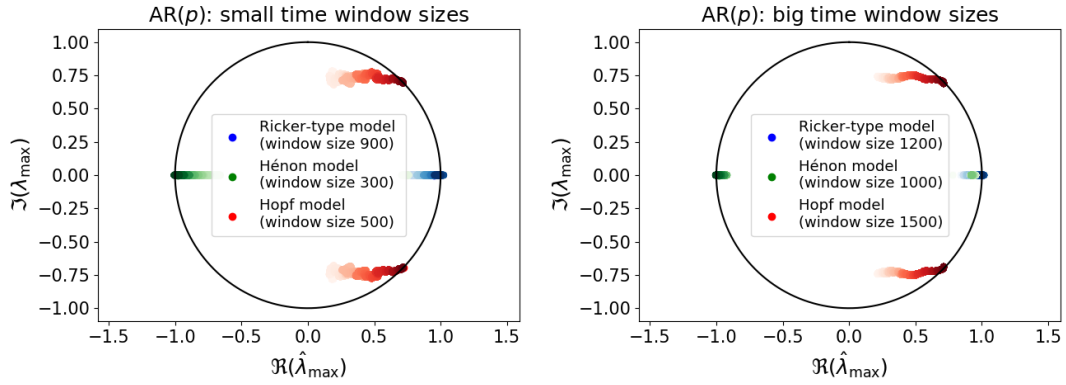
(c) ROC lead time $k = 800$

Figure 4.16: The predictive power of the dominant eigenvalue compared to the standard statistical measures for higher lead times, $k = 200$, $k = 400$ and $k = 800$ respectively, is shown in 4.16(a), 4.16(b) and 4.16(c). The standard deviation is a good leading indicator for smaller lead times, as shown in 4.15(c). For higher lead times the standard deviation falls off in quality, whereas the autocorrelation and the DEV are still good leading indicators. They stay in a comparable range of the ROC plane.

of the measures. Nonetheless, the autocorrelation and the dominant eigenvalue stay in the same quality range regarding the ROC curves, whereas the standard deviation just improves a bit. For higher time window sizes the described trend stabilizes and the fluctuations decrease significantly. The results for bigger time windows can be found in subsection A.1.2.

4.3.5 Classification of bifurcation types

The estimation of dominant eigenvalues with the method, introduced in subsection 3.2.1, serves also as classification parameter for various bifurcation types. The order p of the AR(p) process is chosen equal to the embedding dimension d of the three models, specified in subsection 4.3.1. The autoregression is calculated with the pre-implemented python routine *ar_ model* of the *statsmodels.tsa* package with a fixed time lag of unity for the regression scheme. The calculations are performed for the set of three models, whereby the Ricker-type model undergoes a fold bifurcation, the Hénon model exhibits a period doubling bifurcation and the Hopf model passes a Hopf bifurcation. The time evolution of the dominant eigenvalues in the complex plane is presented in Figure 4.17 for each of that models. The



(a) Complex DEV trend: window size 900, 300 and 500 for the Ricker-type, the Hénon and the Hopf model

(b) Complex DEV trend: window size 1200, 1000 and 1500 for the Ricker-type, the Hénon and the Hopf model

Figure 4.17: The estimated dominant eigenvalues for the three investigated models exhibit a trend in the Gaussian complex plane according to the theory of local bifurcation types, summarized in subsection 2.3.2. The color shading visualizes the time dimension. The color from light to dark indicates early and later calculated dominant eigenvalues from the data. This visualization fits the theory accurately and provides a possibility to classify the type of an uprising bifurcation. The Ricker-type model exhibits a fold bifurcation: The DEVs reach unity in the complex plane. The Hénon model undergoes a period doubling bifurcation: The DEVs reach minus unity. The Hopf model has complex conjugated eigenvalues with an absolute value of unity when the Hopf bifurcation occurs.

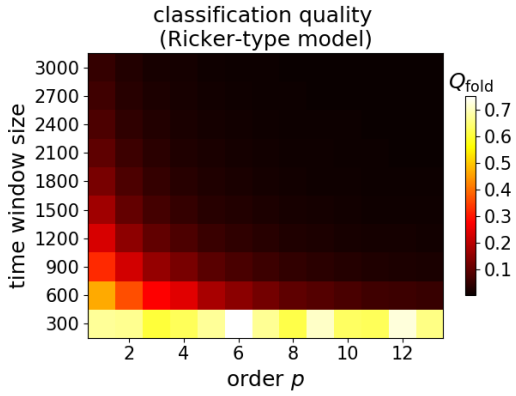
small and big window sizes are chosen in analogy to subsection 4.3.2. The time dimension is resolved by a color shading from light to dark corresponding to a time evolution from early to later calculated dominant eigenvalues. For both, the small and the big chosen windows,

the classification of the different local bifurcation types is possible via the theory of local bifuraction types, summarized in subsection 2.3.2. As stated in theory the Ricker-type model has purely real eigenvalues, signed in blue, that reach unity and the Hénon model has purely real eigenvalues that approach minus unity at the bifurcation point. The Hopf model exhibits complex conjugated eigenvalues that reach unity in their absolute values at the bifurcation. For higher window sizes less scattering in the Gaussian plane is found as long as the numerics remain stable for the other parameters. Bigger windows lead to higher absolute values of the dominant eigenvalue of the three models, as noted in Figure 4.13. Therefore, bigger window sizes correspond to eigenvalues which start already in higher vicinity to the unit circle in the Gaussian plane.
The flip bifurcation of the Hénon model is well classified with the green eigenvalues. Nevertheless, the first eigenvalues start in vicinity of unity. This should not result in a classification problem because they are just the earliest calculated eigenvalues and the general trend leads to minus unity. This undesired scattering towards plus unity is emphasized for bigger time windows. Very small time windows and cancelling out the initial transient of the data widely solve this scattering phenomenon and thus, almost all estimated eigenvalues of the Hénon model start for a window size of 100 time steps at the negative real axis of the complex plane.

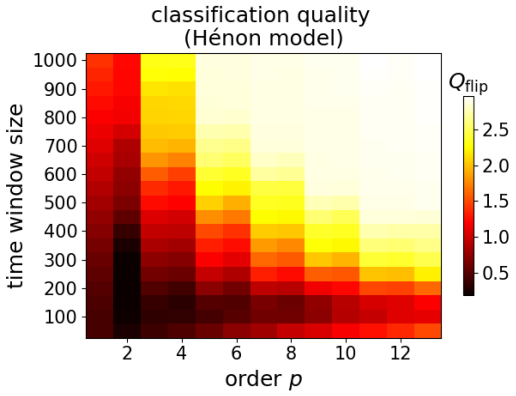
4.3.6 Robustness of the bifurcation classification

The classification of bifurcation types, presented in subsection 4.3.5, is investigated regarding the robustness of the estimation method towards window sizes and orders p of the AR(p) process. For a systematic investigation the quality measures, presented in subsection 3.3.1, are calculated for each of the three models, investigated in subsection 4.3.5, for a wide range of time window sizes and orders p including the chosen parameter values of subsection 4.3.5. The smaller the quality measure value the better the parameter range. The results are shown in Figure 4.18. In 4.18(a) the colormap, calculated with $\gamma = 1$, indicates some unstable scattering in the complex plane for time window sizes around 300 time steps. This is actually found as visible in 4.19(a) and discussed in detail in subsection 4.3.7. For sufficient great time windows the method is stable and yields more eigenvalues in the vicinity of unity with higher window size and higher order p. This trend leads to very small values of the quality measure and indicates that almost all eigenvalues are very close to unity. Thus, the time evolution is less resolved for very big time window sizes and high orders p. This corresponds directly to the observation that the whole absolute value curve of the eigenvalues is positively shifted for these parameter changes. A good trade-off between both, the resolving problem for big time window sizes and orders p and the stability problem for small time window sizes is found in a small parameter range in the lower left corner. The time window size of 900 and $p = 3$, used in subsection 4.3.5, yield a quality measure of $Q_{\text{fold}} \approx 0.25$ whereat the maximum value is around 0.8. That is a good quality value for identifying the fold bifurcation of the Ricker-type model.
As noticed in earlier analyses, the flip bifurcation of the Hénon model is identified better for respectively small time windows and a small order p as visible in 4.18(b), calculated with $\gamma = 1$. The almost white upper right corner represents the phenomenon of eigenvalues that approach unity and not minus unity, which is emphasized for big time window sizes. The quality measure

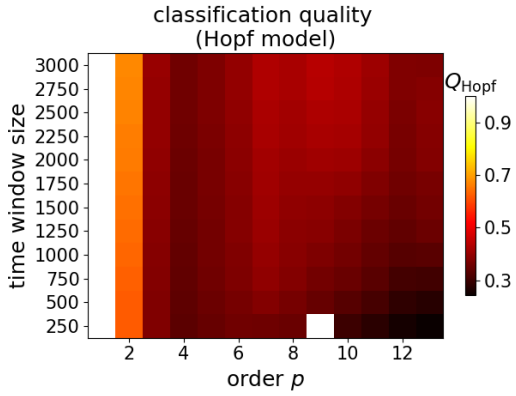
for the Hénon model yields approximately $Q_{\text{flip}} \approx 0.6$ for a time window size of 300 time steps and an order $p = 2$, which is chosen for the calculations in subsection 4.3.5. In comparison to the maximum value of around 3, that is a very good parameter choice. Furthermore, note that the order $p = 2$, which is chosen for the analysis, is extraordinarily good for the classification quality.
The Hopf model does not show a strong dependency of the time window size as visible in 4.18(c), calculated with $\gamma = 5$ for a better colormap resolution. It is more sensitive to variations in the order p and provides better quality results for higher orders. The very low quality values in the lower right corner for high orders p, respectively, correspond to some instability artefact: The model order is too high for such small amount of data per window and yields early estimated scattering values all around the unit circle with absolute values slightly smaller than unity. Nonetheless, after the first few eigenvalues scatter around the unit circle, the desired trend of complex conjugated eigenvalues approaching the unit circle, is observable. The



(a) Classification quality for the Ricker-type model



(b) Classification quality for the Hénon model



(c) Classification quality for the Hénon model

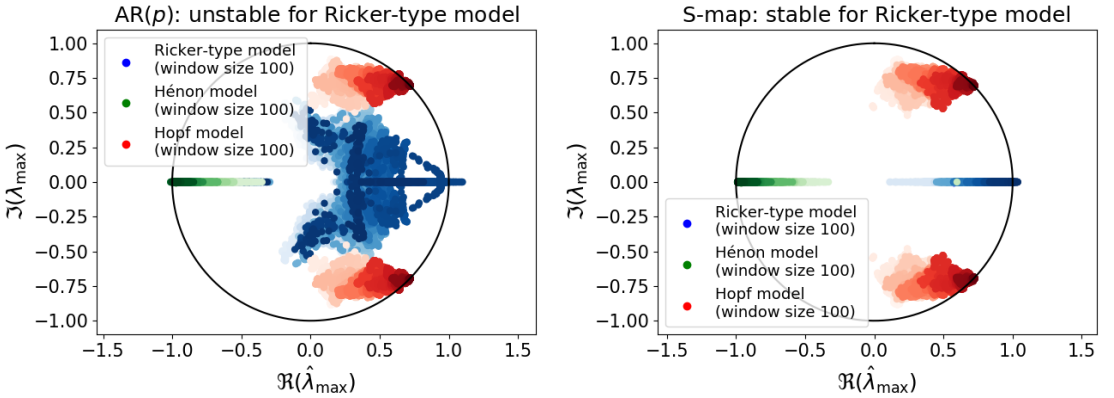
Figure 4.18: The figures show the quality measures for the bifurcation type classification of the three investigated models for a wide range of time window sizes and AR(p) orders. The quality measures, defined in subsection 3.3.1, can be interpreted as average distance of the eigenvalues to the final values of the classifying trend weighted with proceeding time. Thus, they are a measure of scattering effects in the complex plane which do not represent the desired classification trend. Generally speaking, the smaller the quality measure the better the bifurcation type classification. The 4.18(a) and 4.18(b) are calculated with a time weight factor $\gamma = 1$, the calculations for the Hopf model were performed with $\gamma = 5$ for better distinguishing of the colormap range. Further details of the interpretation can be found in the running text.

classification is therefore still possible. A clear instability is also denoted for the white field of window size 300 and $p = 9$. Note, that for a chosen order $p = 4$ that corresponds to the optimal embedding dimension d, calculated in subsection 4.3.1, the quality measure is better than for all other orders p, except for the unstable area in the lower right corner. In the end, the order $p = 4$ together with a window size of 500, used in subsection 4.3.5, yield a quality measure of $Q_{\text{Hopf}} \approx 0.3$, whereat the maximum value is around unity.

4.3.7 Comparison between AR(p) and S-map eigenvalue estimation

The calculations of the dominant eigenvalues, their absolute value trend and their visualisation in the complex plane is repeated for each of the three models, investigated in subsection 4.3.2, by estimation with the S-map method. The embedding dimension d and the time lag τ are chosen following the optimisation process in subsection 4.3.1. The window size is chosen in analogy to the previous subsections to simplify the comparison. The nonlinearity factor θ has no significant influence on the results of the three models and is therefore chosen low as $\theta = 0.2$. The implementation of the S-map approach follows the instructions of item 2 in subsection 3.2.5. This saves computation time and gives very similar results to the much more time consuming calculations with the implementation of item 1. The S-map calculations are performed for an equilibrium average state space vector in each single time window for all available next neighbours of a given window size.
The results are widely the same as for the AR(p) method: The absolute value trend is very similar to the results for AR(p). The trends are in the same range and have roughly the same shape. For higher window sizes there occurs also a slight shift of the bifurcation time prediction of the Hopf model and its eigenvalues are also slightly lower as in the AR(p) estimated data. There is less volatility in the Ricker-type eigenvalue trend for the small window sizes.
The classification of bifurcation types in the Gaussian plane yields the same results as its AR(p) counterparts. The classification is clear and for bigger time windows the scattering in the plane decreases, while the absolute values are in higher vicinity to the unit circle at the same time. For the Hénon model a few early estimated eigenvalues near positive unity are also observed. For a detailed presentation of the data see subsection A.1.1.
The main difference is the stability for the Ricker-type model at small window sizes as seen in Figure 4.19. Both, the AR(p) and the S-map estimation yield stable results with moderate scattering for the Hénon and the Hopf model, but the eigenvalues of the Ricker-type model begin to scatter widely and develop already imaginary conjugates for window sizes of around 800 points and smaller if they are estimated with AR(p). For window size 800 the scattering DEV values look similar to the two lately estimated complex conjugated branches in the form of a horseshoe in 4.19(a). For even smaller rolling time window sizes, as visible in 4.19(a), the unstable scattering trend increases significantly.
In conclusion, the S-map estimation is definitely more stable for smaller window sizes and exhibits less volatility for the absolute values of the Ricker-type model's DEVs. This should be rather obvious because the S-map algorithm works on the reconstructed state space vectors in each time window. Therefore, more data per window is available for the S-map algorithm than for the AR(p) algorithm due to the pre-processing of the time delay embedding of the data. Thus, if very small time windows are necessary for any practical reason, the S-map

approach could turn out to be the better choice. Apart from that the results are almost identical to the results computed with AR(p) estimation. A striking point is the simplicity of the AR(p) algorithm itself. The results of the S-map method depend on the embedding dimension d, the time lag τ and the nonlinearity factor θ. Even if θ does not play an important role in the investigated datasets, the results are just reproducible for a small range or even one particular time lag as in the case of the Hénon model. For these reasons and, additionally, because of the higher implementation effort and the higher computation costs for bigger window sizes, the S-map method is not applied to the empirical data of section 4.4.



(a) Unstable Ricker-type eigenvalues for small window sizes with AR(p)

(b) Stable Ricker-type eigenvalues for small window sizes with S-maps

Figure 4.19: The figure compares the results of the eigenvalue estimation in the complex plane for the AR(p) method and the S-map method at small window sizes of 100 time series sampling steps. The Hopf and the Hénon model are stable for very small window sizes of both methods. They exhibit moderate scattering and a clear trend without numerical artefacts. The Ricker-type model, however, shows an unstable eigenvalue estimation with the AR(p) method for a time window of 100 points, whereas the S-map estimation method stays stable even for this window size. The AR(p) does not stabilize the eigenvalue estimation of the Ricker-type model before time window sizes of around 900 time series sampling steps.

4.4 Application of the AR(p) eigenvalue estimation to empirical data

In this section the AR(p) eigenvalue estimation method, described in subsection 3.2.1, is applied to various real times series data in order to investigate its predictive ability in practical applications. Furthermore, the real world time series data of subsection 4.4.1 and subsection 4.4.2 are known to undergo a fold and a Hopf bifurcation, respectively. Thus, it is tried to classify the bifurcation types of the empirical time series via the DEV estimation approach.

4.4.1 Microcosm experiment

In [Ver+11] a *microcosm experiment*[4] that undergoes a fold bifurcation is presented. Various early warning measures as the autocorrelation, the recovery rate and the variance are tested in [Ver+11] and can identify the uprising critical transition.

The investigated system is a population of cyanobacteria, precisely *aphanizomenonflos-aquae*, that is cultured in a chemostat microcosm. The population is exposed day per day to a slightly increasing light level. The cyanobacteria culture needs light for photosynthesis and growth in a specified range. If the light exposure exceeds the upper limit, the light inhibits the growth of the bacteria population. If the bacteria population is dense enough, it can provide the necessary shading for its own growth and compensate negative light effects. This leads to a positive feedback of the population growth. If the light exposure is further increased, the light overcompensates the shading effect, the population cannot grow any more and every loss of biomass further decreases the shading: The system destabilizes in a fold bifurcation.

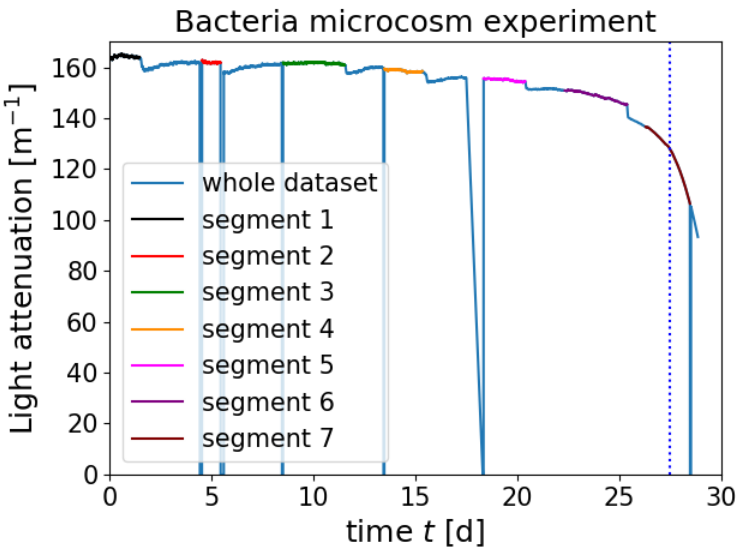


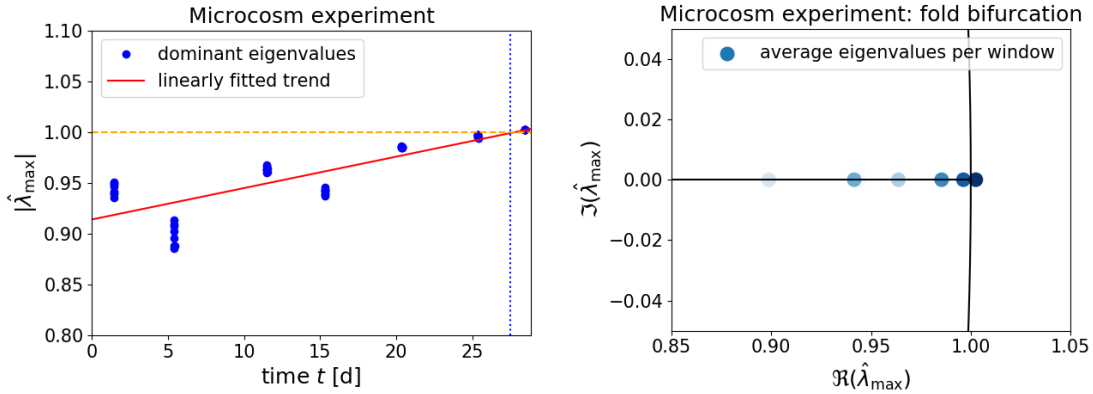
Figure 4.20: The light attenuation coefficient that is correlated with the cyanobacteria population is shown. The last 257 data points of every colour marked data segment are used for the analysis. The sudden decrease of population around 3 to 5 % after each data segement are due to dilution disturbance.

Every four or five days the cyanobacteria culture is perturbed by flushing with medium that reduces the biomass around 3 to 5 %. These perturbations are needed in order to study the response of the ecological system in terms of leading indicators. The dataset of the microcosm experiment is presented in Figure 4.20. The light attenuation coefficient correlates with the cyanobacteria density. Each of the coloured segments represents the data before the population is perturbed. The analysis is performed with the last 257 data points of each data segment. The dominant eigenvalue is estimated in time windows of 250 time steps with an AR of the order[5] $p = 3$ for the best results. The results, presented in 4.21(a), are robust against various

[4]The experimental data is provided by Prof. Dr. Arndt Telschow from the University Osnabrück, institute for mathematics and informatics (Barbarastr. 12, 49076 Osnabrück)(arndt.telschow@uos.de).

[5]The optimal order is determined by the false nearest neighbour algorithm. The results can be found in subsection A.2.2.

window sizes and orders p. For each slice of 257 data points the eight absolute values of the dominant eigenvalues of the time windows are shown in blue. Thus, a linearly fitted trend with the computed eigenvalues is equally weighted at each slice of data. The linear trend does not represent the actual evolution of the eigenvalues in time. This evolution is unknown. The linear trend is only a tool to illustrate the positive evolution of the absolute values of the dominant eigenvalue. Moreover, the linear trend can give a rough estimate of the time at which the bifurcation occurs. But note, that one has to be careful with the time estimate because it will just be valid if nothing else in the parameter evolution of the system changes significantly. For example, if in a simulated dataset the control parameter is increased linearly in the first half of the simulation and in an exponential manner later on, an extrapolation of the eigenvalue trend based on the first half of the dataset will doubtless lead to a wrong bifurcation time prediction.
In 4.21(b) the dominant eigenvalues are shown in the complex plane. For every slice of data the eight eigenvalues are averaged and drawn into the Gaussian plane. The shading from light to dark corresponds to the time evolution from early to later eigenvalues. The trend in the complex plane gives reasonable results for the expected fold bifurcation of the cyanobacteria population system. The second and third point do not fully agree with the expected trend towards unity, but the overall trend is clearly visible. If one uses the full data segments, marked with different colours in Figure 4.20, the trend in the complex plane is completely correct. Nonetheless, the analysis is performed with equally sized data segments in order to avoid unequal weighting of the various data segments.



(a) Microcosm experiment: dominant eigenvalue trend

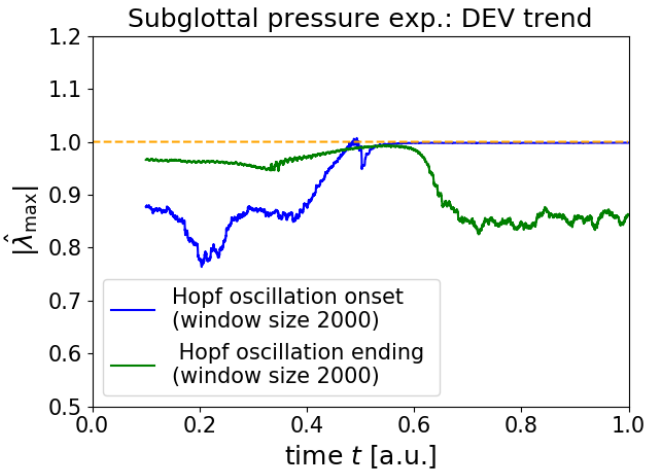
(b) Microcosm experiment: bifurcation type

Figure 4.21: In 4.21(a) the absolute value trend of the dominant eigenvalues for the last 257 data points of the data segments in Figure 4.20 are shown. The calculations are performed with an AR($p = 3$) scheme. The data is linearly fitted to illustrate the positive trend. In 4.21(b) the dominant eigenvalue trend is shown in the complex plane. The eigenvalues are averaged for each window. The pure real eigenvalues approach unity from early to later time segments, indicated by the shading from light to dark.

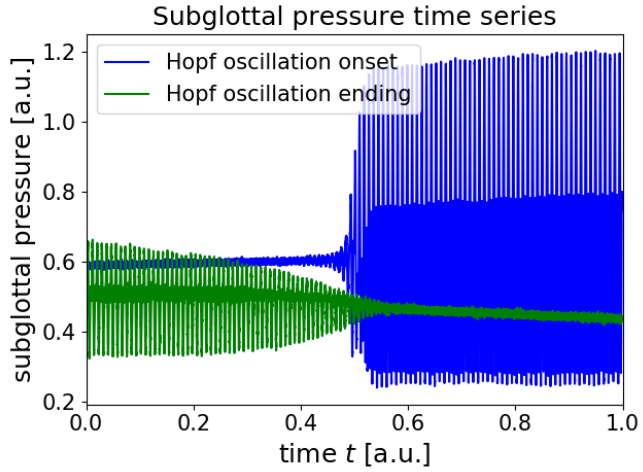
4.4.2 Physiological Hopf experiment

The time series of the experiment[6] for the onset of the Hopf bifurcation and the opposite case, shown in 4.22(b), are investigated with rolling time windows of 2000 points and an AR of the order[7] $p = 5$. The results for the absolute values of the dominant eigenvalue are shown in 4.22(a). The absolute values of the dominant eigenvalues, marked by the blue curve, stay around 0.87 for the onset data, before they begin to increase just before the Hopf bifurcation. For the subsiding oscillations of the second dataset, marked in green, the eigenvalues are approximately of absolute value unity and decrease, indicating that the oscillations will disappear. The reactions of the dominant eigenvalue trend is rather close to the onset of the bifurcation. Thus, the DEV would give reliable early warnings just with short lead times.

The eigenvalues in the Gaussian plane, shown in Figure 4.23, can identify the expected Hopf bifurcation of the experiment. The onset time series corresponds to the blue dotted eigenvalues in the complex plane. Their time evolution is illustrated by the shading from light to dark coloured points corresponding to eigenvalues early and later in time. The first eigenvalue estimates



(a) Physiological Hopf experiment: dominant eigenvalue trend



(b) Physiological Hopf experiment: time series

Figure 4.22: In 4.22(a) the absolute value trend of the dominant eigenvalues due to a Hopf bifurcation is shown for the onset of the oscillations in blue and the disappearing oscillations in green with a window size of 2000 data points. The corresponding datasets are presented in 4.22(b).

[6]The experiment deals with physiological properties of articulation and speech sound production. More precise information is not available due to an receivable publication of the experiment. The experiment is performed by Prof. Dr. Isao Tokuda from the Ritsumeikan University (1-1-1 Nojihigashi, Kusatsu, Shiga 525-8577, Japan)(isao@fc.ritsumei.ac.jp). The data is provided by Prof. Dr. Arndt Telschow from the University Osnabrück, institute for mathematics and informatics (Barbarastr. 12, 49076 Osnabrück)(arndt.telschow@uos.de).

[7]The optimal order is determined by the false next neighbour algorithm. The results can be found in subsection A.2.2.

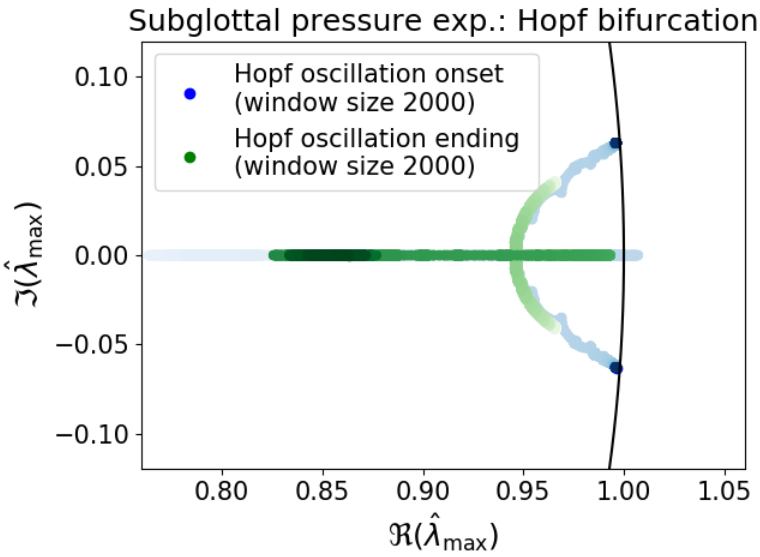


Figure 4.23: The dominant eigenvalue trend in the complex plane is shown for both cases in the same colour coding: The shading from light to dark represents the time evolution from early to later estimated eigenvalues. The typical complex conjugated eigenvalues are observable for both cases. For the onset of the Hopf bifurcation, they are purely real at first, but become imaginary after a few time steps. The complex conjugated eigenvalues approach the unit circle as expected. In the opposite case, marked in green, the complex eigenvalues collapse to purely real eigenvalues and decrease further in magnitude.

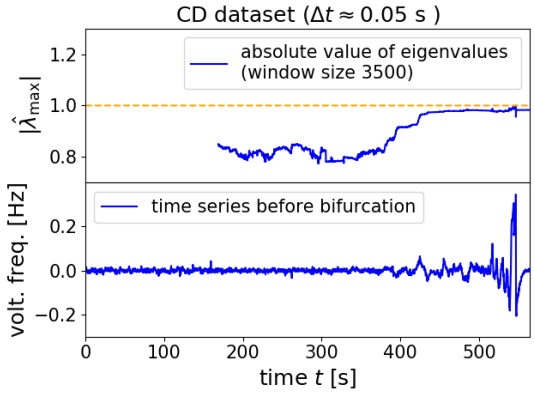
are purely real and not typical for a Hopf bifurcation. After a short time they evolve into complex conjugates and go fast towards the unit circle. This trend indicates the expected Hopf bifurcation in the onset case. The eigenvalues for the second dataset with the disappearing oscillations are shown as green dots. The time evolution is encoded as stated above. The first eigenvalues are complex conjugates as expected for the oscillating behaviour. With increasing time the imaginary parts collapse and the eigenvalues become purely real and decrease further in their absolute value. This is exactly the opposite behaviour in comparison to the eigenvalues for the onset dataset. This way, the direction of the Hopf bifurcation, in other words, whether the oscillations start or disappear, seems to be detectable by the eigenvalue trend in the complex plane.

4.4.3 Power outage data

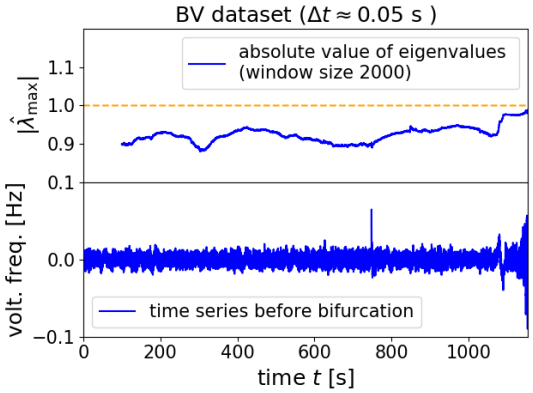
In this subsection two time series of the "1996 Western North America blackouts" are analysed. One dataset was first analysed by Christopher M. Danforth in [CHD12] and provided by him for further studies. This dataset is named the "CD dataset" in this subsection. A second time series that is part of the same power outage event is obtained from the Bonneville Power administration and is called the "BV dataset" from now on.
Both, the CD and the BV dataset yield the time evolution of the bus voltage frequencies until the point of the blackout. At this point a critical transition is assumed. Their analysed detrended versions are shown in the lower graphs of 4.24(a) and 4.24(b). The BV dataset contains a longer time interval than the CD dataset prior to the uprising blackout, but it is less densely sampled. The sampling of the BV dataset is $\Delta t_{\mathrm{BV}} \approx 0.05$ and the sampling of the CD dataset is $\Delta t \approx 0.025$. The sampling of the CD dataset is artificially changed to a less dense sampling of $\Delta t_{\mathrm{CD}} \approx 0.05$ in order to create comparable results. This way, one can exclude that

numerical sampling issues cause different results for the datasets.

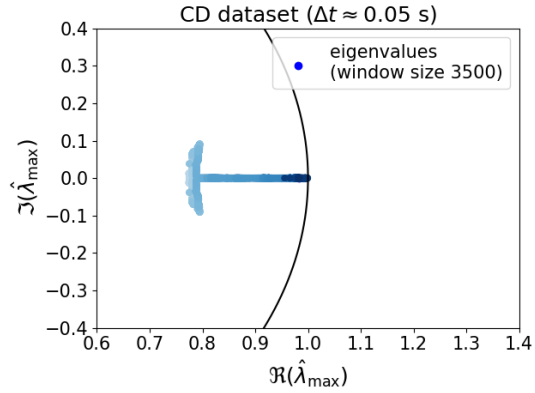
Both, the CD and the BV dataset are investigated with AR($p = 4$)[8] and a time window size



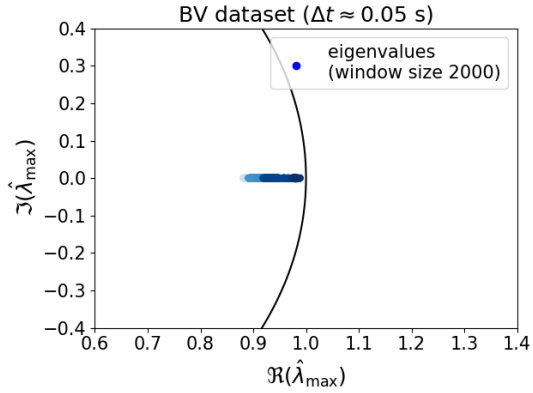
(a) CD power outage data and DEV trend



(b) BV power outage data without DEV trend



(c) CD power outage data: DEV trend in the complex plane



(d) BV power outage data: DEV trend in the complex plane

Figure 4.24: In the upper graphs of 4.24(a) and 4.24(b) the absolute value trend of the dominant eigenvalues is shown for the CD and BV dataset, drawn in the lower graphs. In 4.24(c)and 4.24(d) the dominant eigenvalue trend in the complex plane is shown for both datasets. Early and later estimated eigenvalues are indicated by the shading from light to dark. After an initially unstable behaviour of the CD eigenvalues they approach unity. This is a sign for an uprising bifurcation. The BV eigenvalues do not show such a clear trend and the reason for the blackout remains in question.

of 3500 time steps for the CD and 2000 time steps for the BV dataset. The rolling windows are chosen in order to maintain reasonable results without using too much data per window. The results for the absolute value of the dominant eigenvalue of the CD dataset, presented in the upper graph of 4.24(a), indicate a critical transition in the CD dataset by a positive trend. The eigenvalues even reach unity in magnitude and yield a similar curve shape as the

[8]The optimal order $p = 4$ is chosen based on the false nearest neighbour algorithm. The results can be found in subsection A.2.2.

autocorrelation at a lag of unity, shown in [Ehe17]. Moreover, there seems to be some plateau at a time around $t = 430\,\mathrm{s}$ until $t = 520\,\mathrm{s}$. The plateau of the dominant eigenvalue prior to the blackout could be due to inherent stabilizing mechanisms of the power system that were trying to prevent the power outage. In contrast, the dominant eigenvalue shows no trend for the BV dataset in the upper graph of 4.24(a). Because numerical issues are excluded due to an adapted sampling, the critical transition in the BV dataset should need another explanation than the one of the CD dataset.
The dominant eigenvalues in the Gaussian plane are shown in 4.24(c) and 4.24(d). The early eigenvalues are the light blue dots, the later estimated eigenvalues are the darker ones. The first eigenvalues of the CD dataset are unstable and scatter a bit in the complex plane. After a few time steps they become purely real and approach unity. Thus, the critical transition in the CD dataset should not be a Hopf or a flip bifurcation, but another type, e.g. a simple fold bifurcation. The dominant eigenvalues give no further reliable information of the critical transition occurring in the BV dataset. The eigenvalues are purely real and seem to approach unity, but they do not reach unity. Furthermore, the whole trend is confined in the 10^{-2} regime over the whole time range. In conclusion, the trend could indicate a bifurcation that may be observable for a longer BV time series, but the poorly pronounced trend for the eigenvalues could just as well indicate a critical transition that is not caused by a bifurcation.

4.4.4 Greenhouse-icehouse earth transition

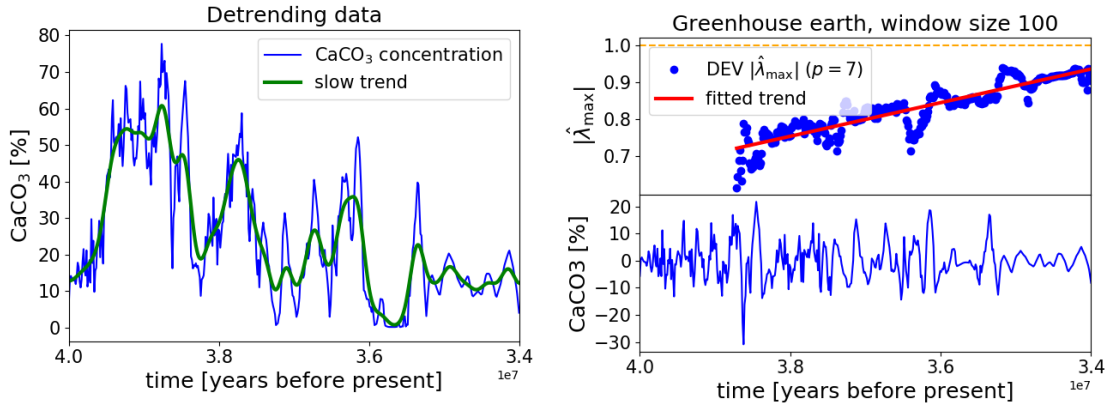
In the history of the earth's climate various situations are found in which the climate conditions change significantly and rather suddenly in time. These phenomena are nowadays believed to occur due to critical transitions including tipping points caused by bifurcations. More precisely, if some external climate parameters change over a critical threshold, the climate system underoes a bifurcation and, in consequence, changes its behaviour remarkably.[Dak+08]
This happened for example around 34 million years ago when the earth developed from an ice free climate state, the *greenhouse earth*, to a climate state with iced pole caps, the so-called *icehouse earth*. A time series of the $CaCO_3$ concentration in a tropical Pacific sediment core is investigated in [Dak+08]. This time series records the $CaCO_3$ concentration prior to the greenhouse-icehouse transition. Previous investigations in [Dak+08] have shown a significant increase of the autocorrelation $\rho_{X_t,X_{t-1}}$ prior to the climate transition. The original time series[9] is depicted in blue in 4.25(a).
In order to analyse the fast dynamics of the $CaCO_3$ concentration, a slow trend is fitted with a Gaussian filter, shown in green in 4.25(a). It is subtracted from the original time series. The result is shown in the lower graph of 4.25(b). The dominant eigenvalues are calculated for rolling time windows of hundred sample points and an AR order[10] $p = 5$. The results are stable for suboptimal orders p.[11] In the upper graph of 4.25(b), the positive dominant eigenvalue trend is

[9] The data is provided by Prof. Dr. Arndt Telschow from the University Osnabrück, institute for mathematics and informatics (Barbarastr. 12, 49076 Osnabrück)(arndt.telschow@uos.de).

[10] The order for the detrended dataset is determined via the false nearest neighbour algorithm, described in subsection 3.1.1, and can be found in subsection A.2.2.

[11] The robustness against suboptimal orders p is confirmed by the results of subsection A.1.3.

clearly observable. It is illustrated by a linear fit, shown in red.[12] An uprising critical transition is detected with the dominant eigenvalue as a leading indicator. The dominant eigenvalue does not reach unity, depicted by the orange dashed horizontal line, in the time interval of the investigated time series. This could indicate that the critical transition takes place in a time beyond the time series interval. Another simple explanation for this observation are uncertainties due to noise or numerical errors that influence the quality of the numerial methods negatively.



(a) Greenhouse earth: data detrending

(b) Greenhouse earth: DEV trend

Figure 4.25: In 4.25(a) the original time series of the $CaCO_3$ concentration is shown in blue. An underlying slow trend is fitted in green using a Gaussian filter. It is subtracted from the original data. The result is visible in the lower graph of 4.25(b). The detrended time series is analysed with the dominant eigenvalue approach. The dominant eigenvalues show a significant positive trend, guided by the red line, in the upper graph of 4.25(b). They do not reach the instability threshold of unity, marked by the orange dashed horizontal line. This could indicate that the actual transition takes place beyond the time series end or it is due to numerical uncertainties. The rolling time window size is 100 sampling points for an autoregression order $p = 5$.

[12] Note, that the linear trend is just a tool for illustration and does not necessarily correspond to the real time evolution law of the dominant eigenvalue.

5 Methodological investigations in R-tipping systems

In [RS16] the leading indicators autocorrelation $\rho_{X_t,X_{t-1}}$ and standard deviation σ are applied to an R-tipping system, simulated with a saddle-node normal form. The measures seem to increase prior to the actual R-tipping point and thus, are considered as early warning signals in this system. Regarding the theoretical R-tipping mechanism, described in subsection 2.2.2, there is no clear reason for the occurrence of CSD, presented in subsection 2.5.1, prior to the R-tipping point. Speaking in the mechanical analogue of subsection 2.5.1, CSD is based on the flattening effect of a stable state's potential valley prior to the B-tipping point. In the case of R-tipping the potential does not explicitly change its shape, but stable and unstable states in the phase space drift much faster than the system's typical relaxation time as explained in detail in subsection 2.2.2. The theory does not give any hint to CSD prior to an R-tipping point. Therefore, the reason for the positive trend of autocorrelation $\rho_{X_t,X_{t-1}}$ and standard deviation σ in the R-tipping system is investigated. Additionally, the DEV $\lambda_{\max}$ is estimated with an autoregression scheme AR(p) and the flow $\dot{x}$ estimation method, introduced in subsection 3.2.1 and subsection 3.2.2. Also the DEV $\lambda_{\max}$ can only be used as a leading indicator if CSD arises prior to the tipping point. In order to answer the question, why a positive trend is observed and whether it is really an early warning signal in this case, the R-tipping saddle-node model of [RS16] and an R-tipping transcritical model are introduced in section 5.1. The stable counterparts of these models without R-tipping are also presented in section 5.1 in order to compare the results. The positive trend is found for both, the stable and the unstable R-tipping variants. In a second step the analysis is performed again for the detrended datasets in order to answer the question if the positive trend is caused by the rapid shift of the time series data or has its origin in the occurrence of an R-tipping point as it would be necessary for a leading indicator. This would imply that a similar phenomenon as CSD occurs prior to R-tipping. As indicated by the theoretical absence of CSD, no positive trend is found for the detrended datasets.

5.1 Synthetic R-tipping data

In this section the saddle-node model, following the simulation instructions of [RS16], and an imperfect transcritical model, are introduced with and without an R-tipping point for the further analysis of the leading indicators in R-tipping systems.

5.1.1 R-tipping saddle-node normal form

The simulated R-tipping saddle-node normal form can be found in [RS16]. The saddle-node normal form can be written as

$$\dot{x} = (x + \zeta)^2 - 1 \tag{5.1}$$

with the normal form parameter set to unity. The stable and unstable equilibria are $x^{(s/u)}_{equi}(\zeta) = -\zeta \mp 1$. They are presented in Figure 5.1. The simulations are performed with a Runge-Kutta 4 method with a time step size of $dt = 0.001$ in an interval of $t \in [-10, 10]$ and additional Gaussian white noise $\mathcal{N}(\mu = 0, \sigma^2 = 0.0625) \cdot \sqrt{dt}$ as described in subsection 3.4.2. The parameter ζ is varied following the function

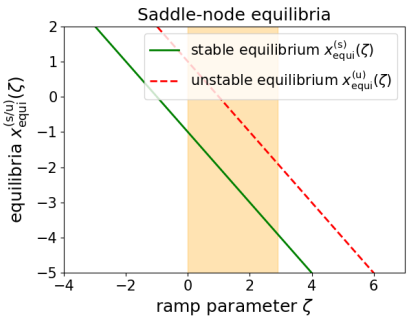
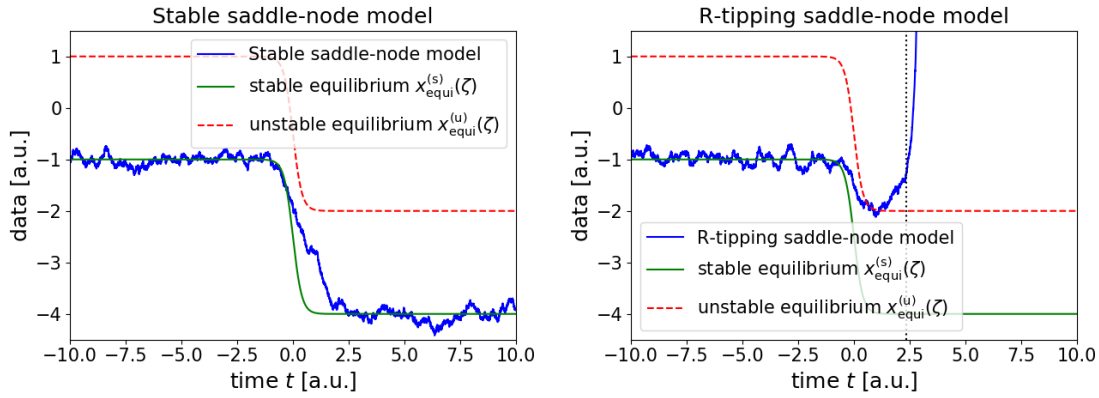


Figure 5.1: The stable and unstable equilibria $x^{(s/u)}_{equi}(\zeta)$ are shown depending on the ramp parameter ζ. The orange rectangle marks the interval of the ramp parameter change that obeys the time evolution law of Equation 5.2.

$$\zeta(t) = \frac{\zeta_{max}}{2}\left[\tanh\left(\frac{\zeta_{max} \cdot \varepsilon \cdot t}{2}\right) + 1\right] \tag{5.2}$$

with the ramp distance $\zeta_{max} = 3$, the ramp speed ε and the initial condition $\zeta_0 = 0$. The initial condition is $x_0 = -1$. The saddle-node dataset without R-tipping is simulated with the ramp speed $\varepsilon = \frac{2}{3}$. It is shown in 5.2(a). The stable equilibrium is drawn as green solid line, the unstable one as red dashed line. The model is able to follow the rapidly shifted stable branch. The R-tipping saddle-node model[1] is calculated with a ramp speed of $\varepsilon = \frac{5}{3}$. The resulting dataset is presented in 5.2(b). The slightly perturbed data, indicated by the blue solid line, follows the stable equilibrium until it is driven to the unstable one due to the rapid shift of the parameter ζ. Note, that the system diverges with a time delay after crossing the unstable equilibrium $x^{(u)}_{equi}(\zeta)$ and not as intuitively assumed, at the time when the stable and unstable equilibria $x^{(s/u)}_{equi}(\zeta)$ are closest to each other. This phenomenon of a delayed R-tipping is investigated in detail in [RS16]: The maximum escape rate to infinity is found to be delayed after the system has crossed the unstable equilibrium $x^{(u)}_{equi}(\zeta)$. This can be analogously observed in 5.2(b). The actual R-tipping point is approximately depicted by the black dotted vertical line.

[1] The time series is cut for values greater than 100000 for practical reasons.



(a) Saddle-node model without R-tipping

(b) R-tipping saddle-node model

Figure 5.2: Both, the stable dataset in 5.2(a) and the R-tipping saddle-node model in 5.2(b) are shown in blue. The green solid line indicates the stable equilibrium, the red dashed line the unstable one. The models are simulated with Equation 5.1. The stable saddle-node model is calculated with the ramp speed $\varepsilon = \frac{2}{3}$ and the R-tipping model is computed using the ramp speed $\varepsilon = \frac{5}{3}$. The ramp parameter ζ is shifted governed by Equation 5.2 from 0 to 3. The initial condition is $x_0 = -1$. Gaussian white noise of the form $\mathcal{N}(\mu = 0, \sigma^2 = 0.0625) \cdot \sqrt{\mathrm{d}t}$ is added. The actual R-tipping point is approximately depicted by the black dotted vertical line. Further details of the simulations are described in the running text.

5.1.2 R-tipping imperfect transcritical model

An imperfect transcritical normal form, written as

$$\dot{x} = x \cdot ((a+1) \cdot r - x) + h - \frac{a^2 r^2}{4} - \frac{2ar^2}{4} \tag{5.3}$$

is used to simulate an R-tipping dataset with the control parameter r, the imperfection h and an additional imperfection coefficient a that tilts the asymptote away from zero. The stable and unstable equilibria of the system are

$$x^{(\mathrm{s/u})}_{\mathrm{equi}}(r) = \frac{(a+1) \cdot r}{2} \pm \sqrt{\frac{r^2}{4} + h}. \tag{5.4}$$

They are shown in Figure 5.3. The additional coefficient a is needed to guarantee the existence of fixed points that can be stable or unstable depending on the control parameter r without undergoing a bifurcation. The parameter r is used as the so-called ramp parameter of the R-tipping system.

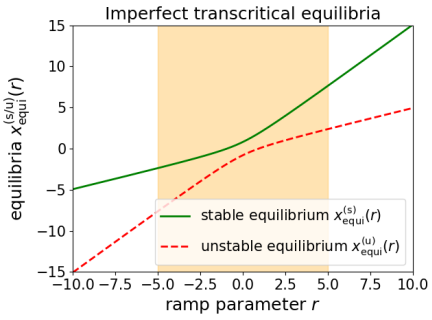
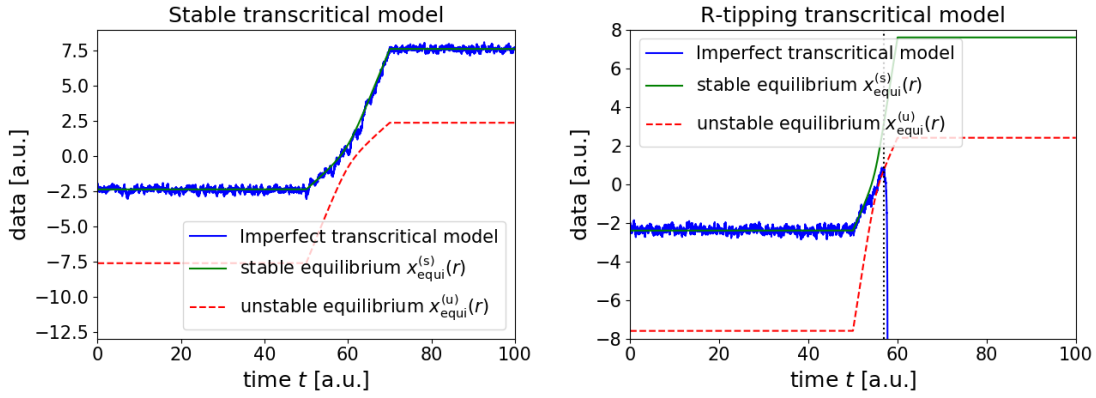


Figure 5.3: The stable and unstable equilibria $x^{(\mathrm{s/u})}_{\mathrm{equi}}(r)$ as a function of the control parameter r, also called ramp parameter, are shown for the simulated imperfect transcritical model. The orange rectangle denotes the interval of the rapid linear shift of the ramp parameter r.

The simulation of the datasets of Equation 5.3 is done with a Runge-Kutta 4 method with a time step size of $\mathrm{d}t = 0.001$ and additional Gaussian white noise $\mathcal{N}(\mu = 0, \sigma^2 = 0.25) \cdot \sqrt{\mathrm{d}t}$ as described in subsection 3.4.2. The control parameter r is linearly increased in the interval of -5

to 5, marked in orange in Figure 5.3. The ramp shift begins at time $t = 50$. The stable imperfect transcritical model is simulated with a rate of change of around $\kappa \approx \frac{1}{2}$ per time step of unity. It is presented in 5.4(a). The stable equilibrium is depicted by the green solid line, the unstable one is marked by the red dashed line. For $t < 50$ the stable and unstable branches remain at their fixed positions for $r = -5$, visible in Figure 5.3. At $t = 50$ the parameter r changes rapidly and pushes the two branches in a short time of around 10 time steps towards their new fixed positions for $r = 5$. The time series data, shown in blue, is able to follow the rapidly shifted stable equilibrium. The imperfection coefficients are chosen as $a = 1$ and $h = 0.65$. The initial condition is $x_0 = -2$.
The simulations of the R-tipping imperfect transcritical model[2] are performed with a rate of change of around $\kappa \approx 1$ per time step of unity. The imperfection coefficients are chosen as defined above for the stable transcritical system. The R-tipping dataset is shown in 5.4(b). The doubled rate of change leads to the R-tipping effect: The system cannot follow its equilibrium state as fast as the state changes until the previously stable state develops unstable and the system diverges towards minus infinity.



(a) Imperfect transcritical model without R-tipping

(b) R-tipping imperfect transcritical model

Figure 5.4: Both, the stable dataset in 5.4(a) and the R-tipping imperfect transcritical model in 5.4(b) are shown in blue. The green solid line indicates the stable equilibrium, the red dashed line the unstable one. The models are simulated with Equation 5.3. The stable transcritical model is calculated with the rate of change $\kappa = \frac{1}{2}$ per time step of unity and the R-tipping model is computed using the doubled rate of change $\kappa = 1$ per time step of unity. The ramp parameter r is linearly shifted starting at time $t = 50$ from -5 to 5. The imperfection coefficients are chosen as $a = 1$ and $h = 0.65$. The initial condition is $x_0 = -2$. Gaussian white noise of the form $\mathcal{N}(\mu = 0, \sigma = 0.5) \cdot \sqrt{\mathrm{d}t}$ is added. The actual R-tipping point is approximately depicted by the black dotted vertical line. Further details of the simulations can be found in the running text.

[2]The time series is cut for values less than -100000 for practical reasons.

5.2 Leading indicators

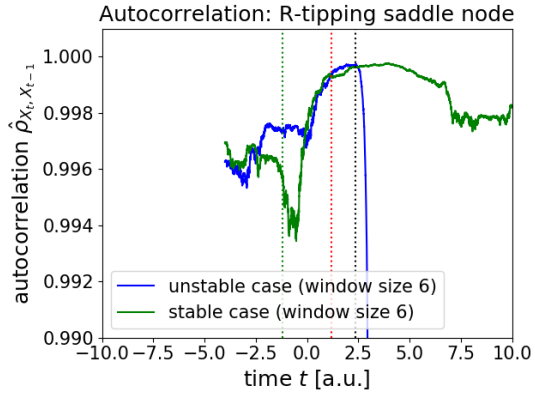
The leading indicators autocorrelation $\rho_{X_t,X_{t-1}}$, standard deviation σ and the DEV $\lambda_{\max}$ with autoregression AR(p) and the flow $\dot{x}$ estimation method are applied to the stable models and their R-tipping counterparts of section 5.1. In a second step the datasets are detrended in order to investigate if the time series data shift alone causes the increase in the leading indicators. If the detrended data would yield positive trends for the R-tipping datasets and no trend for the stable datasets, a similar phenomenon as CSD for B-tipping systems should arise prior to R-tipping. This is not implied by the theoretical description of R-tipping in subsection 2.2.2, but suggested by the investigations in [RS16].

5.2.1 Application to the R-tipping models

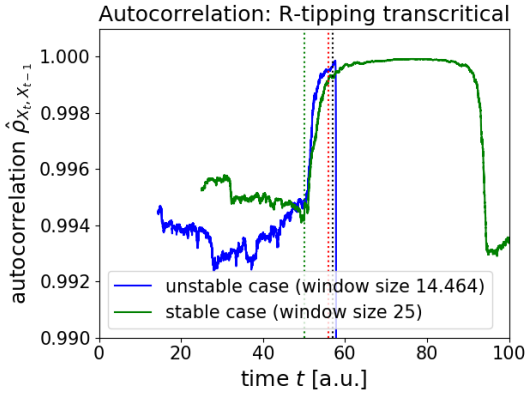
The autocorrelation $\rho_{X_t,X_{t-1}}$ estimation method is applied to the saddle-node and the imperfect transcritical models of subsection 5.1.1 with and without an R-tipping point. A rolling time window size of 6 time steps is chosen for the saddle-node dataset. The calculations of the imperfect transcritical model are performed with a rolling time window of $\frac{1}{4}$ of the size of the whole investigated time series. The results are shown in 5.5(a) and 5.5(b). The green dotted vertical line indicates the beginning of the ramp shift, the red dotted vertical line marks the time at which the unstable equilibrium $x^{(\mathrm{u})}_{\mathrm{equi}}$ is reached by the unstable system and the black dotted vertical line approximately shows the actual R-tipping point of the unstable system. Both, the stable and the unstable R-tipping systems show some random fluctuations without any clear trend prior to the ramp shift as expected and increase in a very similar way after the ramp shift's start time. In the unstable case the autocorrelation $\rho_{X_t,X_{t-1}}$ diverges suddenly to minus infinity at the time of the R-tipping point. In the stable case the autocorrelation $\rho_{X_t,X_{t-1}}$ also reaches approximately unity at that time and remains at that value over the range of one rolling time window size, before it decreases. These results imply that the autocorrelation $\rho_{X_t,X_{t-1}}$ that is used in the case of B-tipping systems react rather to the rapid ramp shift in the time series data than to some specific phenomenon prior to the R-tipping point.
A similar behaviour is observed for the standard deviation σ with the same rolling time window size as illustrated by 5.5(c) and 5.5(d). The standard deviation σ increases in both, the stable and unstable case. The increase is smaller for the unstable datasets because of the smaller ramp shift in the time series data due to the uprising R-tipping point. At the tipping point the standard deviation σ diverges to infinity, whereas the stable datasets reach a maximum in the standard deviation before they decrease towards zero. For both measures in Figure 5.5, the reaction to the ramp shift is delayed in the R-tipping saddle-node model due to a big time window size. Nevertheless, it is chosen to stabilize the numerical results.

In order to complete the set of investigated leading indicators for the R-tipping models the DEV $\lambda_{\max}$ is estimated via an autoregression scheme AR(p) with the optimal order[3] $p = 3$ for the saddle-node model and $p = 4$ for the imperfect transcritical model and via the flow $\dot{x}$

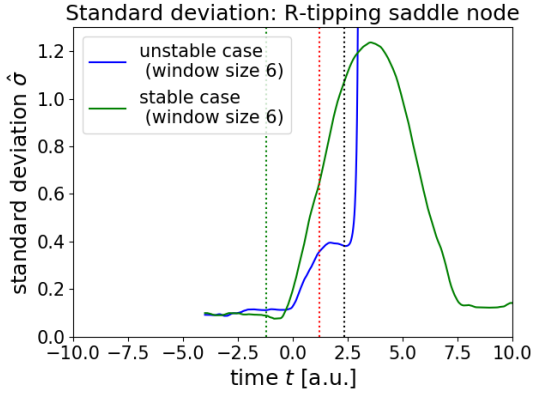
[3]The optimal AR(p) order is determined by applying the false nearest neighbour algorithm, presented in subsection 3.1.1. The optimal embedding dimension is chosen as the optimal order p. The results of the false nearest neighbour algorithm can be found in subsection A.2.3.



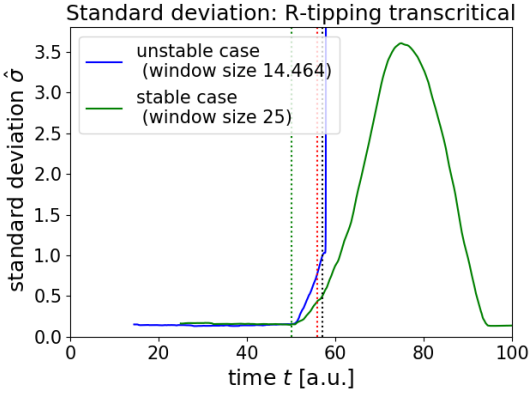
(a) Autocorrelation for the saddle node model



(b) Autocorrelation for the imperfect transcritical model



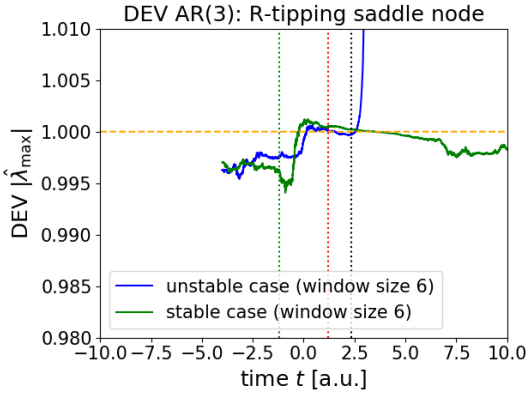
(c) Standard deviation for the saddle node model



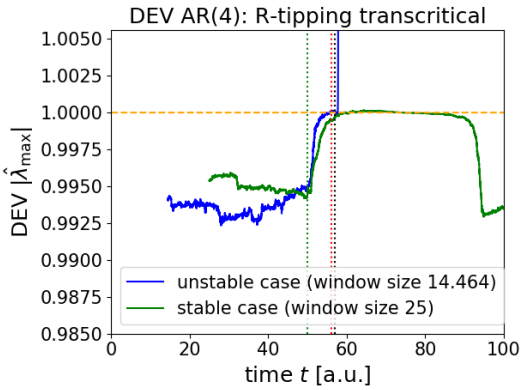
(d) Standard deviation for the imperfect transcritical model

Figure 5.5: The time evolution of the autocorrelation $\rho_{X_t,X_{t-1}}$ and the standard deviation σ is evaluated for the stable and unstable R-tipping saddle-node and imperfect transcritical models. The rolling time window sizes are $\Delta t = 6$ for the saddle-node and $\frac{1}{4}$ of the time series length for the transcritical models. The results for the saddle-node model are presented in 5.5(a) and 5.5(c). The calculations for the imperfect transcritical model are shown in 5.5(b) and 5.5(d). The green dotted vertical line indicates the beginning of the ramp shift, the red dotted vertical line marks the time at which the R-tipping datasets reach the unstable equilibrium $x^{(u)}_{equi}$ and the black dotted vertical line approximately shows the actual R-tipping point. The trends of the evaluated measures are positive in both, the stable and unstable cases during the ramp shift. The delayed response of the autocorrelation $\rho_{X_t,X_{t-1}}$ and the standard deviation σ for the saddle-node model is due to the big time window size. Based on these results it is impossible to anticipate whether R-tipping occurs or not.

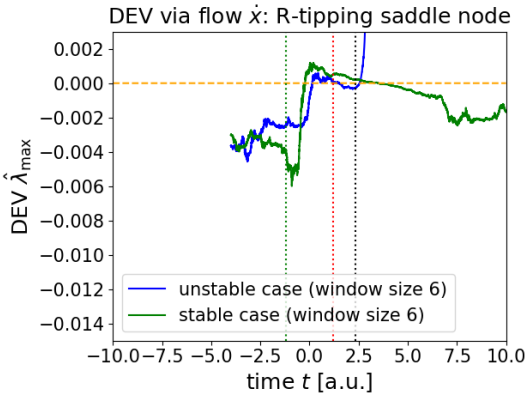
estimation method, described in subsection 3.2.2. The chosen rolling time window sizes are the same as in the calculations above. The results are presented in Figure 5.6.
The analogous behaviour as for the autocorrelation $\rho_{X,t,X_{t-1}}$ and the standard deviation σ is observed for the DEV λ_{max} trends. The unstable as well as the stable datasets show a positive



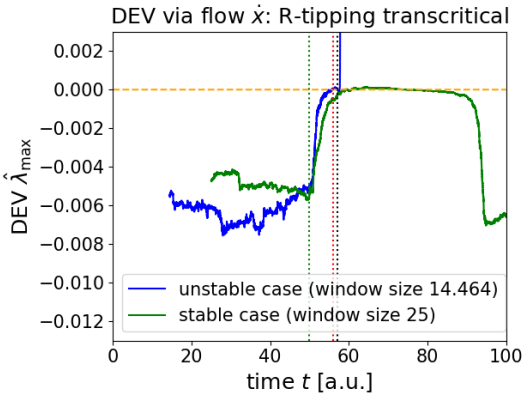
(a) DEV AR(p) trend for the saddle-node model



(b) DEV AR(p) trend for the imperfect transcritical model



(c) DEV flow $\dot{x}$ trend for the saddle-node model



(d) DEV flow $\dot{x}$ trend for the imperfect transcritical model

Figure 5.6: The time evolution of the DEV λ_{max} is evaluated for the stable and unstable R-tipping saddle-node and imperfect transcritical models. The rolling time window sizes are $\Delta t = 6$ for the saddle-node and $\frac{1}{4}$ of the time series length for the transcritical models. The DEV λ_{max} evaluation for the saddle-node model with the autoregression AR(p) and the flow $\dot{x}$ approach is shown in 5.6(a) and 5.6(c). The analysis for the imperfect transcritical model is shown in 5.6(b) and 5.6(d). The optimal AR(p) orders are chosen as $p = 3$ for the saddle-node and $p = 4$ for the imperfect transcritical model based on the fnn algorithm, introduced in subsection 3.1.1. The green dotted vertical line indicates the beginning of the ramp shift, the red dotted vertical line marks the time at which the R-tipping datasets reach the unstable equilibrium $x^{(u)}_{equi}$ and the black dotted vertical line approximately shows the actual R-tipping point. The DEV λ_{max} trends are positive in both, the stable and unstable cases during the ramp shift. The DEV λ_{max} diverges at the R-tipping point for the unstable model and decreases approximately one time window size after crossing the stability threshold, denoted by the orange dashed horizontal line, for the stable datasets. The delayed response of the DEV λ_{max} to the ramp shift in the saddle-node models is due to the big time window size. Based on these results it is impossible to anticipate whether R-tipping occurs or not.

trend in the DEV λ_{max} during the ramp shift of the time series data. The DEV approaches the typical instability thresholds in vicinity of the unstable equilibrium $x^{(u)}_{equi}$ and diverges

approximately at the actual R-tipping point of the system. In the stable case the DEV $\lambda_{\max}$ decreases approximately one time window size after crossing the instability threshold.
These results imply that the investigated measures cannot be seen as leading indicators of R-tipping points, but only as artefacts of the sudden time series data shift due to overlapping time windows. In order to exclude the possibility that a small influence on the investigated measures that is caused by a phenomenon similar to CSD, but prior to R-tipping points, is hidden by the artificial effects of the rapid time series data shift, the analysis is repeated for the detrended variants of the datasets in subsection 5.2.2.

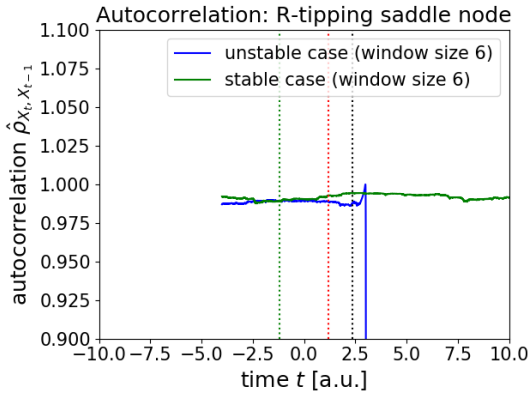
5.2.2 Application to the detrended R-tipping models

The investigations in subsection 5.2.1 suggest that the leading indicators autocorrelation $\rho_{X_t,X_{t-1}}$, standard deviation σ and the DEV λ respond to the rapid shift in the time series data with a positive trend. They do not react to the occurrence of an R-tipping point. Thus, the investigated stable and unstable saddle-node and imperfect transcritical models, introduced in section 5.1, are detrended and the analysis of subsection 5.2.1 is analogously performed to these detrended datasets[4]. The procedure and the parameters are chosen analogously to the evaluations in subsection 5.2.1.
The results for the autocorrelation $\rho_{X_t,X_{t-1}}$ and the standard deviation σ are presented in Figure 5.7. The green dotted vertical line depicts the starting point of the R-tipping ramp shift. The red dotted vertical line indicates the time at which the R-tipping datasets cross the unstable equilibrium $x^{(\mathrm{u})}_{\mathrm{equi}}$ and the black dotted vertical line marks the actual R-tipping point. Both measures do not react to the approaching R-tipping point of the two example systems. The measures remain constant apart from random fluctuations before the R-tipping occurs and thus, prior to R-tipping points no effect exists that could cause a reliable trend in the autocorrelation $\rho_{X_t,X_{t-1}}$ or the standard deviation σ similar to the CSD in B-tipping systems. The divergence of the measures after the R-tipping point is caused by the divergence in the detrended datasets at the tipping point.

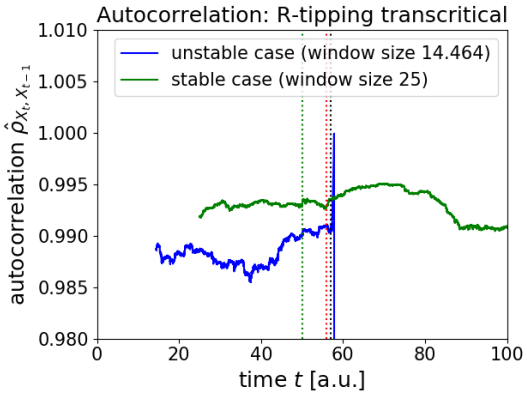
In order to complete the investigations of the leading indicators in R-tipping systems, the DEV $\lambda_{\max}$, estimated with the AR(p) and the flow $\dot{x}$ approach, is evaluated for the detrended datasets in Figure 5.8. The optimal AR(p) order[5] for the saddle-node models is chosen to be $p = 3$ and $p = 4$ for the imperfect transcritical model. No DEV $\lambda_{\max}$ trend is observable in Figure 5.8. Similar to the standard statistical measures, the DEV $\lambda_{\max}$ trend depends on the existence of some effect similar to CSD and thus, does not serve as leading indicator in the R-tipping systems.

[4]The detailed procedure of the detrending approach and the detrended datasets are presented in subsection A.2.4.

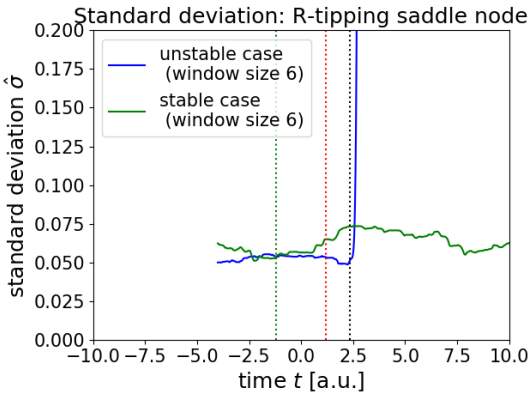
[5]The optimal AR(p) order is determined by applying the false nearest neighbour algorithm, presented in subsection 3.1.1. The optimal embedding dimension is chosen as the optimal order p. The results of the false nearest neighbour algorithm can be found in subsection A.2.3.



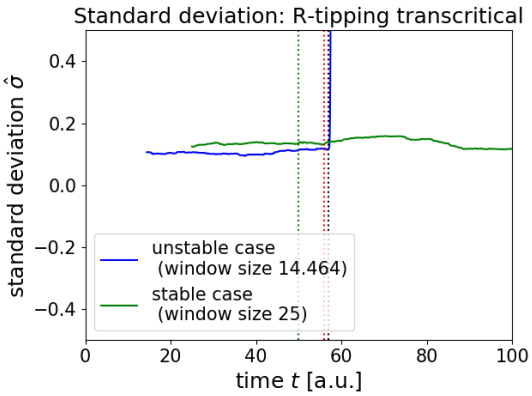
(a) Autocorrelation for the detrended saddle-node model



(b) Autocorrelation for the detrended transcritical model

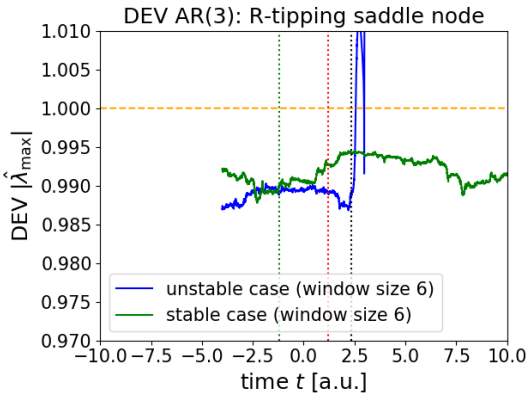


(c) Standard deviation for the detrended saddle-node model

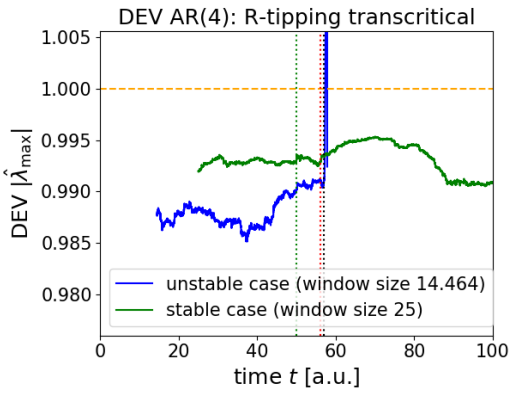


(d) Standard deviation for the detrended transcritical model

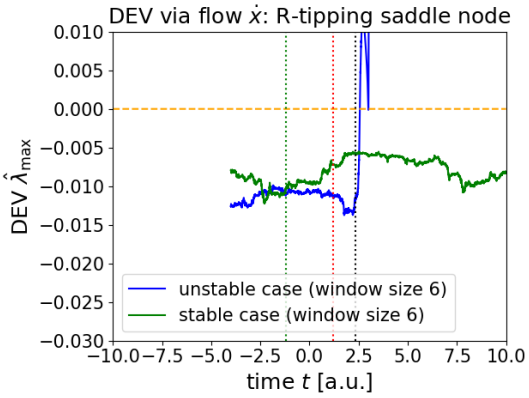
Figure 5.7: The time evolution of the autocorrelation $\rho_{X_t,X_{t-1}}$ and the standard deviation σ for the detrended versions of the stable and unstable saddle-node models is presented in 5.7(a) and 5.7(c). The results for the detrended imperfect transcritical models are shown in 5.7(b) and 5.7(d). The rolling time window size for the saddle-node model is $\Delta t = 6$. The rolling time windows for the imperfect transcritical models are chosen as $\frac{1}{4}$ of the whole time series length. The green dotted vertical line indicates the starting ramp shift, the red dotted vertical line is the time at which the unstable equilibrium $x^{(u)}_{equi}$ is crossed by the R-tipping datasets and the black dotted vertical line marks the actual R-tipping point. No trend is visible neither for the autocorrelation $\rho_{X_t,X_{t-1}}$ nor for the standard deviation σ. The divergence of the measures for the unstable datasets is due to the divergence of the time series data around the R-tipping point. Thus, the measures do not serve as leading indicators of R-tipping points.



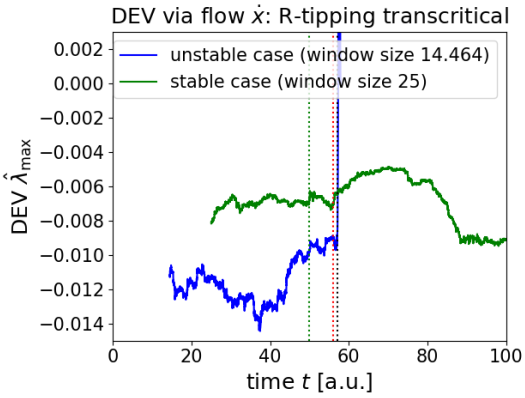
(a) DEV AR(p) trend for the detrended saddle-node model



(b) DEV AR(p) trend for the detrended transcritical model



(c) DEV flow $\dot{x}$ trend for the detrended saddle-node model



(d) DEV flow $\dot{x}$ trend for the detrended transcritical model

Figure 5.8: The time evolution of the DEV $\lambda_{\max}$ for the detrended versions of the stable and unstable saddle-node models, estimated via an autoregression scheme AR($p = 3$) and via the flow $\dot{x}$ approach, is presented in 5.8(a) and 5.8(c). The results for the detrended imperfect transcritical models are shown in 5.7(b) and 5.7(d). The AR order $p = 4$ is chosen for the transcritical model. The rolling time window size for the saddle-node model is $\Delta t = 6$. The rolling time windows for the imperfect transcritical models are chosen as $\frac{1}{4}$ of the whole time series length. The green dotted vertical line indicates the starting ramp shift, the red dotted vertical line is the time at which the unstable equilibrium $x^{(u)}_{\text{equi}}$ is crossed by the R-tipping datasets and the black dotted vertical line marks the actual R-tipping point. No trend is visible for the DEV $\lambda_{\max}$. The divergence of the measures for the unstable datasets is due to the divergence of the time series data around the R-tipping point. Thus, the measures do not serve as leading indicators of R-tipping points.

6 Conclusion and outlook

This Master's thesis deals with the idea of anticipating critical transitions in complex dynamical systems. Various standard statistical approaches are investigated, the prediction quality and numerical robustness are evaluated and new numerical methods in order to estimate the dominant eigenvalue trend as leading indicator of a system are introduced. In order to complete this work, a final conclusion of the key results is given in this chapter.

The commonly observed relation between rolling time window size, volatility and delay of the estimated measures is reproduced for the standard statistical measures of autocorrelation $\rho_{X_t,X_{t-1}}$, standard deviation σ and skewness Γ as well as for the introduced DEV λ_{max}.

The capacity of the autocorrelation $\rho_{X_t,X_{t-1}}$ considering B-tipping datasets is controlled. It works as leading indicator in most of the investigated cases. It deserves relatively low amounts of data to be stable estimated and it provides significant predictive power for a wide range of lead times k. A limit of the simple increasing autocorrelation trend has been observed for the Hénon map and its period doubling bifurcation: The trend is decreasing for a time lag of unity and fixed around unity for a time lag of two.

The standard deviation σ is not applicable in such a comprehensive way as the autocorrelation $\rho_{X_t,X_{t-1}}$ or the DEV λ_{max} estimation. The standard deviation σ yields no positive trend for two of the simulated B-tipping datasets although they undergo a bifurcation. In the cases of positive standard deviation trends, the estimation needs a bit more data than the autocorrelation, but decreases significantly in its predictive power with increasing lead times k. It seems to be a rather bad leading indicator in comparison to the autocorrelation $\rho_{X_t,X_{t-1}}$ and the DEV λ_{max} estimation.

The skewness Γ is a special leading indicator for B-tipping systems and can provide useful information only in the case of flickering datasets. Providing that one needs to investigate an obviously flickering system or wants to obtain information, whether the system flickers or not, the skewness Γ yields a useful tool.

The estimation of the DEV λ_{max} using autoregression AR(p) turns out to be a rather useful tool to predict critical transitions. It provides comparable predictive power as the autocorrelation $\rho_{X_t,X_{t-1}}$ over a wide range of lead times k and is robust against suboptimal choice of the AR(p) order p. Although it needs more data than the autocorrelation $\rho_{X_t,X_{t-1}}$ and standard deviation σ in order to be significant and stable estimated, two advantages are predominant:

1. The derived theoretical formalism in subsection 2.3.1 guarantees an upper limit or threshold $\lambda_{threshold} = 1$ at which a critical transition takes place. This can be used to get an idea of the vicinity to a critical transition and the undesired consequences that originate from that vicinity. Furthermore, if it can reasonably assumed that the trend continues, the time of the tipping point can be roughly predicted. Even though an upper limit of unity also exists for the autocorrelation, it is not necessarily reached before the critical transition an thus, cannot be interpreted in terms of the vicinity to the tipping point.

2. The theory of local bifurcations provides the opportunity to classify various kinds of

bifurcations due to the DEV $\lambda_{\max}$ evolution in the complex plane. This feature is unique in the field of leading indicators and enables to gain more information over possible consequences of the critical transition like uprising oscillations of an approaching Hopf bifurcation.

Furthermore, the DEV $\lambda_{\max}$ estimation method with autoregression process AR(p) has been applied to four real time series from different areas of research. The method is able to serve as early warning signal in the four different real world systems and it is just slightly more complicated than the standard statistical measures in its application. Note, that the classification of the bifurcation is also possible for these empirical time series.
The DEV $\lambda_{\max}$ estimation method is also implemented using S-maps instead of a simple autoregression scheme. The two methods lead to more or less equal results, but the S-map approach is rather complex and difficult to implement in python compared to the simple predefined autoregression approach. Besides, the S-map approach is explicitly developed to estimate the time derivatives of systems that are non-stationary [Dey+16], but stationarity is assumed within each of the single rolling time windows. In the numerical methods this fact is taken into account by averaging the eigenvalues in each rolling time window or by averaging over the the state space vectors of a rolling window in order to get the mean equilibrium state vector before calculating the DEV of the equilibrium state vector's time evolution. Anyway, it is a possible procedure. But note, that one of the great advantages of the S-map method, namely the possibility to track the time derivatives of a non-stationary system, is lost by applying that procedure. Despite that fact, in some cases the DEV $\lambda_{\max}$ estimation with S-maps can still be useful for two reasons:

1. The S-map approach makes predictions of future states based on similar states in state space, whereas the AR(p) process is based on states that are past in time. Thus, the two methods represent two different ways to gain information of the available data and differ in their heuristic ideas. Their difference is visualized in Figure 2.8.

2. The S-map approach makes direct use of Takens' embedding theorem. The embedding and the idea of estimating future states from a point cloud in state space seems to provide more information out of the same amount of data. In consequence, for some systems the S-map approach stays stable for small window sizes whereas the AR(p) process yields unstable results, shown in Figure 4.19.

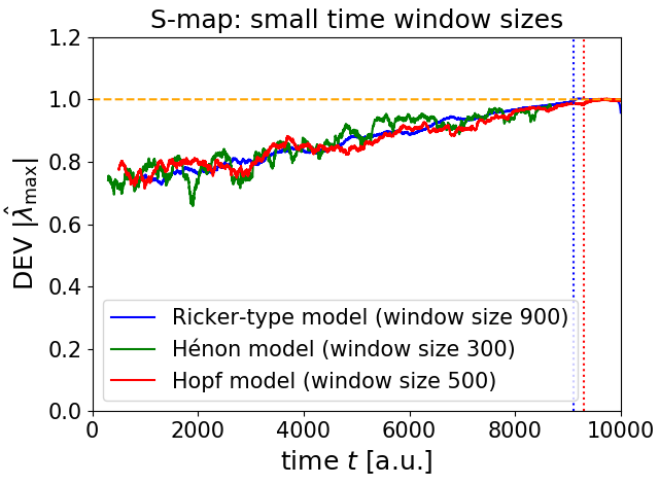
In order to guarantee an optimal choice of numerical embedding parameters the false nearest neighbour algorithm and the average distance from diagonal algorithm are applied to the time series prior to the further analysis. The false nearest neighbour algorithm turned out to give reasonable suggestions for a good embedding dimension d for the DEV S-map method and suitable orders $p = d$ in the case of the DEV AR(p) approach. It corresponds directly to the best orders p for a classification of the different bifurcation types as analysed in subsection 4.3.6. This is an advantage of the false nearest neighbour method in terms of the computation of the optimal choice of the autoregression order p.
The average distance from diagonal algorithm yields reasonable results for the time lag τ for the DEV S-map method. The algorithm gives also some inside in the system's dynamical structure as it shows periodic patterns for the Hopf data or alternating behaviour with period two for

the Hénon map with a period doubling bifurcation.
Additionally, it should be shortly noted, that the conditions for a good bifurcation type classification, as investigated in subsection 4.3.6, are complex. For the Ricker-type model with a fold bifurcation a bigger time window corresponds to better classification trends, for the Hénon model a rather small time window size yields the best results, because of some widely spreading eigenvalues near unity instead of minus unity for big rolling windows. The Hopf model does not seem to depend strongly on the rolling window size at all apart from at least the minimally desired amount of data to obtain stable estimates.
The presence of R-tipping and N-tipping systems leads to a general limitation of the considered leading indicators. In chapter 5 the applicability of the autocorrelation $\rho_{X_t,X_{t-1}}$ and the standard deviation σ to R-tipping systems is investigated. Additionally, the DEV $\lambda_{\max}$ estimation with an autoregression scheme AR(p) and the flow $\dot{x}$ is applied to the R-tipping data. The autocorrelation $\rho_{\mathrm{X_t,X_{t-1}}}$, the standard deviation σ and the DEV $\lambda_{\max}$ are not able to predict an uprising instability due to R-tipping. The autocorrelation $\rho_{\mathrm{X_t,X_{t-1}}}$ and the standard deviation σ as early warning signals are - supported by the results of chapter 5 - erroneously applied to an R-tipping system in [RS16]. The measures only react to the rapid ramp shift of the time series data. This result is in principal agreement to the theory of R-tipping systems: The investigated leading indicators are based on the occurrence of CSD prior to the tipping point, but CSD does not arise in the vicinity of an R-tipping point. In consequence, the positive trend of the considered measures does not depend on whether an R-tipping instability actually occurs or not.
Obviously, the same results are obtained for N-tipping systems. A system that undergoes a critical transition based on random fluctuations does not exhibit any changing measure at all because the system itself does not change in the vicinity of the N-tipping point. This result is illustrated for a simulated dataset in subsection A.1.4. In order to judge whether a critical transition that is caused by noise could occur, it would be necessary to gain information about the noise process itself corresponding to the distance of stable and unstable states of the system.
If it could be possible to anticipate R-tipping and N-tipping, could be the topic of future research. Another current idea is detecting the time at which a system's control parameter starts to change based on Bayesian statistics. This would give further inside into the evolution of a system and serves as early warning signal for changing system dynamics as well. In combination with the presented leading indicators that could be a more reliable tool to detect precociously critical transitions. Further investigative aims could be the development of methods that are able to determine whether a control parameter changes continuously and how that affects the evolution of the DEV $\lambda_{\max}$. This could be used to extrapolate the DEV trend prior to the tipping point and give a rough idea of the time at which the critical transition arises.
Finally, the application to times series of currently important systems has to be pushed ahead. The leading indicators could be applied to tension and stress data of faults where earthquakes emerge. The ocean with its sea currents is an important part of the climate system and a critical transition in the ocean currents could have complex consequences for the climate and ecological ocean systems.

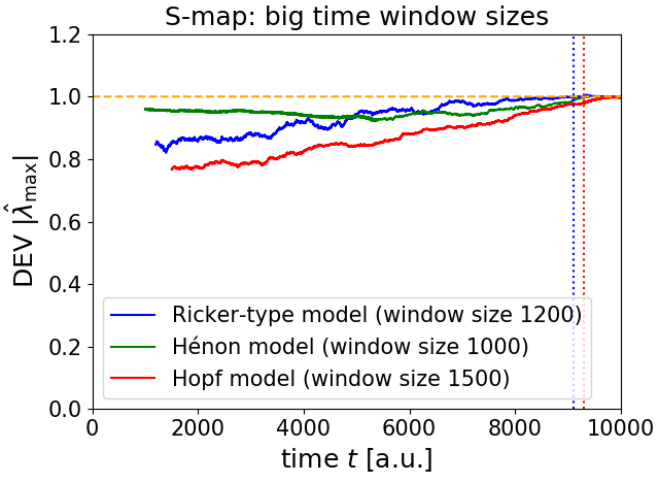
A Appendix

A.1 Additional results

In the following subsections additional results are presented that are useful in terms of comparison and completion of the main results of this Master's thesis.



(a) Absolute DEV trend (S-map): Window size 900 for the Ricker-type, 300 for the Hénon and 500 for the Hopf model.

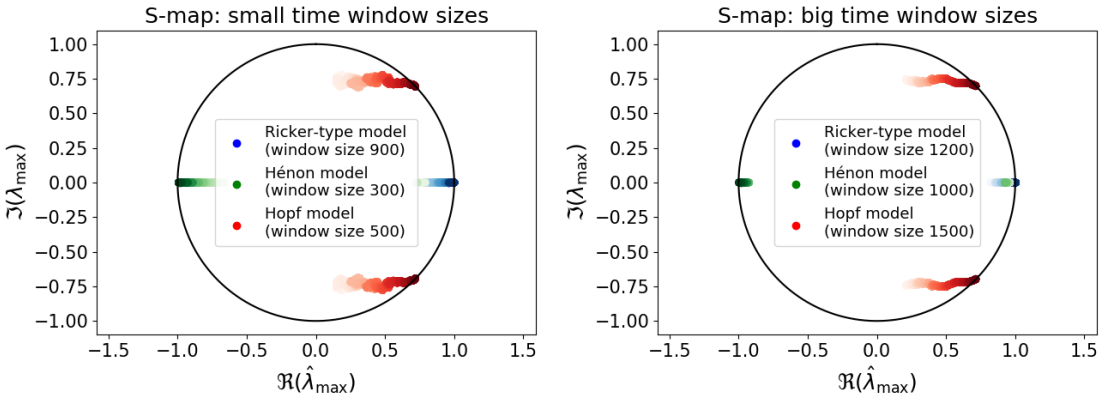


(b) Absolute DEV trend (S-map): Window size 1200 for the Ricker-type, 1000 for the Hénon and 1500 for the Hopf model.

A.1.1 Eigenvalue estimation with S-maps

The results are presented in Figure A.1 and Figure A.2. The absolute value trend in A.1(a) and A.1(b) is very similar to the results in Figure 4.13. The classification of bifurcation types in the Gaussian plane, shown in A.2(a) and A.2(b), yields widely the same results as its AR(p) counterparts in Figure 4.17.

Figure A.1: The presented results in A.1(a) and A.1(b) are calculated with S-maps for the parameters that are used in subsection 4.3.2. The results are very similar to the dominant eigenvalue estimation with AR(p).



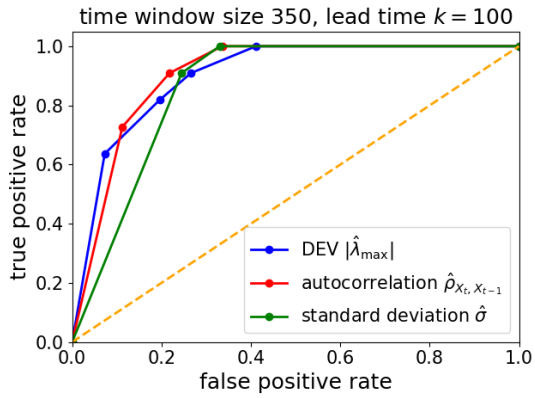
(a) Complex DEV trend (S-map): Window size 900 for the Ricker-type, 300 for the Hénon and 500 for the Hopf model.

(b) Complex DEV trend (S-map): Window size 1200 for the Ricker-type, 1000 for the Hénon and 1500 for the Hopf model.

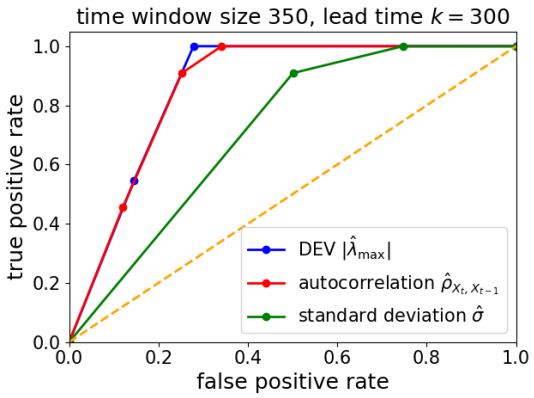
Figure A.2: The presented results in A.2(a) and A.2(b) are calculated with S-maps for the parameters that are used in subsection 4.3.5. The results are very similar to the dominant eigenvalue estimation with AR(p).

A.1.2 ROC curves with bigger time window size

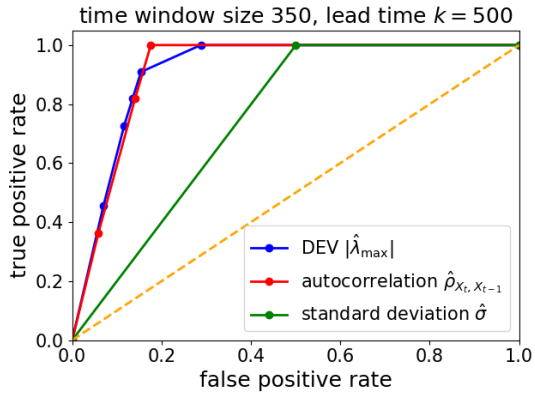
The calculations of 4.16(a), 4.16(b) and 4.16(c) are performed again for a bigger window size of 350 time steps. The trend, described in subsection 4.3.4, is observable without fluctuations in Figure A.3.



(a) ROC curve with lead time $k = 100$



(b) ROC curve with lead time $k = 300$

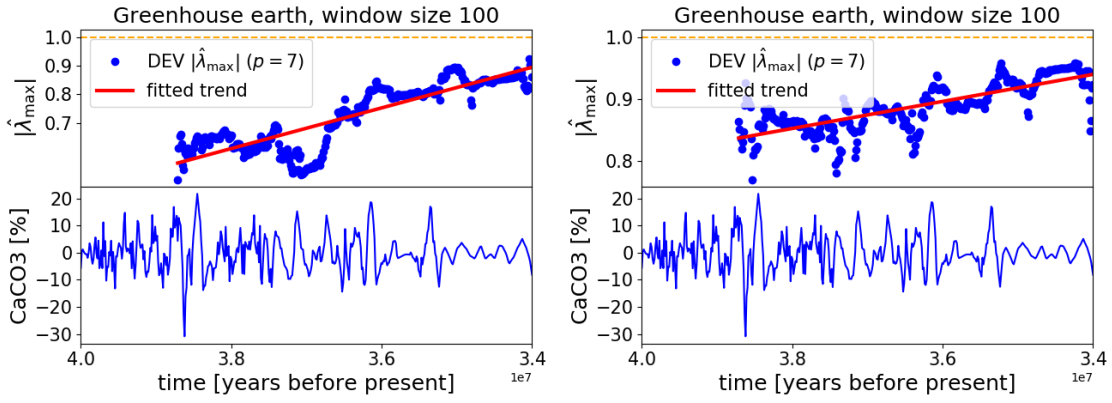


(c) ROC curve with lead time $k = 500$

Figure A.3: The graphs show the results of the calculations in subsection 4.3.4 with a time window size of 350 time steps. The trend, described in subsection 4.3.4, is clearly visible. The lead times k are adjusted because of the bigger window size. For a time series of 1000 points and a time window size of 350 points the first indicator is estimated 650 time steps prior to the bifurcation.

A.1.3 Greenhouse-icehouse transition with suboptimal AR(p) order p

The results of the greenhouse-icehouse transition remain stable for suboptimal AR(p) orders p. In Figure A.4 the increasing dominant eigenvalue trend is shown exemplarily for an order of $p = 3$ and $p = 7$. The optimal order is $p = 5$.



(a) Greenhouse data: DEV with AR order $p = 3$.

(b) Greenhouse data: DEV with AR order $p = 7$.

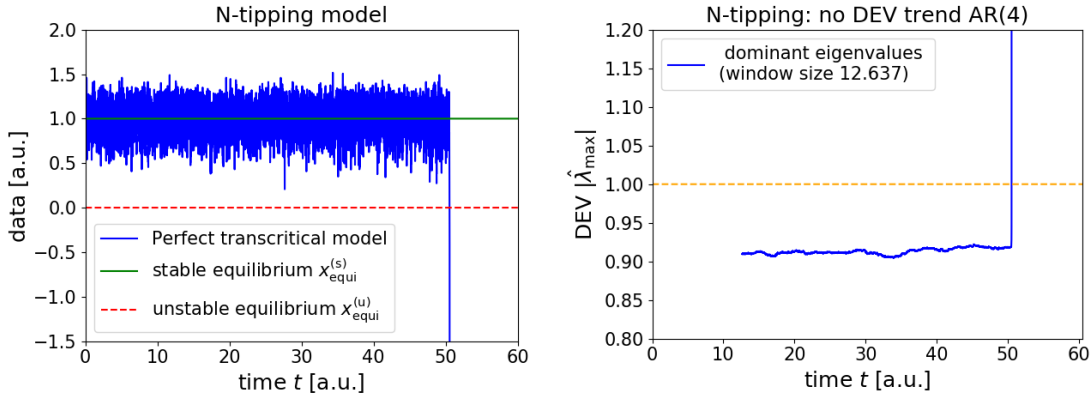
Figure A.4: The positive dominant eigenvalue trends remains stable for suboptimal orders $p = 3, 7$ with the AR(p) based method. The optimal order is $p = 5$.

A.1.4 Leading indicators in N-tipping systems

The simulated N-tipping dataset, presented in A.5(a), is calculated using the Euler-Maruyama method, described in subsection 3.4.1, and a perfect transcritical normal form, written as

$$\dot{x} = x(r - x) + \nu \mathrm{d}W. \tag{A.1}$$

The parameters are chosen to be $r = \nu = 1$ and a Gaussian white noise process $\mathrm{d}W$ with a standard deviation $\sigma = 0.2$ is added in order to enable the system to leave the stable state, marked by the green solid line in A.5(a) and switch to the unstable branch, marked by the red dashed line. The simulation is performed over a time interval $t \in [0, 10000]$ with a time step of $\mathrm{d}t = 0.1$. An analysis of the time series with the AR(p) DEV $\lambda_{\max}$ estimation approach gives no significant results as shown in A.5(b).



(a) N-tipping dataset

(b) DEV for N-tipping dataset

Figure A.5: In A.5(a) the simulation of an N-tipping dataset is presented. It is based on the perfect transcritical normal form of Equation A.1 with the parameter $r = 1$ and additive Gaussian white noise with the standard deviation $\sigma = 0.2$. It is chosen in order to provide the small possibility of driving the system that fluctuates around the stable state towards the unstable state due to a random shock. The stable state is marked by the green solid line. The unstable one is marked by the red dashed line. In A.5(b) the results of a DEV $\lambda_{\max}$ analysis are presented. As expected, the leading indicator cannot rise prior to the critical transition. The phenomenon of N-tipping is unpredictable with the DEV as leading indicator.

A.2 Data preparation for the numerical analysis

In the following subsections the parameter analyses of the investigated data of chapter 4 and chapter 5 are collected and summarized.

A.2.1 Optimal AR(p) order of the ROC dataset

In this section the results of the false nearest neighbour algorithm for the CSD ROC dataset, used in subsection 4.3.3, are shortly presented. The dataset consists of sequences of B-tipping CSD datasets. The false nearest neighbour algorithm is exemplarily applied to one sequence in order to determine the optimal embedding dimension d that is chosen as the optimal order p of the AR(p) process. The results are presented in Figure A.6.

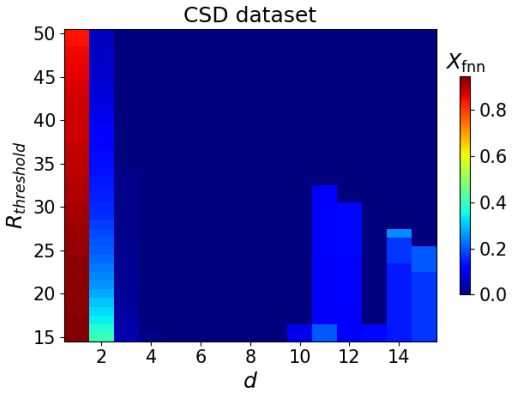
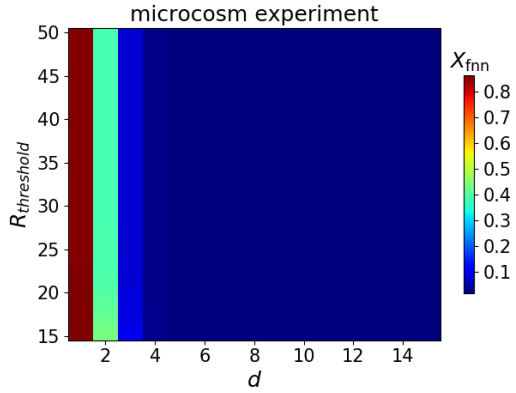


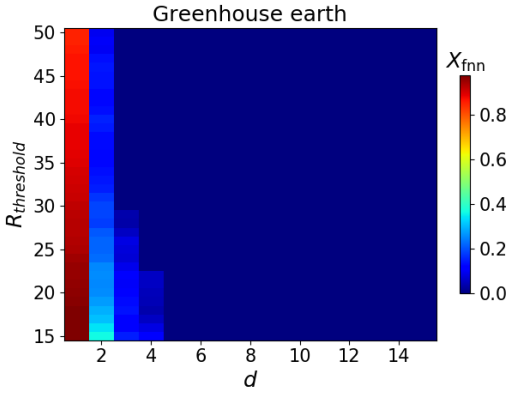
Figure A.6: The results of the false nearest neighbour algorithm are presented for one CSD dataset simulated analogously to the instructions of subsection 4.1.4. The optimal order is chosen as $p = 3$.

A.2.2 Optimal AR(p) order of the experimental data

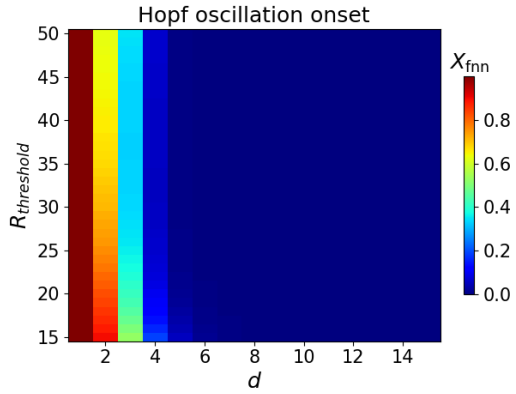
The experimental data of section 4.4 is investigated based on the optimal parameters that are determined by the false nearest neighbour algorithm. The results are presented in Figure A.7 on the next page.



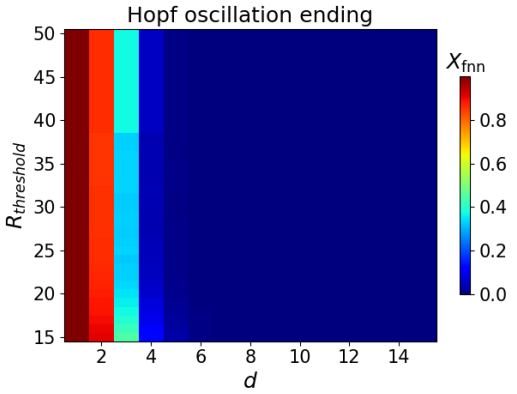
(a) Optimal AR(p) order for the Veraart experiment



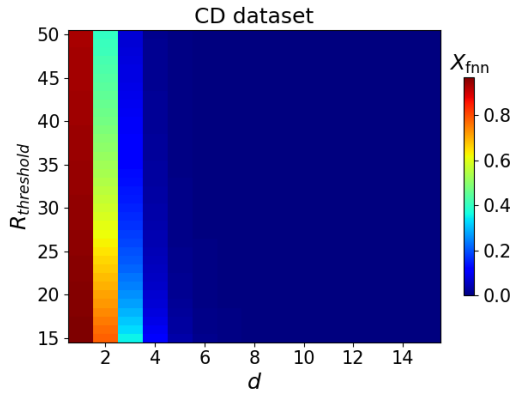
(b) Optimal AR(p) order for the greenhouse-icehouse earth transition



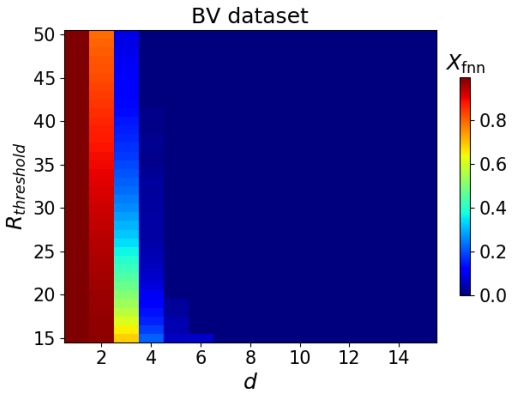
(c) Optimal AR(p) order for the physiological Hopf onset experiment



(d) Optimal AR(p) order for the physiological Hopf ending experiment



(e) Optimal AR(p) order for the CD power outage data

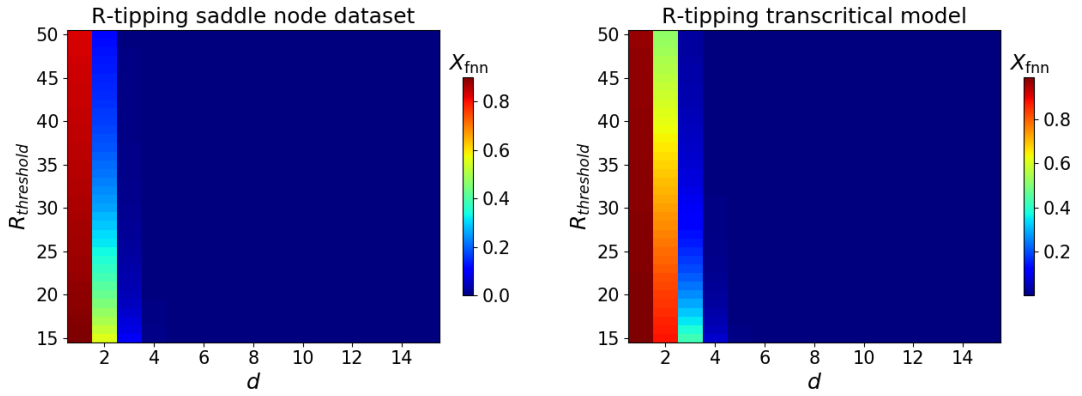


(f) Optimal AR(p) order for the BV power outage data

Figure A.7: The colormaps represent the results of the false nearest neighbour algorithm for the investigated experimental datasets of section 4.4. In each subsection of section 5.1 the optimal AR(p) order p is chosen as the optimal embedding dimension d of these colormaps.

A.2.3 Optimal AR(p) order of the R-tipping datasets

The R-tipping systems of section 5.1 are investigated based on the optimal parameters that are determined by the false nearest neighbour algorithm. The results are presented in Figure A.8.



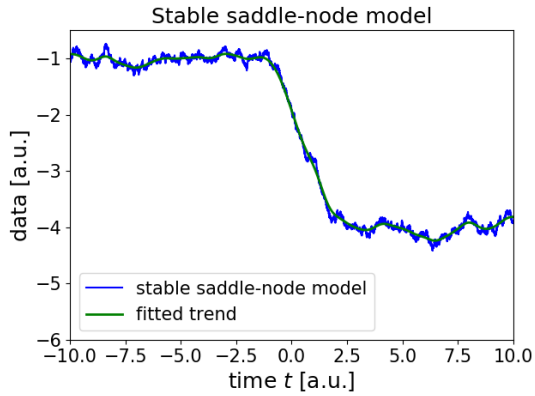
(a) Optimal AR(p) order for the R-tipping saddle-node model

(b) Optimal AR(p) order for the R-tipping transcritical model

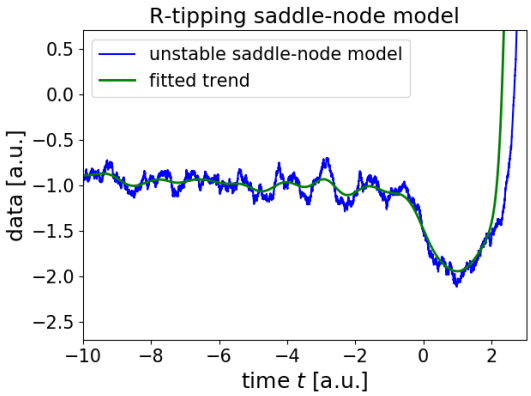
Figure A.8: The colormaps represent the results of the false nearest neighbour algorithm for the investigated R-tipping datasets of section 5.1. In each subsection of section 5.1 the optimal AR(p) order p is chosen as the optimal embedding dimension d of these colormaps.

A.2.4 Detrending of the R-tipping datasets

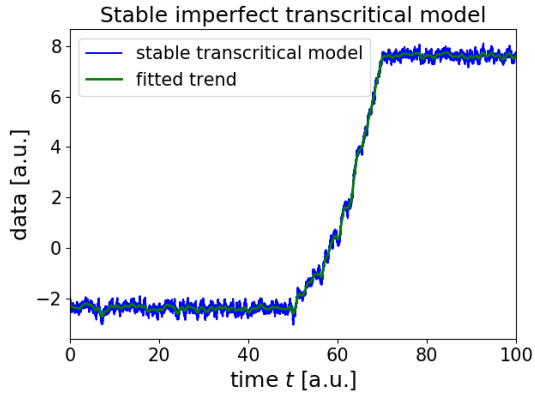
The detrended stable and R-tipping saddle-node and imperfect transcritical models of subsection 5.2.2 are obtained by subtracting a fitted slow trend from the original time series. The slow trend is calculated by convolution of the time series with a Gaussian kernel. The convolution is performed using the python routine *scipy.ndimage.filters.gaussian_filter*. The results for the fits are shown in Figure A.9.



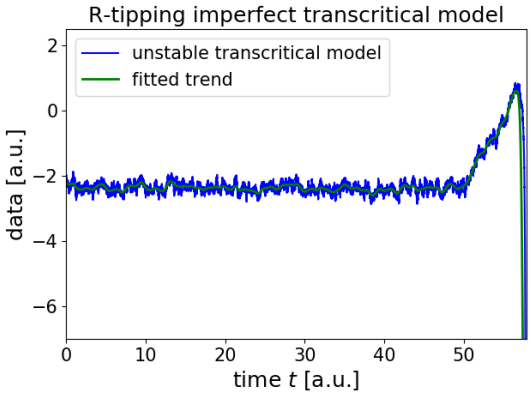
(a) Preparation of the detrended stable saddle-nodemodel



(b) Preparation of the detrended R-tipping saddle-node model



(c) Preparation of the detrended stable imperfect transcritical model



(d) Preparation of the detrended R-tipping imperfect transcritical model

Figure A.9: The stable and the R-tipping saddle-node model are shown in A.9(a) and A.9(b). The green coloured trend is fitted using a Gaussian filter with the kernel's standard deviation $\sigma = 300$ for the stable case and $\sigma = 150$ for the unstable case. The procedure is performed using the pre-implemented python routine *scipy.ndimage.filters.gaussian_filter* with the default parameter values apart from the standard deviation σ. In A.9(c) and A.9(d) the Gaussian filtering procedure is applied to the stable and R-tipping imperfect transcritical datasets with the kernel's standard deviation $\sigma = 250$. The fitted trends are subtracted from the original time series in order to obtain the detrended datasets.

Bibliography

[AAH13] O Ozgur Aybar, I Kusbeyzi Aybar, and AS Hacinliyan. "AS Stability and Bifurcation in the Hénon Map and its Generalizations". In: *Chaotic Modeling and Simulation (CMSIM)* 4 (2013), pp. 529–538.

[Ash+12] P. Ashwin et al. "Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 370.1962 (Jan. 2012), pp. 1166–1184. DOI: 10.1098/rsta.2011.0306.

[Box+15] George E. P. Box et al. *Time Series Analysis : Forecasting and Control.,Incorporated, 2015. ProQuest Ebook Central, https://ebookcentral.proquest.com/lib/ulbmuenster/detail.action?docID=2064681.* Wiley Series in Probability and Statistics Ser. John Wiley & Sons, Incorporated, 2015. ISBN: 9781118675021 (print) 9781118674918 (e-book). URL: https://ebookcentral.proquest.com/lib/ulbmuenster/detail.action?docID=2064681.

[Bro+08] I. N. Bronstein et al. *Taschenbuch der Mathematik*. 7th ed. Frankfurt a. M.: Harri Deutsch, 2008.

[CB11] S. R. Carpenter and W. A. Brock. "Early warnings of unknown nonlinear shifts: a nonparametric approach". In: *Ecology* 92.12 (Dec. 2011), pp. 2196–2201. DOI: 10.1890/11-0716.1.

[Cen85] Robert M. Centor. "Receiver Operating Characteristic (ROC) Curve Analysis Using Microcomputer Spreadsheets". In: *Proc Annu Symp Comput Appl Med Care*. Nov. 1985, pp. 207–211.

[Cha14] Thomas Chadefaux. "Early warning signals for war in the news". In: *Journal of Peace Research* 51.1 (Jan. 2014), pp. 5–18. DOI: 10.1177/0022343313507302.

[CHD12] Eduardo Cotilla-Sanchez, Paul D. H. Hines, and Christopher M. Danforth. "Predicting Critical Transitions From Time Series Synchrophasor Data". In: *IEEE Transactions on Smart Grid* 3.4 (Dec. 2012), pp. 1832–1840. DOI: 10.1109/tsg.2012.2213848.

[Dak+08] Vasilis Dakos et al. "Slowing down as an early warning signal for abrupt climate change". In: *Proceedings of the National Academy of Sciences* 105.38 (2008), pp. 14308–14312. ISSN: 0027-8424. DOI: 10.1073/pnas.0802430105. eprint: http://www.pnas.org/content/105/38/14308.full.pdf. URL: http://www.pnas.org/content/105/38/14308.

[Dak+12] Vasilis Dakos et al. “Methods for Detecting Early Warnings of Critical Transitions in Time Series Illustrated Using Simulated Ecological Data”. In: *PLoS ONE* 7.7 (July 2012). Ed. by Bülent Yener, e41010. DOI: 10.1371/journal.pone.0041010.

[Dak+17] Vasilis Dakos et al. “Elevated nonlinearity as an indicator of shifts in the dynamics of populations under stress”. In: *Journal of The Royal Society Interface* 14.128 (Mar. 2017), p. 20160845. DOI: 10.1098/rsif.2016.0845.

[Dao07] Yair Daon. *Simplex and S-map algorithms*. Nov. 2007. URL: https://cran.r-project.org/web/packages/rEDM/vignettes/rEDM-algorithms.pdf.

[Dey+16] E. R. Deyle et al. “Tracking and forecasting ecosystem interactions in real time.” In: *Proc. R. Soc. B*. Vol. 283: 20152258. 2016. URL: http://dx.doi.org/10.1098/rspb.2015.2258.

[DM04] Volker Dose and Annette Menzel. “Bayesian analysis of climate change impacts in phenology”. In: *Global Change Biology* 10.2 (Feb. 2004), pp. 259–272. DOI: 10.1111/j.1529-8817.2003.00731.x.

[DM06] Volker Dose and Annette Menzel. “Bayesian correlation between temperature and blossom onset data”. In: *Global Change Biology* 12.8 (Aug. 2006), pp. 1451–1459. DOI: 10.1111/j.1365-2486.2006.01160.x.

[Ehe17] Frank Ehebrecht. “Anticipation of critical transitions in complex systems”. MA thesis. Westphälische Wilhelms-Universität Münster, 2017.

[Ger10] Christian Gerthsen. *Gerthsen Physik*. Ed. by Dieter Meschede. Springer-Lehrbuch. Springer-Verlag Berlin Heidelberg, 2010.

[Gse+16] Alena Sonia Gsell et al. “Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems”. In: *Proceedings of the National Academy of Sciences* 113.50 (Nov. 2016), E8089–E8095. DOI: 10.1073/pnas.1608242113.

[Has07] Uwe Hassler. *Stochastische Integration und Zeitreihenmodellierung*. Springer Berlin Heidelberg, 2007. DOI: 10.1007/978-3-540-73568-7.

[Hor07] Ray Horak. *Webster’s New World Telecom Dictionary*. Hungry Minds Inc, 2007. ISBN: 047177457X.

[Izr+00] Igor Izrailtyan et al. “Early detection of acute allograft rejection by linear and nonlinear analysis of heart rate variability”. In: *The Journal of Thoracic and Cardiovascular Surgery* 120.4 (2000), pp. 737–745. ISSN: 0022-5223. DOI: https://doi.org/10.1067/mtc.2000.108930. URL: http://www.sciencedirect.com/science/article/pii/S0022522300329877.

[Jur+17] Jan Jurczyk et al. “Measuring critical transitions in financial markets”. In: *Scientific Reports* 7.1 (Sept. 2017). DOI: 10.1038/s41598-017-11854-1.

[KBA92] Matthew B. Kennel, Reggie Brown, and Henry D. I. Abarbanel. "Determining embedding dimension for phase-space reconstruction using a geometrical construction". In: *Phys. Rev. A* 45 (6 Mar. 1992), pp. 3403–3411. DOI: 10.1103/PhysRevA.45.3403. URL: https://link.aps.org/doi/10.1103/PhysRevA.45.3403.

[KLZ14] Andreas Klöckner, Franciscus L. J. van der Linden, and Dirk Zimmer. "Noise Generation for Continuous System Simulation.Proceedings of the 10th International Modelica Conference - Lund, Sweden - Mar 10-12, 2014". In: *10th International Modelica Conference*. Ed. by Hubertus Tummescheit and Karl-Erik Arzén. Linköping Electronic Conference Proceedings 96. Lund, Sweden: Linköping University Electronic Press, Dec. 2014, pp. 837–846. URL: https://elib.dlr.de/88584/.

[Koh05] Wolfgang Kohn. *Statistik. Datenanalyse und Wahrscheinlichkeitsrechnung.* Springer-Verlag Berlin Heidelberg, 2005. ISBN: ISBN 3-540-21677-4.

[KP92] Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*. Springer Berlin Heidelberg, 1992. DOI: 10.1007/978-3-662-12616-5.

[KS04] Holger Kantz and Thomas Schreiber. *Nonlinear Time Series Analysis*. Second edition. Cambridge University Press, 2004. ISBN: 0251 82150 9 hardback 0251 52902 6 paperback.

[KS08] Y. A. Kuznetsov and R. J. Sacker. "Neimark-Sacker bifurcation". In: *Scholarpedia* 3.5 (2008). revision #91556, p. 1845. DOI: 10.4249/scholarpedia.1845.

[LE11] Joosup Lim and Bogdan I. Epureanu. "Forecasting a class of bifurcations: Theory and experiment". In: *Physical Review E* 83.1 (Jan. 2011). DOI: 10.1103/physreve.83.016203.

[Lee+13] I. A. van de Leemput et al. "Critical slowing down as early warning for the onset and termination of depression". In: *Proceedings of the National Academy of Sciences* 111.1 (Dec. 2013), pp. 87–92. DOI: 10.1073/pnas.1312114110.

[Liv+13] V.N. Livina et al. "Forecasting the underlying potential governing the time series of a dynamical system". In: *Physica A: Statistical Mechanics and its Applications* 392.18 (Sept. 2013), pp. 3891–3902. DOI: 10.1016/j.physa.2013.04.036.

[LKL10] V. N. Livina, F. Kwasniok, and T. M. Lenton. "Potential analysis reveals changing number of climate states during the last 60 kyr". In: *Climate of the Past* 6.1 (Feb. 2010), pp. 77–82. DOI: 10.5194/cp-6-77-2010.

[LMF15] V. N. Livina, T. M. Vaz Martins, and A. B. Forbes. "Tipping point analysis of atmospheric oxygen concentration". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25.3 (Mar. 2015), p. 036403. DOI: 10.1063/1.4907185.

[Loh12] Hans Lohninger. *Grundlagen der Statistik*. German. TU Wien. Nov. 2012. URL: http://www.statistics4u.com/fundstat_germ/ee_classifier_performance_metrics.html.

[Man+09] P. Manshour et al. “Turbulencelike Behavior of Seismic Time Series”. In: *Physical Review Letters* 102.1 (Jan. 2009). DOI: 10.1103/physrevlett.102.014101.

[RCL94] Michael T. Rosenstein, James J. Collins, and Carlo J. De Luca. “Reconstruction expansion as a geometry-based framework for choosing proper delay times”. In: *Physica D: Nonlinear Phenomena* 73.1 (1994), pp. 82–98. ISSN: 0167-2789. DOI: https://doi.org/10.1016/0167-2789(94)90226-7. URL: http://www.sciencedirect.com/science/article/pii/0167278994902267.

[Rik+16] Marcel G. M. Olde Rikkert et al. “Slowing Down of Recovery as Generic Risk Marker for Acute Severity Transitions in Chronic Diseases”. In: *Critical Care Medicine* 44.3 (Mar. 2016), pp. 601–606. DOI: 10.1097/ccm.0000000000001564.

[RS16] Paul Ritchie and Jan Sieber. “Early-warning indicators for rate-induced tipping”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 26.9 (Sept. 2016), p. 093116. DOI: 10.1063/1.4963012.

[RS17] Paul Ritchie and Jan Sieber. “Probability of noise- and rate-induced tipping”. In: *Physical Review E* 95.5 (May 2017). DOI: 10.1103/physreve.95.052209.

[RW15] Hui Ren and David Watts. “Early warning signals for critical transitions in power systems”. In: *Electric Power Systems Research* 124 (July 2015), pp. 173–180. DOI: 10.1016/j.epsr.2015.03.009.

[Sch+09] Marten Scheffer et al. “Early-warning signals for critical transitions”. In: *Nature* 461.7260 (Sept. 2009), pp. 53–59. DOI: 10.1038/nature08227.

[Sch+12] Marten Scheffer et al. “Anticipating Critical Transitions”. In: *Science* 338.6105 (2012), pp. 344–348. ISSN: 0036-8075. DOI: 10.1126/science.1225244. eprint: http://science.sciencemag.org/content/338/6105/344.full.pdf. URL: http://science.sciencemag.org/content/338/6105/344.

[Sta+97] J. Stark et al. “Takens embedding theorems for forced and stochastic systems”. In: *Nonlinear Analysis: Theory, Methods & Applications* 30.8 (1997). Proceedings of the Second World Congress of Nonlinear Analysts, pp. 5303–5314. ISSN: 0362-546X. DOI: https://doi.org/10.1016/S0362-546X(96)00149-6. URL: http://www.sciencedirect.com/science/article/pii/S0362546X96001496.

[Str15] Steven H. Strogatz. *Nonlinear dynamics and chaos. With Applications to Physics, Biology, Chemistry and Engineering.* Second edition. Westview Press, 2015.

[Sug94] George Sugihara. “Nonlinear forecasting for the classification of natural time series”. In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 348.1688 (1994). Ed. by and and, pp. 477–495. ISSN: 0962-8428. DOI: 10.1098/rsta.1994.0106. eprint: http://rsta.royalsocietypublishing.org/content/348/1688/477.full.pdf. URL: http://rsta.royalsocietypublishing.org/content/348/1688/477.

[Tab+] M.R.R. Tabar et al. “Short-Term Prediction of Mediumand Large-Size Earthquakes Based on Markov and Extended Self-Similarity Analysis of Seismic Data”. In: *Modelling Critical and Catastrophic Phenomena in Geoscience*. Springer Berlin Heidelberg, pp. 281–301. DOI: 10.1007/3-540-35375-5_11.

[Tak81] Floris Takens. “Detecting strange attractors in turbulence”. In: *Dynamical Systems and Turbulence, Warwick 1980*. Ed. by David Rand and Lai-Sang Young. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981, pp. 366–381. ISBN: 978-3-540-38945-3.

[Ver+11] Annelies J. Veraart et al. “Recovery rates reflect distance to a tipping point in a living system”. In: *Nature* 481.7381 (Dec. 2011), pp. 357–359. DOI: 10.1038/nature10723.

[XCM10] Baogui Xin, Tong Che, and Junhai Ma. “Neimark-Sacker Bifurcation in a Discrete-Time Financial System”. In: *Discrete Dynamics in Nature and Society* 2010 (2010), p. 12.

[ZKH15] Xiaozhu Zhang, Christian Kuehn, and Sarah Hallerberg. “Predictability of critical transitions”. In: *Phys. Rev. E* 92 (5 Nov. 2015), p. 052905. DOI: 10.1103/PhysRevE.92.052905. URL: https://link.aps.org/doi/10.1103/PhysRevE.92.052905.

[ZW06] Eric Zivot and Jiahui Wang. *Modeling Financial Time Series with S-PLUS*. Vol. 2. Springer-Verlag New York, 2006. ISBN: 978-0-387-32348-0 (e-book) 978-0-387-27965-7 (Softcover). DOI: 10.1007/978-0-387-32348-0.

Register of illustrations

Acknowledgements

When I look back on the time prior to my Master's studies in physics, I am glad that I have chosen the University of Münster. I was especially happy with the studies' conception as well as with numerous lectures. So, for example, the module "Nonlinear modelling in natural sciences" or the didactic organization of the subsidiary subject "Geophysics" with remarkably good lectures in "Fluid dynamics", lectured by Prof. Dr. U. Hansen, and "Advanced geophysics II", lectured by Dr. S. Stellmach, AOR.
Besides, I found a topic for my Master's thesis that met exactly my expectations. I have been fascinated by the mathematical understanding of complex dynamical systems and the practical, analytical approach to understand these systems based on time series analysis for some time. I want to thank my supervisors Dr. O. Kamps and PD Dr. Gurevich as well as Prof. Dr. U. Thiele, the leader of the research group, for the opportunity of writing my Master's thesis in this area of research. Furthermore, I want to thank Dr. O. Kamps for the time he spent with me on scientific discussions.
I like to remember making the acquaintance of Tobias Frohoff-Hülsmann, Tobias Schemmelmann and Felix Tabbert during the physics workshop in Zaferna. At this point I want to thank again Tobias Frohoff-Hülsmann and Dominik Wittmann, a good friend for many years, for proofreading the english text. I also want to thank my colleagues Jonas Brink, Frederik Edens and Daniel Lahrmann for a good working atmosphere and inspiring scientific discussions.
Special thanks go to my family, in particular my mother Ruth, my sister Stephanie and my brother Jonas Heßler, his girlfriend Alisa Weiland as well as my grandparents Ute Mieles and Helmuth Schmidt. Besides, special thanks go to my friends Marie-Christin Heißenbüttel, André Dachwitz, Felicitas Landau, Jan Kampik, Alejandro Jurado, Johanna Seelbach and Theresa Menacher as well as Bianca Schmitt and Leon Dörfler for the private support and the good time we spent together. Furthermore, I want to thank my friends Michela, Paolo and Caterina Gargiani for their visit that developed into a small holiday and made the final stage of my work more comfortable. I feel as close to them as to my family.
Finally, in continuation of the acknowledgements of my Bachelor's thesis I want to remember my father Joachim Heßler who always supported me, gave good advice to me and always found the right balance of paternal care and seriousness. Especially during the last two years, I noticed how much I miss him. I am grateful for every memory with him.

Danksagung

Wenn ich auf die Zeit vor Beginn meines Master-Studiums in Physik zurückblicke, bin ich froh, dass meine Wahl auf die Westphälische Wilhelms-Universität Münster gefallen ist. Sowohl der Studienaufbau als auch viele der Module stießen auf mein besonderes Interesse. So zum Beispiel die „Nichtlineare Modellierung in den Naturwissenschaften“ oder die Gestaltung des Nebenfachs „Geophysik“mit außerordentlich guten Vorlesungen zur „Fluiddynamik“ bei Prof. Dr. U. Hansen und „Geophysik für Fortgeschrittene II“ bei Dr. S. Stellmach, AOR.
Außerdem habe ich ein Thema für die Masterarbeit gefunden, das genau meinen Vorstellungen entsprach. Das mathematische Verständnis komplexer dynamischer Systeme und der praktisch analytische Ansatz zum Verständnis derselben auf Basis von Zeitreihenanalyse faszinieren mich seit längerer Zeit. Bei meinen Betreuern Dr. O. Kamps und PD Dr. S. Gurevich sowie dem AG-Leiter Prof. Dr. U. Thiele möchte ich mich recht herzlich für die Möglichkeit bedanken, meine Masterarbeit auf diesem Gebiet zu verfassen. Darüberhinaus möchte ich mich bei Dr. O. Kamps für die Zeit bedanken, die er sich für wissenschaftliche Diskussionen mit mir genommen hat.
In besonders guter Erinnerung werden mir auch die Bekanntschaften mit Tobias Frohhoff-Hülsmann, Tobias Schemmelmann und Felix Tabbert auf dem Physik-Workshop auf der Zaferna-Hütte bleiben. An dieser Stelle möchte ich mich auch noch einmal für die anregenden, wissenschaftlichen Diskussionen sowie das Korrekturlesen des englischen Textes bei Tobias Frohoff-Hülsmann und meinem langjährigen Freund Dominik Wittmann bedanken. Außerdem möchte ich mich bei meinen Bürokollegen Jonas Brink, Frederik Edens und Daniel Lahrmann für eine angenehme Arbeitsatmosphäre und anregende wissenschaftliche Diskussionen bedanken.
Besonderer Dank gilt auch meiner Familie, vor allem meiner Mutter Ruth, meiner Schwester Stephanie und meinem Bruder Jonas Heßler, seiner Freundin Alisa Weiland sowie meinen Großeltern Ute Mieles und Helmuth Schmidt. Darüberhinaus möchte ich mich besonders bei meinen Freundinnen und Freunden Marie-Christin Heißenbüttel, André Dachwitz, Felicitas Landau, Jan Kampik, Alejandro Jurado, Johanna Seelbach und Theresa Menacher sowie Bianca Schmitt und Leon Dörfler für die private Unterstützung und viele schöne gemeinsame Stunden bedanken. Außerdem möchte ich mich ganz herzlich für den lieben Besuch meiner Freunde Michela, Paolo und Caterina Gargiani bedanken, der mir zu einem kleinen Urlaub wurde und mir die Endphase der Arbeit angenehmer gemacht hat. Ich fühle mich ihnen in Gedanken nah wie meiner Familie.
Zuletzt möchte ich mich anknüpfend an die Danksagung zu meiner Bachelor-Arbeit meines Vaters Joachim Heßler erinnern, der mir Zeit seines Lebens unterstützend und mit immer offenem Ohr und gutem Rat zur Seite stand und das rechte Maß zwischen Fürsorge und väterlichem Ernst zu halten wusste. Gerade in den letzten zwei Jahren merke ich, wie sehr er mir fehlt. Ich bin für jede gemeinsame Erinnerung dankbar.

Declaration of Authorship

I hereby declare that this Master's thesis, that deals with "Leading indicators in B- and R-tipping systems with focus on eigenvalue estimation", is my own unaided work, that no other direct and indirect sources and tools than the quoted ones are used and that the parts of the Master's thesis that are literally or logically taken from other works - also electronic media - are definitely marked as borrowing with indication of the source.

Münster, 26th September 2018
place, date

Martin Heßler

I give my consent to the comparison of this Master's thesis with other texts in order to exclude plagiarism. Furthermore, I give my consent to saving the Master's thesis in a database for this purpose.

Münster, 26th September 2018
place, date

Martin Heßler

Plagiatserklärung der/ des Studierenden

Hiermit versichere ich, dass die vorliegende Arbeit über „Leading indicators in B- and R-tipping systems with focus on eigenvalue estimation“ selbstständig verfasst worden ist, dass keine anderen Quellen und Hilfsmittel als die angegebenen benutzt worden sind und dass die Stellen der Arbeit, die anderen Werken - auch elektronischen Medien - dem Wortlaut oder Sinn nach entnommen wurden, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht worden sind.

Münster, 26.09.2018
Ort, Datum

Martin Heßler

Ich erkläre mich mit einem Abgleich der Arbeit mit anderen Texten zwecks Auffindung von Übereinstimmungen sowie mit einer zu diesem Zweck vorzunehmenden Speicherung dieser Arbeit in einer Datenbank einverstanden.

Münster, 26.09.2018
Ort, Datum

Martin Heßler