

**Technische Universität Berlin**

Institut für Mathematik  
Fachgebiet Technomathematik

Fakultät II - Mathematik und Naturwissenschaften  
Skr. BEL 1  
Marchstr. 6  
10587 Berlin



Master Thesis

**Data-driven approaches to study the dynamical  
stability of the stably stratified boundary layer**

Amandine Kaiser

Matriculation Number: 386024  
02.07.2019

Supervised by  
Jun.-Prof. Dr. Nikki Vercauteren

Assistant Supervisor  
Prof. Dr. Wilhelm Stannat



Hereby I declare that I wrote this thesis myself with the help of no more than the mentioned literature and auxiliary means.

Berlin, 02.07.2019

A handwritten signature in blue ink, appearing to read 'A. Kaiser', positioned above a horizontal dotted line.

*(Amandine Kaiser)*



# Zusammenfassung

In vielen experimentellen und natürlichen Systemen treten kritische Übergänge, d.h. abrupte Wechsel von einem dynamischen Zustand in einen anderen, auf. Zum Beispiel in Klimasystemen können in der untersten Schicht der Atmosphäre, genauer gesagt der atmosphärischen Grenzschicht, abrupte Übergänge zwischen gänzlich turbulenten Zuständen und stabilen festgestellt werden. Solche abrupten Übergänge werden insbesondere in Polarregionen und in der Nacht beobachtet. Sie sind äußerst relevant für den Grad der Durchmischung mit den höheren Ebenen der Grenzschicht. Um die stabil geschichtete Grenzschicht analysieren zu können verwenden viele Ansätze eine Klassifizierung der Zustände in der Grenzschicht in schwach und sehr stabile Zustände. Aus diesem Grund ist eine Methode zur Erkennung der Übergänge unerlässlich.

In dieser Masterarbeit werden diese Übergänge in der stabil geschichteten Grenzschicht der Atmosphäre mit Hilfe von verschiedenen Methoden der statistischen Modellierung und aus dem Bereich der Dynamischen Systeme untersucht. Die Grundlage für die Analyse bilden ein Indikator für die Stabilität (Resistenz gegenüber Störungen) und ein Modell für Übergänge von Temperaturinversionen in Bodennähe in der Nacht und in arktischen Konditionen. Ein besonderer Fokus liegt dabei auf Bifurkationspunkte, Punkte in denen sich die Stabilität des Systems drastisch verändert. Die Zuverlässigkeit des Stabilitätsindikators wird mittels simulierter Daten und echter Daten, gemessen in der Nacht und in der Antarktis, überprüft. Die Resultate zeigen, dass die Übergänge zuverlässig erkannt werden.



# Abstract

Many experimental or natural systems undergo critical transitions, i.e. sudden shifts from one dynamical regime to another. For example in the climate system, the lower atmospheric layer, namely the atmospheric boundary layer, can experience sudden transitions between fully turbulent states and stable, quasi-laminar states. Such rapid transitions are observed in Polar regions and at night, and have important consequences in the level of mixing with the higher levels of the atmosphere. To analyse the stable boundary layer many approaches rely on the identification of regimes, i.e. weakly and very stable regimes. Therefore, it is crucial to detect the transitions between the regimes.

In this master thesis a combination of methods from dynamical systems and statistical modelling are applied to study these regime transitions. The analysis is based on an indicator for the dynamical stability (i.e. the resilience to perturbations) and a conceptual model for regime transitions of near-surface temperature inversion at night as well as in Arctic conditions. A focus lies on bifurcation points in the dynamics, points in which the stability of the system changes drastically. The performance of the stability indicator is assessed by applying it to simulated and observation data, provided from nighttime and Polar meteorological measurements.



# Contents

<b>Zusammenfassung</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>1. Introduction</b>	<b>1</b>
<b>2. Statistics / ARMA(p,q) Processes</b>	<b>3</b>
2.1. Introduction to ARMA Processes . . . . .	3
2.2. Modeling with ARMA Processes . . . . .	9
<b>3. Regime Transitions in Near-Surface Temperature Inversions: A Conceptual Model</b>	<b>15</b>
3.1. Derivation of the Solution of the Simplified Model . . . . .	17
3.1.1. Analytic Derivation . . . . .	18
3.1.2. Numerical Integration . . . . .	21
3.2. Bifurcation Analysis . . . . .	23
<b>4. Dynamical Stability of Near-Surface Temperature Inversions</b>	<b>33</b>
4.1. The Stability Indicator . . . . .	33
4.2. Dependence of $\Upsilon$ on the window length $\tau$ . . . . .	38
4.2.1. Approach 1: K-Means . . . . .	39
4.2.2. Approach 2: Anderson-Darling Normality Test . . . . .	39
4.3. Alternative Definitions for a Stability Indicator . . . . .	40
<b>5. Testing of <math>\Upsilon</math> on Controlled and Observational Data</b>	<b>43</b>
5.1. Simplified Model by van de Wiel with Added Noise . . . . .	43
5.2. Dumosa . . . . .	50
5.3. Dome C Data . . . . .	61
<b>6. Outlook</b>	<b>65</b>
<b>7. Conclusion</b>	<b>71</b>
<b>8. Acknowledgement</b>	<b>73</b>
<b>A. BIC vs AIC vs AICc</b>	<b>75</b>

<b>B. Properties of Equation 3.1</b>	<b>79</b>
B.1. Continuity . . . . .	79
B.2. Differentiability . . . . .	79
<b>C. Proof: Linear Process with Gaussian White Noise is Stationary and Gaussian     Itself</b>	<b>81</b>
<b>List of Figures</b>	<b>85</b>
<b>List of Tables</b>	<b>89</b>
<b>Bibliography</b>	<b>91</b>

# 1. Introduction

The atmospheric boundary layer (BL) is the part of the troposphere that is directly influenced by the earth's surface [Stull (2009)]. It is the area in which we live and most human activities take place. The list of direct and indirect influences of the atmospheric boundary layer (BL) on our lives is very long. Therefore, a deep knowledge of the processes that take place therein are important for various research areas and are of high interest for us.

The BL thickness varies between hundreds of meters and a few kilometres. One of its key characteristics is the change of its structure with the diurnal cycle. In figure 1.1 the diurnal evolution of the BL in high pressure regions over land is shown. At sunrise

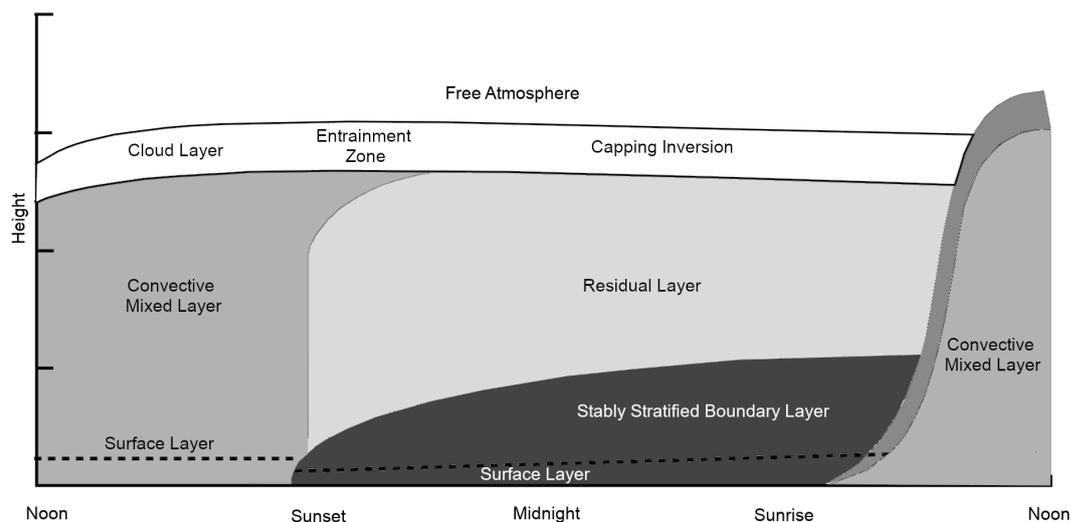


Figure 1.1.: diurnal evolution of the BL [on the basis of graphic by Stull (2009)]

a turbulent mixed layer starts to form due to the heating of the grounds and begins to grow in depth. Its maximum depth is reached in the late afternoon and about half an hour before sunset turbulence decays leaving a residual layer in place of the mixed layer. As night progresses, the bottom portion of the residual layer is transformed by its contact with the ground into a stable boundary layer (SBL). Generally speaking, the boundary layer can become stably stratified by the advection of warm air over a colder surface. The density of air changes with the variation of its temperature. Thus, in the

SBL there is a denser flow at the bottom and layered above is a less dense one. This is called a stable stratification. As a result of long-wave radiative cooling a stably stratified boundary layer often forms at night over land, where it is known as nocturnal boundary layer (NBL), or in Arctic conditions due to the presence of cold snow and the ice surface. If the density stratification is great enough near the the surface to cause temperatures to increase with height, then that part of the SBL is classified as a temperature inversion.

The dynamics in the SBL are very complex and as a result hard to describe and model. Nonetheless, several approaches have been made. The approaches generally rely on identifying multiple regimes of the dynamics in the SBL. It is often broadly classified in weakly stable and very stable regimes. We use the definitions by [Mahrt \(2014\)](#): *"The weakly stable regime usually includes a well-defined boundary layer in which the turbulence decreases with height and becomes small at the top of the boundary layer. The turbulence is relatively continuous in both time and space. Weakly stable conditions occur with either cloud cover or significant airflow.[...]The very stable regime occurs with strong stratification and weak winds and does not follow the traditional concept of a boundary layer. For example, the turbulence may increase with height and reach a maximum in a layer only intermittently coupled to the surface. The very stable regime will include a variety of different scenarios and vertical structures, and a unifying conceptual picture is not available."* There exist more detailed classification schemes which may include a transition regime between the weakly and very stable regime and an extremely stable regime. We focus on weakly and very stable regimes.

Van de Wiel et. al. define a conceptual model in their paper [[van de Wiel et al. \(2017\)](#)] to understand the regime transitions of near-surface temperature inversions at night as well as in Arctic conditions. Moreover, Nevo et. al. define in their paper [[Nevo et al. \(2017\)](#)] an indicator for the hydro-dynamical stability of the SBL using a combination of dynamical systems concepts and stochastic processes tools. We concentrate on these two approaches.

This master thesis is organized as follows: In chapter 2 the essential concepts of autoregressive moving average, ARMA(p,q), modelling, which are needed in the following chapters, are introduced. Then, in chapter 3 the model for the regime transitions and its equilibrium properties are analysed. In chapter 4 the stability indicator is introduced. Moreover, we analyse its mathematical background and we explain how a window length which needs to be defined to calculate the stability indicator can be chosen. Thereafter, in chapter 5 we apply the stability indicator to controlled data generated by the conceptual model for regime transitions and two observational data sets. Finally, in chapter 6 we give an outlook for another approach to choose the window length.

## 2. Statistics / ARMA(p,q) Processes

This chapter gives an overview over the statistical concepts, especially ARMA(p,q) related ones, which are needed in this master thesis. In particular, in chapter 4 ARMA(p,q) processes play an important role.

### 2.1. Introduction to ARMA Processes

First of all a time series is the same as a stochastic process indexed by integers. The term "time series" is nonetheless often used to refer to the realization of a time-series process. Because it will be clear from the context of the discussion the term time series will be used in this thesis to refer to both the processes and the realization.

**Definition 2.1.1. (stochastic process)** [Klenke (2008)]

Let  $I \subset \mathbb{R}$ . A family of random variables  $X = \{x_t | t \in I\}$  (on the probability space  $(\Omega, \mathcal{F}, Pr)$  where  $\Omega$  is the sample space,  $\mathcal{F}$  a filtration and  $Pr$  a probability measure) with values in  $(E, \mathcal{E})$  is called a stochastic process with index set  $I$  and range  $E$ .

If not specified otherwise the random variables will be real-valued, i.e.  $(E, \mathcal{E}) = (\mathbb{R}, \mathbb{B}(\mathbb{R}))$  with  $\mathbb{B}(\mathbb{R})$  being the Borel  $\sigma$ -algebra on  $\mathbb{R}$ . To shorten notation we write  $\{x_t\}$  for the stochastic process.

Often there exists some sort of regularity in the behaviour of a time series. The notion of regularity is introduced by the concept of stationarity.

**Definition 2.1.2. (strictly stationary)** [Shumway (2017)]

A strictly stationary time series is one for which the probabilistic behaviour of every collection of variables

$$\{x_{t_1}, x_{t_2}, \dots, x_{t_k}\}$$

is identical to that of the time shifted set

$$\{x_{t_1+h}, x_{t_2+h}, \dots, x_{t_k+h}\}.$$

That is

$$Pr\{x_{t_1} \leq c_1, \dots, x_{t_k} \leq c_k\} = Pr\{x_{t_1+h} \leq c_1, \dots, x_{t_k+h} \leq c_k\}$$

for all  $k = 1, 2, \dots$ , all time points  $t_1, t_2, \dots, t_k$ , all numbers  $c_1, c_2, \dots, c_k$  and all time shifts  $h = 0, \pm 1, \pm 2, \dots$ .

This version of stationarity is often too strong and it is difficult to assess with only one data set. Hence, a milder version is commonly used.

**Definition 2.1.3. (weakly stationary)** [Shumway (2017)]

A weakly stationary time series,  $\{x_t\}$  is a finite-variance process, i.e.  $var(x_t) < \infty$ , such that

- i. the mean value function,  $\mu_t = E(x_t) = \int_{-\infty}^{\infty} x f_t(x) dx$ , is constant and does not depend on time  $t$ , and
- ii. the autocovariance function,  $\gamma(s, t) = cov(x_s, x_t) = E[(x_s - \mu_s)(x_t - \mu_t)]$ , depends on  $s$  and  $t$  only through their difference  $|s - t|$ .

If it is not specified otherwise the term stationary will be used instead of weakly stationary.

A strictly stationary process with finite second moments i.e.  $E[x_t^2] < \infty$ , is stationary but the converse is not true. There is one important case however in which stationarity implies strict stationarity.

**Definition 2.1.4. (Gaussian time series)**[Brockwell (1991)]

The process  $\{x_t\}$  is a Gaussian time series if and only if the distribution functions of  $\{x_t\}$  are all multivariate normal.

By defining the  $n$ -dimensional vector  $X = (x_{t_1}, \dots, x_{t_n})^T$ , the  $n \times 1$  mean vector  $\mu = (\mu_{t_1}, \dots, \mu_{t_n})^T$  and the  $n \times n$  covariance matrix  $\Sigma = \{\gamma(t_i, t_j) | i, j = 1, \dots, n\}$ , which is assumed to be positive-definit, the multivariate density function of  $X \in \mathbb{R}$  can be written as

$$f(X) = (2\pi)^{-\frac{n}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(X - \mu)^T \Sigma^{-1} (X - \mu)\right)$$

where  $|\cdot|$  denotes the determinant.

If  $\{x_t\}$  is a stationary Gaussian process  $\{x_t\}$  is also strictly stationary, since for all  $n \in \mathbb{N}$  and for all  $h, t_1, t_2, \dots \in \mathbb{Z}$ , the random vectors  $(x_{t_1}, \dots, x_{t_n})^T$  and  $(x_{t_1+h}, \dots, x_{t_n+h})^T$  have

the same mean and covariance matrix and, hence the the same distribution.

To define an autoregressive model it is essential to explain white noise first. White noise is a time series generated from a collection of uncorrelated <sup>1</sup> random variables  $w_t$ , with mean 0 and finite variance  $\sigma_w^2$ . The process is denoted as  $\{w_t\} \sim wn(0, \sigma_w^2)$ . A white noise process is stationary but not necessarily strictly stationary. Moreover, Gaussian white noise is independent. Because in the case of Gaussianity uncorrelated random variables are also independent.

**Definition 2.1.5. (autoregressive model)** [Shumway (2017)]

An autoregressive model of order  $p$ , abbreviated AR( $p$ ), is of the form

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t \quad (2.1)$$

$$= \sum_{i=1}^p \phi_i x_{t-i} + w_t \quad (2.2)$$

where  $\{x_t\}$  is stationary,  $\{w_t\}$  is white noise and  $\phi_1, \phi_2, \dots, \phi_p$  are constants ( $\phi_p \neq 0$ ). We can assume that  $E(x_t) = \mu = 0$  since we can otherwise replace  $x_t$  by  $x_t - \mu$ . Hence,

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \dots + \phi_p(x_{t-p} - \mu) + w_t$$

or one can write

$$x_t = \alpha + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t$$

where  $\alpha = \mu(1 - \phi_1 - \dots - \phi_p)$ .

**Note:** If  $\mu = 0$  it follows directly that  $\alpha = 0$ . Furthermore, it holds that

$$\begin{aligned} \mu &= E[x_t] \\ &= E[\alpha + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t] \\ &= \alpha + \sum_{i=1}^p \phi_i \mu \\ &= \mu - \mu \sum_{i=1}^p \phi_i + \mu \sum_{i=1}^p \phi_i \\ &= \mu. \end{aligned}$$

Thus, this is well defined.

<sup>1</sup>That means  $\rho(s, t) = \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}} = 0$  for  $s \neq t$ , with  $\rho(s, t)$  being the autocorrelation function.

**Example 2.1.6. (the AR(1) model)** [Shumway (2017)]

The AR(1) model is given by

$$x_t = \phi x_{t-1} + w_t. \quad (2.3)$$

Provided that  $|\phi| < 1$  and  $\sup_t \text{var}(x_t) < \infty$ , an AR(1) model can be represented as a linear process given by

$$x_t = \sum_{i=0}^{\infty} \phi^i w_{t-i}. \quad (2.4)$$

This can be shown by iterating equation 2.3. The mean of the process defined in 2.4 is

$$\begin{aligned} E[x_t] &= \sum_{i=0}^{\infty} \phi^i E[w_{t-i}] \\ &= 0 \end{aligned}$$

and the autocovariance function is

$$\begin{aligned} \gamma(h) &= \text{cov}(x_{t+h}, x_t) \\ &= E\left[\left(\sum_{j=0}^{\infty} \phi^j w_{t+h-j}\right)\left(\sum_{k=0}^{\infty} \phi^k w_{t-k}\right)\right] \\ &= E[(w_{t+h} + \dots + \phi^h w_t + \phi^{h+1} w_{t-1} + \dots)(w_t + \phi w_{t-1} + \dots)] \\ &\stackrel{1}{=} E[(\phi^h w_t)w_t + (\phi^{h+1} w_{t-1})(\phi w_{t-1}) + \dots] \\ &= \text{var}(w_t)\phi^h + \text{var}(w_{t-1})\phi^{h+1}\phi + \dots \\ &= \sigma_w^2 \sum_{j=0}^{\infty} \phi^{h+j} \phi^j \\ &= \sigma_w^2 \phi^h \sum_{j=0}^{\infty} \phi^{2j} \\ &\stackrel{2}{=} \frac{\sigma_w^2 \phi^h}{1 - \phi^2}, \text{ for } h \geq 0. \end{aligned}$$

Hence, the process given by 2.4 is indeed stationary.

**Definition 2.1.7. (moving average model)** [Shumway (2017)]

The moving average model of order  $q$ , or MA( $q$ ) model, is defined as

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \dots + \theta_q w_{t-q} \quad (2.5)$$

where  $\{w_t\} \sim \text{wn}(0, \sigma_w^2)$  and  $\theta_1, \theta_2, \dots, \theta_q$  ( $\theta_q \geq 0$ ) are constants.

<sup>1</sup>Note:  $0 = \rho(j, k) = \frac{\text{cov}(x_j, x_k)}{\sqrt{\text{var}(x_j)\text{var}(x_k)}} = \frac{\text{cov}(x_j, x_k)}{\sigma_w^2}$  and that implies  $E[w_j w_k] = 0$  for  $j \neq k$

<sup>2</sup>geometric series

Unlike the autoregressive process the moving average process is stationary for any values of the parameters  $\theta_1, \dots, \theta_q$ .

**Definition 2.1.8. (ARMA(p,q))** [Shumway (2017)]

A time series  $\{x_t, t = 0, \pm 1, \pm 2, \dots\}$ , is an ARMA(p,q) process if it is stationary and

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \dots + \theta_q w_{t-q} \quad (2.6)$$

with  $\phi_p \geq 0$ ,  $\theta_q \geq 0$  and  $\sigma_w^2 > 0$ . The parameters p and q are called the autoregressive and moving average orders, respectively. If  $\{x_t\}$  has a non-zero mean  $\mu$ , we set  $\alpha = \mu(1 - \phi_1 - \dots - \phi_p)$  and write the process as  $x_t = \alpha + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \dots + \theta_q w_{t-q}$  where  $\{w_t\} \sim wn(0, \sigma_w^2)$ .

The polynomials

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p, \phi_p \neq 0$$

and

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q, \theta_q \neq 0,$$

with  $z \in \mathbb{C}$ , are required to have no common factors to avoid parameter redundant models. This ensures that the model can't be reduced to a simpler one.

It is often convenient to use the abbreviated form of 2.6

$$\phi(B)x_t = \theta(B)w_t$$

where  $\phi(\cdot)$  and  $\theta(\cdot)$  are the before mentioned polynomials and B is the backshift operator, i.e.  $B^j x_t = x_{t-j}$ ,  $B^j w_t = w_{t-j}$  for  $j = 0, \pm 1, \pm 2, \dots$ .

An important part of the Definition 2.1.8 is the requirement that  $\{x_t\}$  is stationary. The following theorem states the conditions for a unique solution of 2.6 to exist.

**Theorem 2.1.9. (existence and uniqueness)**[Brockwell (2016)]

A stationary solution  $\{x_t\}$  of equation 2.6 exists (and is also the unique stationary solution) if and only if

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \neq 0, \text{ for all } |z| = 1, z \in \mathbb{C}$$

**Note:** The region defined by the set of complex  $z$  such that  $|z| = 1$  is referred to as the unit circle.

Proof: (sketch)

The idea of the proof is to show that the solution  $x_t$  of 2.6 can be written as

$$x_t = \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}$$

if  $\phi \neq 0$  for  $|z| = 1$ . This is done by showing that  $\phi^{-1}$  exists and can be defined as the linear filter with absolutely summable coefficients

$$\chi(B) := \frac{1}{\phi(B)} = \sum_{j=-\infty}^{\infty} \chi_j B^j.$$

Applying the operator  $\chi(B)$  to both sides of  $\phi(B)x_t = \theta(B)w_t$  we get

$$\begin{aligned} x_t &= \chi(B)\phi(B)x_t \\ &= \chi(B)\theta(B)w_t \\ &= \psi(B)w_t \end{aligned}$$

where  $\psi(z) := \chi(z)\theta(z) = \sum_{j=-\infty}^{\infty} \psi_j z^j$ . It can be shown that  $x_t = \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}$  is indeed the unique solution and stationary. □

Further details are for example given in Brockwell (2016) and Brockwell (1991).

We say that the solution  $\{x_t\}$  of an ARMA process is causal or a causal function of  $\{w_t\}$  if  $\{x_t\}$  can be expressed in terms of the current and past values of  $\{w_s\}$ ,  $s \leq t$ . If  $\{w_t\}$  can be expressed in terms of  $\{x_s\}$ ,  $s \leq t$  we say the ARMA process is invertible.

**Definition 2.1.10. (causality)** [Brockwell (2016)]

An ARMA(p,q) process  $\{x_t\}$  is causal, or a causal function of  $\{w_t\}$ , if there exist constants  $\{\psi_j\}$  such that  $\sum_{j=0}^{\infty} |\psi_j| < \infty$  and

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} \text{ for all } t.$$

Causality is equivalent to the condition

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \neq 0 \text{ for all } |z| \leq 1. \quad (2.7)$$

The proof of the equivalence between causality and 2.7 is for example given in the book by Shumway (2017).

**Definition 2.1.11. (invertibility)** [Brockwell (2016)]

An ARMA(p,q) process  $\{x_t\}$  is invertible if there exist constants  $\{\pi_j\}$  such that  $\sum_{j=0}^{\infty} |\pi_j| < \infty$  and

$$w_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} \text{ for all } t.$$

Invertibility is equivalent to the condition

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q \neq 0 \text{ for all } |z| \leq 1. \quad (2.8)$$

The proof of the equivalence between invertibility and 2.7 is analogous to the one of the equivalence between causality and 2.7.

## 2.2. Modeling with ARMA Processes

In chapter 4 it will be essential to determine the best fitting ARMA(p,q) model to represent an observed stationary time series. This involves several interrelated problems like the choice of  $p$  and  $q$  (order selection) and the estimation of the mean, the coefficients  $\{\phi_i, i = 1, \dots, p\}$ ,  $\{\theta_i, i = 1, \dots, q\}$ , and the white noise variance  $\sigma_w^2$ . Especially the selection of the appropriate order of the ARMA(p,q) model is crucial for applying the method defined in chapter 4. However, there are several procedures for selecting the model which fits best. The ones used in chapter 4 are information criteria. The criteria are based on the assumption that for a fixed number of observations the successive increase of the orders  $p$  and  $q$  increases the fit of the model. In order to compensate for the criteria's tendency to over fitting a penalty is introduced. This penalty term depends on the number of free parameters and on the number of observations. The criteria uses the Maximum Likelihood estimator. This is a method to estimate an unknown value  $\theta$  which is in a given set  $\Theta$  after observing random variables  $x_1, \dots, x_n$  whose distribution can be different for different values of  $\theta$ .

**Definition 2.2.1. (likelihood function, maximum likelihood estimator)**[Dudewicz (1976)]

Let  $x_1, \dots, x_n$  be random variables (not necessarily independent or identically distributed) with joint distribution function  $F(c_1, \dots, c_n; \theta) = Pr\{x_1 \leq c_1, \dots, x_n \leq c_n; \theta\}$  where  $\theta \in \Theta$  is unknown. The likelihood function is

$$L(\theta) = \begin{cases} f(x_1, \dots, x_n; \theta) & \text{if } F \text{ has a density } f \\ p(x_1, \dots, x_n; \theta) & \text{if } F \text{ has a probability mass function } p \end{cases} \quad (2.9)$$

Any  $\hat{\theta}_{MLE} = \hat{\theta}_{MLE}(x_1, \dots, x_n) \in \Theta$  such that

$$L(\hat{\theta}_{MLE}) = \sup\{L(\theta) | \theta \in \Theta\} \quad (2.10)$$

is called a maximum likelihood estimator (MLE) of  $\theta$ .

Often the natural logarithm of  $L$  is used instead of  $L$  as it simplifies calculations. The logarithm is a monotone transformation thus the values that maximise  $L(\theta)$  also maximise  $l(\theta) := \ln(L(\theta))$ . This function is referred to as the log-likelihood function.

To calculate the maximum likelihood estimator for an ARMA(p,q) process we must make some assumptions (the same as in chapter 4). Namely, we require  $\{x_t\}$  to be a linear, causal and invertible process and  $\{w_t\}$  to be a Gaussian white noise process.

**Definition 2.2.2. (linear process)**[Brockwell (2016)]

The time series  $\{x_t\}$  is a linear process if it has the representation

$$x_t = \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}$$

for all  $t$ , where  $\{w_t\} \sim \text{wn}(0, \sigma_w^2)$  and  $\{\psi_j\}$  is a sequence of constants with

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$

Let  $\beta := (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$  and  $\theta := (\beta, \sigma_w^2)$  be the vector of unknown and to be estimated parameters and  $X := (x_1, \dots, x_n)^T$ . Under the given assumptions the admissible set for  $\beta$ ,  $\Theta$ , is given by

$$\Theta = \{\beta \in \mathbb{R}^{p+q} | \phi(z)\theta(z) \neq 0 \text{ for } |z| \leq 1, \phi_p, \theta_q \neq 0, \phi(z) \text{ and } \theta(z) \text{ have no common zeros}\}$$

The idea of the maximum likelihood method is to find  $\beta$  and  $\sigma_w^2$  so that the probability to observe  $X$  is maximal. Given the above made assumptions the likelihood function of the ARMA(p,q) model is

$$L(\theta) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2}(X - \mu)^T \Sigma^{-1}(X - \mu)\right) \quad (2.11)$$

where  $\mu$  is the expected value and  $\Sigma$  the non-singular covariance matrix of  $X$  depending on  $\theta$ .

Note: In Appendix C we show that a linear process  $\{x_t\}$  is stationary and Gaussian if  $\{w_t\}$  is Gaussian.

Without loss of generality we can also assume that  $E(x_t) = 0$ , i.e.  $\mu = 0$ , because we can easily transform the data to fit this requirement. This assumption always holds for a linear process. Thus,

$$\begin{aligned} L(\theta) &= \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2} X^T \Sigma^{-1} X\right) \\ &= \frac{1}{\sqrt{(2\pi\sigma_w^2)^n \det(\Gamma)}} \exp\left(-\frac{1}{2\sigma_w^2} X^T \Gamma^{-1} X\right), \end{aligned} \quad (2.12)$$

where  $\Gamma = \frac{1}{\sigma_w^2} \Sigma$ , and

$$\begin{aligned} l(\theta) &= \ln \left[ \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2} X^T \Sigma^{-1} X\right) \right] \\ &= -\frac{1}{2} [n \ln(2\pi) + \ln(\det(\Sigma)) + X^T \Sigma^{-1} X] \\ &= -\frac{1}{2} \left[ n \ln(2\pi\sigma_w^2) + \ln(\det(\Gamma)) + \frac{1}{\sigma_w^2} X^T \Gamma^{-1} X \right] \end{aligned} \quad (2.13)$$

Note:  $\Gamma$ , contrary to  $\Sigma$ , does not depend on  $\sigma_w^2$  and only on  $\beta$ .

The optimal value for  $\sigma_w^2$  can be derived by

$$\frac{\delta l(\theta)}{\delta \sigma_w^2} = -\frac{n}{2} \frac{1}{\sigma_w^2} + \frac{1}{(2\sigma_w^2)^2} X^T \Gamma^{-1} X = 0. \quad (2.14)$$

Solving equation 2.14 we get  $\hat{\sigma}_{w_{MLE}}^2 = \frac{1}{n} X^T \Gamma^{-1} X$ . Inserting this in equation 2.13 we get

$$\begin{aligned} l(\beta) &= -\frac{1}{2} \left[ n \ln(2\pi) + \ln(\det(\Gamma)) + \frac{1}{\frac{1}{n} X^T \Gamma^{-1} X} X^T \Gamma^{-1} X \right] \\ &= -\frac{1}{2} \left[ n \ln(2\pi) + n \ln\left(\frac{1}{n} X^T \Gamma^{-1} X\right) + \ln(\det(\Gamma)) + n \right] \end{aligned} \quad (2.15)$$

To get the maximum likelihood estimator for  $\beta$ ,  $\hat{\beta}_{MLE}$ , we need to find the maximum of this equation. But the computation of  $\det(\Gamma)$  and  $\Gamma^{-1}$  are numerically expensive and are therefore often avoided in practice. This problem can be circumvented by using an innovation algorithm. For the innovation algorithm we need the least squares predictor

$\mathbb{P}_{t-1}$  of  $x_t$  given  $x_{t-1}, \dots, x_1$  and the mean squared forecast error  $v_{t-1}$ .

$$\begin{aligned}\mathbb{P}_{t-1} &= a_{t-1,1}x_{t-1} + \dots + a_{t-1,t-1}x_1 \text{ with } \mathbb{P}_0x_1 = 0, \\ v_{t-1} &= E[(x_t - \mathbb{P}_{t-1}x_t)^2], \\ X &:= (x_1, \dots, x_n)^T \text{ and} \\ \hat{X} &:= (0, \mathbb{P}_1x_2, \dots, \mathbb{P}_{t-1}x_t)^T.\end{aligned}$$

Therefore,  $X - \hat{X} = AX$  with

$$A := \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -a_{1,1} & 1 & 0 & \dots & 0 \\ -a_{2,2} & -a_{2,1} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ -a_{t-1,t-1} & -a_{t-1,t-2} & -a_{t-1,t-3} & \dots & 1 \end{bmatrix}$$

and the inverse of  $A$  is

$$C := \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ c_{1,1} & 1 & 0 & \dots & 0 \\ c_{2,2} & c_{2,1} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ c_{t-1,t-1} & c_{t-1,t-2} & c_{t-1,t-3} & \dots & 1 \end{bmatrix}.$$

Thereby, we can write

$$\begin{aligned}\hat{X} &= X - (X - \hat{X}) \\ &= A^{-1}(x - \hat{X}) - (X - \hat{X}) \\ &= (C - I)(X - \hat{X}).\end{aligned}$$

Due to the fact that the first row of  $C - I$ ,  $I$  being the identity matrix, is 0 it follows

$$\mathbb{P}_t x_{t+1} = \begin{cases} 0 & \text{for } t = 0 \\ \sum_{j=1}^t c_{t,j}(x_{t+1-j} - \mathbb{P}_{t-j}x_{t+1-j}) & \text{for } t = 1, 2, \dots, t-1. \end{cases} \quad (2.16)$$

Thus, we can calculate  $\mathbb{P}_0x_1, \mathbb{P}_1x_2, \dots$  recursively if the coefficients  $c_{i,j}$  are known. These coefficients can be again recursively calculated by

$$\begin{aligned}v_0 &= \gamma(0) \\ c_{t,t-k} &= v_k^{-1}(\gamma(t-k) - \sum_{j=0}^{k-1} c_{k,k-j}c_{t,t-j}v_j) \quad 0 \leq k < t \\ v_t &= \gamma(0) - \sum_{j=0}^{t-1} c_{t,t-j}^2 v_j\end{aligned}$$

From  $X = C(X - \hat{X})$  it follows that  $\Sigma = \text{cov}(X) = CDC^T$ , where  $D = \text{diag}(v_0, v_1, \dots, v_n) = E[(X - \hat{X})(X - \hat{X})^T]$  and thus

$$\begin{aligned} X^T \Sigma^{-1} X &= X^T (CDC^T)^{-1} X \\ &= X^T [X(X - \hat{X})^{-1} D (X(X - \hat{X})^{-1})^T]^{-1} X \\ &= (X - \hat{X})^T D^{-1} (X - \hat{X}) \\ &= \sum_{t=1}^n \frac{(x_t - \mathbb{P}_{t-1} x_t)^2}{v_{t-1}} \end{aligned}$$

and

$$\begin{aligned} \det(\Sigma) &= \det(C)^2 \det(D) \\ &= v_0 v_1 v_{n-1}. \end{aligned}$$

Inserting this in the Likelihood function 2.12 gives

$$L(\theta) = \frac{1}{\sqrt{(2\pi\sigma_w^2)^n (r_0 r_1 \cdots r_{n-1})}} \exp\left(-\frac{1}{2\sigma_w^2} \sum_{t=1}^n \frac{(x_t - \mathbb{P}_{t-1} x_t)^2}{r_{t-1}}\right)$$

with  $r_t := \frac{v_t}{\sigma_w^2}$ .  $\mathbb{P}_{t-1} x_t$  and  $r_t$  do not depend on  $\sigma_w^2$  so that the partial differentiation of the log-likelihood function with respect to the parameters leads to the maximum likelihood estimator. This estimator fulfils the following equations:

$$\begin{aligned} \hat{\sigma}_{w_{MLE}}^2 &= \frac{1}{n} S(\hat{\beta}_{MLE}) \\ &= \frac{1}{n} \sum_{t=1}^n \frac{(x_t - \mathbb{P}_{t-1} x_t)^2}{r_{t-1}} \end{aligned}$$

where  $\hat{\beta}_{MLE}$  denotes the value of  $\beta \in \Theta$  that maximizes the log-likelihood function  $l(\beta)$ . This optimization problem must be solved numerically.

The Bayesian information criterion (BIC) is one of the most widely known and used tools in statistical model selection. It was derived by Schwarz [Schwarz (1978)] to be used as an asymptotic approximation to a transformation of the Bayesian posterior probability of a candidate model [Neath & Cavanaugh (2012)]. It is a modified version of the Akaike information criterion (AIC). The computation of the BIC is based on the maximum likelihood function. Suppose we have a collection of candidate models  $M_1, \dots, M_k$ . These models can be based of different means, variance, covariance and even different distributions. We aim to find the best fitting ARMA(p,q) model for our data. The BIC for a candidate model is defined as

$$\begin{aligned} BIC &= -2\ln(L(\hat{\beta}_{MLE})) + k\ln(n) \\ &= -2\ln(L(\hat{\beta}_{MLE})) + (p + q + 1)\ln(n), \end{aligned} \tag{2.17}$$

where  $k$  is the number of unknown and to be estimated parameters and  $n$  the length of the time series.

In practice, every BIC for every considered model  $M_1, \dots, M_k$  is calculated and then the one corresponding to the minimal BIC is selected. The original derivation assumes that the observations are independent, identically distributed and come from an exponential family. The concept of exponential families is credited to E. J. G. Pitman, G. Darrois and B. O. Koopman in 1935–36. Exponential families are for example important in Bayesian statistics.

**Definition 2.2.3. (exponential family)** [Olive (2014)]

A family of joint probability density functions or joint probability mass functions  $\{f(x; \theta) | \theta \in \Theta\}$  for a random vector  $X$ ,  $x \in X$ , is an exponential family if

$$f(x; \theta) = h(x)c(\theta)\exp\left[\sum_{i=1}^l a_i(\theta)b_i(x)\right]$$

for all  $x$  where  $c(\theta) \geq 0$  and  $h(x) \geq 0$ . The functions  $c, h, b_i$  and  $a_i$  are real valued functions.

With the above mentioned assumptions the conditions for the application of the BIC are given and therefore we can use it in chapter 4.

As mentioned beforehand the BIC is very similar to the AIC. This information criterion is also based on the maximum likelihood estimator but the penalty term is different

$$\begin{aligned} AIC &= -2L(\hat{\beta}_{MLE}) + 2k \\ &= -2\ln(L(\hat{\beta}_{MLE})) + 2(p + q + 1). \end{aligned} \quad (2.18)$$

The third information criterion that is relevant in chapter 4 is the corrected Akaike information criterion, AICc,

$$AICc = -2\ln(L(\hat{\beta}_{MLE})) - \frac{2(p + q + 1)n}{n - p - q - 2}. \quad (2.19)$$

A comparison of the BIC, AIC and AICc is done in section 4.

### 3. Regime Transitions in Near-Surface Temperature Inversions: A Conceptual Model

In the paper by [van de Wiel et al. \(2017\)](#) a conceptual model is introduced to study regime transitions of near-surface temperature inversions in the nocturnal and polar atmospheric boundary layer. The authors were able to determine a connection between the dynamic stability of the inversion and the ambient wind speed  $U$  through their model and measurements. This is for example shown when plotting the temperature inversion between 10m and the surface, as a function of wind speed observed at Dome C, Antarctica, for the period 2014/2015 (see [figure 3.1](#)). The wind speed at height 10m is represented by  $U_{10m}$ ,  $T_{10m}$  is the temperature at height 10m and  $T_s$  the temperature at surface height. The plot shows a sign of two metastable states: one when the wind is weak and  $\Delta T = T_{10m} - T_s$  is big and one for strong wind where  $\Delta T$  is small. The model itself is an initial value problem with a first order ordinary differential equation, abbreviated ODE, which describes the difference between the surface temperature  $T_s$  and the one at a reference height  $T_r$  in one time step. In [chapter 4](#), we present a statistical indicator that was introduced in [Nevo et al. \(2017\)](#) to estimate the dynamical equilibrium properties of time series, based on a combination of dynamical systems concepts and stochastic processes tools introduced in [chapter 2](#). The present conceptual model is used to produce time series of controlled data for which the theoretical equilibrium properties are known. In this chapter we present a linear stability analysis of equilibrium points of the ODE for different values of a bifurcation parameter.

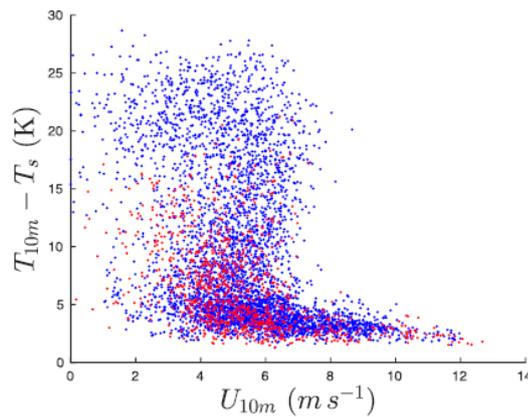


Figure 3.1.: Temperature inversion between 10m and the surface, as a function of wind speed as observed at Dome C, Antarctica, for the period 2014/2015. Each points corresponds to a 30 min averaged measurement. [[van de Wiel et al. \(2017\)](#)]

It is assumed that the wind speed and temperature is constant at a given height  $z_r$ . Then, the following equation describes the evolution of the near-surface inversion strength, based on a simple energy balance at the ground surface:

$$c_v \frac{d\Delta T}{dt} = Q_n - G - H$$

with

- $c_v$ : heat capacity of the soil,
- $\Delta T = T_r - T_s$ : inversion strength between the temperature at height  $z_r$  and at the surface  $z_s$ ,
- $Q_n$ : net long wave radiative flux (an energy loss at the surface that will be set as a constant),
- $G$ : soil heat flux (an energy storage term that will be parametrised as a linear term) and
- $H$ : turbulent sensible heat flux (a non-linear energy transport term that will be parametrised in the following).

After parametrizing the fluxes the model has the form:

$$c_v \frac{d\Delta T}{dt} = Q_i - \lambda \Delta T - \rho c_p c_D U \Delta T f(R_b)$$

with

- $Q_i$ : isothermal net radiation,
- $\lambda$ : lumped parameter, represents feedbacks from soil heat conduction,
- $\rho$ : density of air at constant pressure,
- $c_p$ : heat capacity of air at constant pressure,
- $c_D = \left(\frac{\kappa}{\ln(z_r/z_0)}\right)^2$ : neutral drag coefficient with  $\kappa \approx 0.4$  the von Kármán constant,  $z_0$  the roughness length and  $z_r$  the reference height,
- $U$ : wind speed at height  $z_r$ ,
- $R_b = z_r \left(\frac{g}{T_r}\right) \frac{\Delta T}{U^2}$ : bulk Richardson number and
- $f(R_b)$ : stability function.

The lumped parameter  $\lambda$  corresponds to the linear term in the model as the soil is assumed to respond linearly to the temperature inversion. Moreover,  $\Delta T \cdot f(R_b)$  is the non-linear term due to the non-linear dependence of turbulent diffusion on the vertical temperature gradient.

The model is simplified to an initial value problem

$$\begin{aligned} \frac{dx(t)}{dt} &= \begin{cases} Q_i - \lambda x(t) - Cx(t)(1 - x(t)) & \text{for } x(t) \leq 1 \\ Q_i - \lambda x(t) & \text{for } x(t) > 1 \end{cases} \\ x(t_0) &= x_0 \end{aligned} \quad (3.1)$$

where  $x$  represents  $\Delta T$ ,  $C := \rho c_p c_D U$  and  $f(x) = 1 - x$  the linear form of the stability function of the model, truncated for  $x > 1 \rightarrow f(x) = 0$ . Note that similar types of stability functions are typically used in numerical weather prediction tools. For a deeper discussion of the model, its simplifications and the model parameters the reader is referred to the paper by [van de Wiel et al. \(2017\)](#). In the following, we analyse the behaviour of equilibrium solutions depending on the values of the parameters  $C$ ,  $Q_i$  and  $\lambda$ .

### 3.1. Derivation of the Solution of the Simplified Model

The differential equation  $g(t, x) := \frac{dx}{dt}$ ,  $g : \mathbb{R} \rightarrow \mathbb{R}$ , in the simplified model [3.1](#) is a first order non-linear ordinary differential equation. As it does not depend on  $t$  it is an autonomous ODE, i.e. we can write  $g(t, x) = g(x)$ . Moreover,  $g(x)$  is continuous for all  $x \in \mathbb{R}$  and differentiable for all  $x \in \mathbb{R} \setminus \{1\}$  (see [Appendix B](#) for the proof).

**Theorem 3.1.1.** (*existence and uniqueness of solutions*) [[Hale \(1991\)](#)]

1. If  $g \in C^0(\mathbb{R}, \mathbb{R})$ , then, for any  $x_0 \in \mathbb{R}$ , there is an interval (possibly infinite)  $I_{x_0} = (\alpha_{x_0}, \beta_{x_0})$  containing  $t_0 = 0$  and a solution  $\varphi(t, x_0) := x(t)$  of the initial value problem

$$\begin{aligned} \frac{dx}{dt} &= g(x) \\ x(0) &= x_0 \end{aligned} \quad (3.2)$$

defined for all  $t \in I_{x_0}$ , satisfying the initial condition  $\varphi(0, x_0) := x_0$ .

2. If in addition,  $g \in C^1(\mathbb{R}, \mathbb{R})$ , then  $\varphi(t, x_0)$  is unique in  $I_{x_0}$  and  $\varphi(t, x_0)$  is a  $C^1$  function.

Thus there exists a unique solution of equation [3.1](#).

Note: We use  $t_0 = 0$  in all of the examples. Nonetheless, we derive the solution  $x(t)$  for an arbitrary  $t_0$ . We first derive it analytically and afterwards numerically.

### 3.1.1. Analytic Derivation

Recall that the model is:

$$\begin{aligned} \frac{dx(t)}{dt} &= \begin{cases} Q_i - \lambda x(t) - Cx(t)(1 - x(t)) & \text{for } x(t) \leq 1 \\ Q_i - \lambda x(t) & \text{for } x(t) > 1 \end{cases} \\ x(t_0) &= x_0 \end{aligned}$$

We will look at the two cases  $x \leq 1$  and  $x > 1$  separately and derive the solution for each of these cases.

**1. case:  $x \leq 1$**

$$\begin{aligned} \frac{dx}{dt} &= Q_i - \lambda x - Cx(1 - x) \\ &= Cx^2 - x(\lambda + C) + Q_i \\ \Leftrightarrow 1 &= \frac{1}{Cx^2 - x(\lambda + C) + Q_i} \frac{dx}{dt} \\ \Leftrightarrow \int 1 dt &= \int \frac{1}{Cx^2 - x(\lambda + C) + Q_i} dx \end{aligned} \quad (3.3)$$

After completing the square in the right hand side integral, the integral has the form:

$$= \int \frac{1}{C\left(x - \frac{\lambda+C}{2C}\right)^2 + Q_i - \frac{(\lambda+C)^2}{4C}} dx$$

We substitute  $u := x - \frac{\lambda+C}{2C}$ ,  $du = dx$

$$= \int \frac{du}{Cu^2 - \frac{(C+\lambda)^2}{4C} + Q_i}$$

We factor  $Q_i - \frac{(C+\lambda)^2}{4C}$  out

$$\begin{aligned} &= \int \frac{du}{\left(Q_i - \frac{(C+\lambda)^2}{4C}\right) \left(1 - \frac{4C^2u^2}{C^2+2C\lambda-4CQ_i+\lambda^2}\right)} \\ &= \frac{1}{Q_i - \frac{(C+\lambda)^2}{4C}} \int \frac{du}{1 - \frac{4C^2u^2}{C^2+2C\lambda-4CQ_i+\lambda^2}} \end{aligned}$$

We substitute again  $s = \frac{2iCu}{\sqrt{C^2+2C(\lambda-2Q_i)+\lambda^2}}$ ,  $ds = \frac{2iCdu}{\sqrt{C^2+2C(\lambda-2Q_i)+\lambda^2}}$

$$= -\frac{i\sqrt{C^2+2C(\lambda-2Q_i)+\lambda^2}}{2C\left(Q_i - \frac{(C+\lambda)^2}{4C}\right)} \int \frac{ds}{s^2+1}$$

The integral  $\int \frac{1}{s^2+1} = \tan^{-1}(s)$

$$= -\frac{i \tan^{-1}(s) \sqrt{C^2 + 2C(\lambda - 2Q_i) + \lambda^2}}{2C \left( Q_i - \frac{(C+\lambda)^2}{4C} \right)} + c_1$$

We substitute back for s

$$= -\frac{2 \tanh^{-1} \left( \frac{2Cu}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} \right)}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} + c_1$$

and for u

$$= -\frac{2 \tanh^{-1} \left( \frac{C(2x-1)-\lambda}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} \right)}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} + c_1$$

Hence,

$$\begin{aligned} \frac{dx}{dt} &= Q_i - \lambda x - C(1-x) \\ \Leftrightarrow \int 1 dt &= \frac{-2 \tan^{-1} \left( \frac{C(2x-1)-\lambda}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} \right)}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} + c_1 \\ \Leftrightarrow t + c_2 &= \frac{-2 \tan^{-1} \left( \frac{C(2x-1)-\lambda}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} \right)}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} + c_1 \\ \Leftrightarrow \tilde{c} &= \frac{-2 \tan^{-1} \left( \frac{C(2x-1)-\lambda}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} \right)}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} - t \end{aligned} \quad (3.4)$$

with  $\tilde{c} = c_2 - c_1$  where  $c_1$  and  $c_2$  are arbitrary numbers. By using the initial condition  $x(t_0) = x_0$  we get:

$$\tilde{c} = \frac{-2 \tan^{-1} \left( \frac{C(2x_0-1)-\lambda}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} \right)}{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2}} - t_0$$

Rearranging equation 3.4 gives:

$$x(t) = \frac{1}{2} \left[ \sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2} \tanh \left( -\frac{\sqrt{C^2 + 2(\lambda - 2Q_i) + \lambda^2} (t + \tilde{c})}{2} \right) + \lambda \right] + \frac{1}{2}$$

**2. case:  $x > 1$**

$$\begin{aligned}
 & \frac{dx}{dt} = Q_i - \lambda x \\
 \Leftrightarrow & \int 1 dt = \int \frac{1}{Q_i - \lambda x} dx \\
 \Leftrightarrow & t + a = -\frac{\log(Q_i - \lambda x)}{\lambda} \\
 \Leftrightarrow & -\lambda(t + a) = \log(Q_i - \lambda x) \\
 \Leftrightarrow & \exp(-\lambda(t + a)) = Q_i - \lambda x
 \end{aligned}$$

Thus,

$$x(t) = \frac{Q_i - \exp(-\lambda(t + a))}{\lambda}$$

with  $a$  some arbitrary number. Again by using the initial condition  $x(t_0) = x_0$  we get

$$a = -\frac{\log(Q_i - \lambda x_0)}{\lambda} - t_0$$

Figure 3.2 is a plot of  $x(t)$  with  $Q_i = 3, \lambda = 2$  and  $C = 8$  for three different initial conditions. We can see that all solutions, for the three initial conditions, converge to the

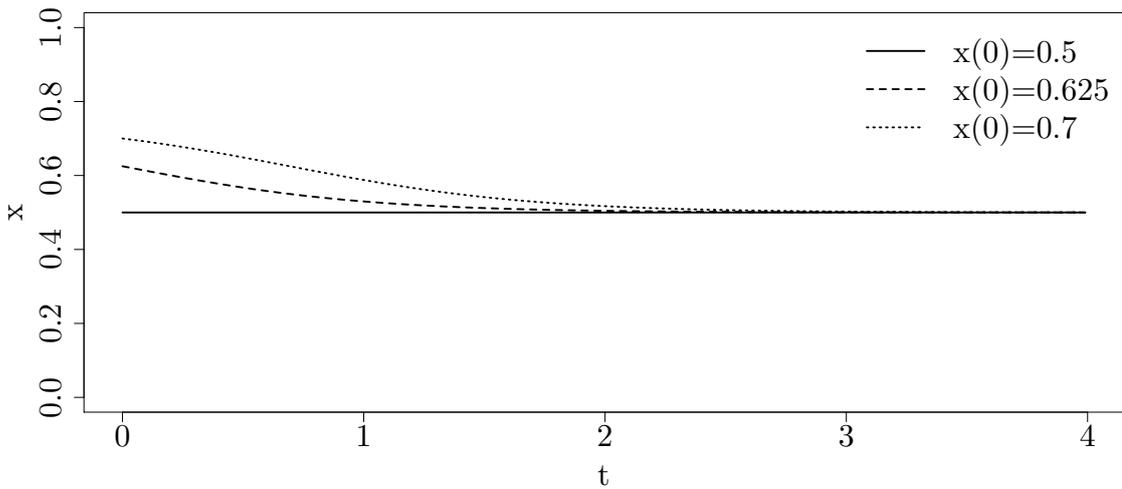


Figure 3.2.: Plot of  $x(t)$

constant solution  $x(t) = 0.5$ . This behaviour is explained in section 3.2.

### 3.1.2. Numerical Integration

To approximate the solution numerically we use a Runge-Kutta method [Corless (2013)]. The idea of Runge-Kutta methods is to evaluate the derivative function  $g(t, x)$  more than once, at different points, and then to use a weighted average of the obtained values as an approximation of the slope of the second. The general Runge-Kutta method is defined as

$$x_{k+1} = x_k + h \sum_{i=1}^s b_i k_i$$

$$k_i = g(t_k + c_i h, x_k + h \sum_{j=1}^{i-1} a_{i,j} k_j)$$

where  $h$  is step size,  $s$  number of stages,  $b_i$  weights and the matrix  $A = [a_{i,j}]$  is called the Runge-Kutta matrix. The information for explicit methods can be summarized in a tableau, called a Butcher tableau, of the form

$$\begin{array}{c|cccccc} & & 0 & 0 & 0 & \cdots & 0 \\ & c_2 & a_{2,1} & 0 & 0 & \cdots & 0 \\ c & c_3 & a_{3,1} & a_{3,2} & 0 & \cdots & 0 \\ & \vdots & \vdots & & & & \\ & c_s & a_{s,1} & a_{s,2} & \cdots & a_{s,s-1} & 0 \\ \hline & & b_1 & b_2 & \cdots & \cdots & b_s \end{array} = \frac{c}{b^T} A$$

The classical Runge-Kutta method, RK4 ( $s=4$ ), has the following Butcher tableau:

$$\begin{array}{c|cccc} 0 & & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ 1 & 0 & 0 & 1 & \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array}$$

It is a fourth order accurate method [Corless (2013)]. The major advantage of the Runge-Kutta 4 method is the ease of implementation, which will be especially important in section 4, while still being accurate. Figure 3.3 shows that the classical Runge-Kutta method, unlike the Euler method, gives a very good approximation for  $x(t)$  and the relative error,  $err := \frac{|x(t_k) - x_k|}{x_k}$ ,  $k = 1, \dots, n$ , is close to zero for  $h = 0.01$  which is due to the fact the RK4 method has better stability properties than the Euler method and is therefore a good choice to integrate the nonlinear ODE from the simplified model. Consequently, we use the RK4 method in chapter 5.

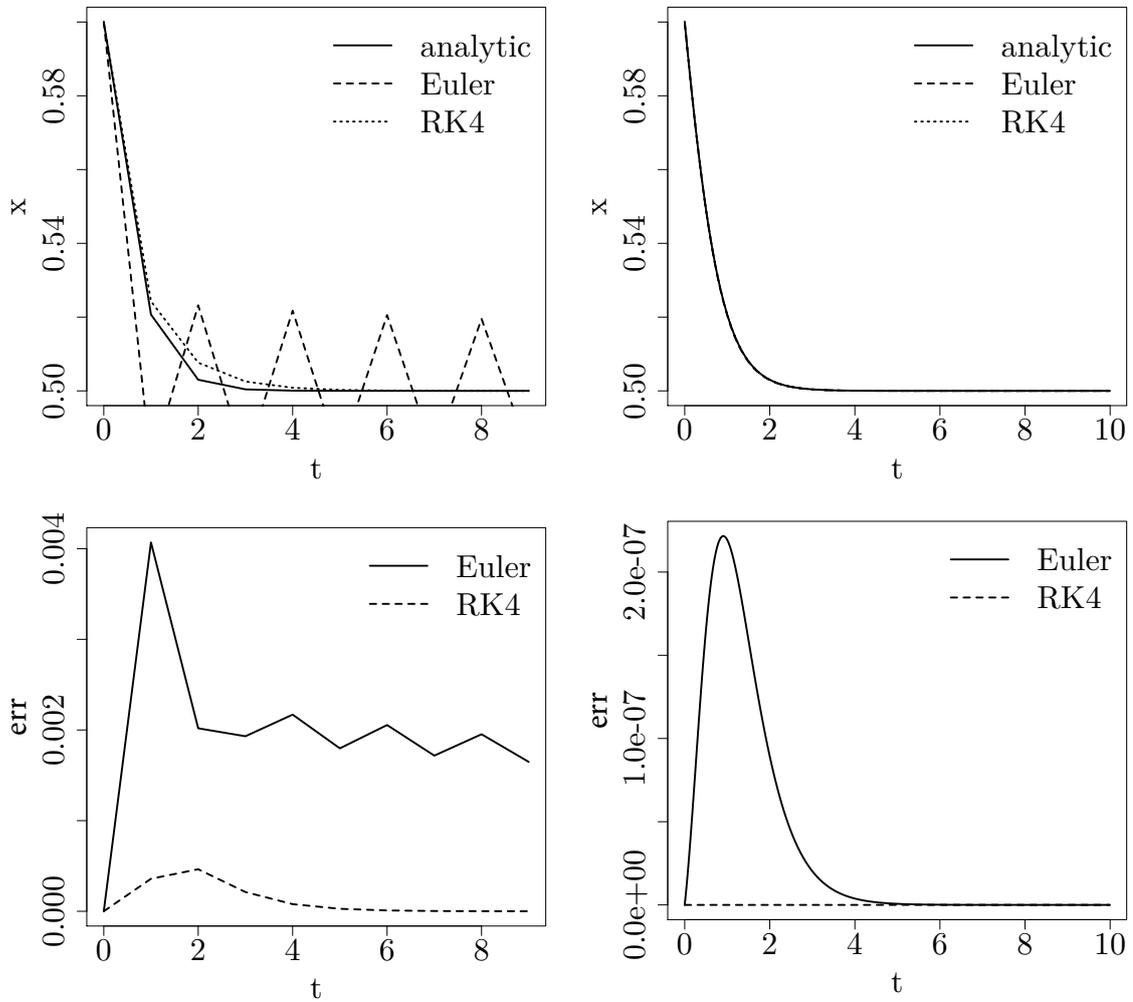


Figure 3.3.: comparison of the analytically and numerically derived solution (top) and their relative errors (bottom) ( $Q_i, \lambda, C$  from example 2 in the paper by [van de Wiel et al. \(2017\)](#)). From left to right:  $h = 1$  and  $h = 0.01$ .

### 3.2. Bifurcation Analysis

The next step is to find the equilibrium solutions of the simplified model and to characterize their stability properties. Roughly speaking an equilibrium point  $x_e$  is stable if all solutions starting near  $x_e$  stay nearby in the course of time. If, in addition, nearby solutions tend to  $x_e$  as  $t \rightarrow \infty$ , then  $x_e$  is asymptotically stable [Hale (1991)]. A precise definition is given later in this section.

**Definition 3.2.1. (equilibrium point)**[Hale (1991)]

An equilibrium point is a solution  $x(t) = x_e$  for which  $\frac{dx(t)}{dt} = 0$ .

Recall:

$$\frac{dx(t)}{dt} = \begin{cases} Q_i - \lambda x(t) - Cx(t)(1 - x(t)) & \text{for } x(t) \leq 1 \\ Q_i - \lambda x(t) & \text{for } x(t) > 1 \end{cases}$$

$$x(t_0) = x_0$$

We look at the cases  $x \leq 1$  and  $x > 1$  separately when calculating the equilibria.

**1. case:  $x \leq 1$**

$$\begin{aligned} 0 &= \frac{dx}{dt} \\ &= Cx^2 - x(\lambda + C) + Q_i \\ &= x^2 - x\left(\frac{\lambda}{C} + 1\right) + \frac{Q_i}{C} \end{aligned}$$

$$\Rightarrow x_{e1} = \frac{\frac{\lambda}{C} + 1}{2} - \sqrt{\left(\frac{1}{2}\left[\frac{\lambda}{C} + 1\right]\right)^2 - \frac{Q_i}{C}} \quad (3.5)$$

$$x_{e2} = \frac{\frac{\lambda}{C} + 1}{2} + \sqrt{\left(\frac{1}{2}\left[\frac{\lambda}{C} + 1\right]\right)^2 - \frac{Q_i}{C}} \quad (3.6)$$

**2. case:  $x > 1$**

$$\begin{aligned} 0 &= \frac{dx}{dt} \\ &= Q_i - \lambda x \end{aligned}$$

$$\Rightarrow x_{e3} = \frac{Q_i}{\lambda} \quad (3.7)$$

**Example 3.2.2.** Let  $\lambda = 2$ ,  $Q_i = 3$  and  $C = 8$ . With the given parameters the simplified model 3.1 has the form

$$\frac{dx}{dt} = \begin{cases} 3 - 2x - 8x(1 - x) & \text{for } x \leq 1 \\ 3 - 2x & \text{for } x > 1 \end{cases}$$

$$x(t_0) = x_0$$

The equilibrium solutions for the model with the given parameters are  $x_{e1} = \frac{1}{2}$ ,  $x_{e2} = \frac{3}{4}$  and  $x_{e3} = \frac{3}{2}$ . They can also be found by plotting the energy demand  $= Q_i = 3$  (dashed line) and the energy supply  $= \lambda x - Cx(1 - x) = 2x + 8x(1 - x)$  (solid line) as a function of  $x$  (see figure 3.4a) or by plotting  $\frac{dx}{dt}$  (see figure 3.4b). Equilibria are represented by the two intersection points of the energy demand and supply graphs or respectively the intersection points with the x-Axis.

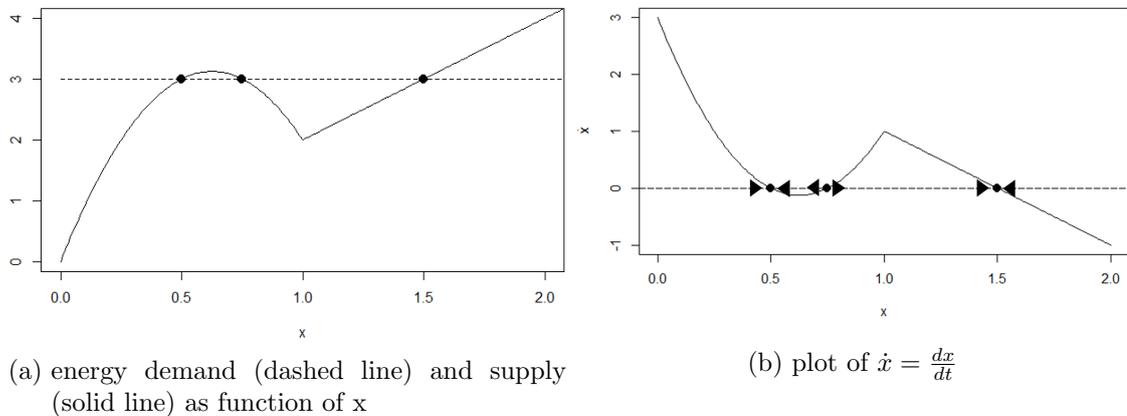


Figure 3.4.: plots of  $\frac{dx}{dt}$

The paper by [van de Wiel et al. \(2017\)](#) presents more examples of the derivation of the equilibrium points for the simplified model with different parameter values.

Now, we define the stability of the equilibrium solutions.

**Definition 3.2.3. (stable)** [[Hale \(1991\)](#)]

An equilibrium point  $x_e$  of equation 3.2 is said to be stable if, for any given  $\epsilon > 0$ , there is a  $\delta > 0$ , depending on  $\epsilon$ , such that, for every  $x_0$  for which  $|x_0 - x_e| < \delta$ , the solution  $\varphi(t, x_0)$  of equation 3.2 through  $x_0$  at 0 satisfies the inequality  $|\varphi(t, x_0) - x_e| < \epsilon$  for all  $t \geq 0$ . The equilibrium point  $x_e$  is said to be unstable if it is not stable.

**Definition 3.2.4. (asymptotically stable)** [Hale (1991)]

An equilibrium point  $x_e$  is said to be asymptotically stable if it is stable and, in addition, there is an  $r > 0$  such that  $|\varphi(t, x_0) - x_e| \rightarrow 0$  as  $t \rightarrow +\infty$  for all  $x_0$  satisfying  $|x_0 - x_e| < r$ .

These two definitions indicate that the stability of an equilibrium point  $x_e$  is a local property of the flow near the equilibrium. In fact, the Grobman-Hartman theorem [Hale (1991)] justifies the use of a linear approximation in a neighbourhood of the fixed point for hyperbolic fixed points (see definition 3.2.6). For a one-dimensional system, that leads to the following theorem.

**Theorem 3.2.5. (asymptotically stable)** [Hale (1991)]

Suppose  $g$  is a  $C^1$  function and  $x_e$  is an equilibrium point of  $\frac{dx}{dt} = g(x)$ , that is  $g(x_e) = 0$ . Suppose also that  $\frac{dg}{dx}(x_e) \neq 0$ . Then, the equilibrium point  $x_e$  is asymptotically stable if  $\frac{dg}{dx}(x_e) < 0$  and unstable if  $\frac{dg}{dx}(x_e) > 0$ .

Proof: Assuming  $x_e$  is an equilibrium point of

$$\dot{x}(t) = \frac{dx}{dt} = g(x)$$

the Taylor expansion of  $g(x)$  at  $x_e$  is

$$\begin{aligned} g(x) &= g(x_e) + g'(x_e)(x - x_e) + o(\|x - x_e\|) \\ &= g'(x_e)(x - x_e) + o(\|x - x_e\|). \end{aligned} \tag{3.8}$$

We use the Landau notation because

$$\lim_{x \rightarrow x_e} \frac{o(\|x - x_e\|)}{\|x - x_e\|} = 0.$$

We set  $\bar{x} = x - x_e$  and thus the linearisation of 3.8 is

$$\dot{\bar{x}} = g'(x_e)\bar{x}. \tag{3.9}$$

From this the statement follows directly. □

If not differently specified we will use the term stable instead of asymptotically stable in the rest of this master thesis.

Here,

$$\frac{dg}{dx} = \begin{cases} C(2x - 1) - \lambda & \text{for } x < 1 \\ -\lambda & \text{for } x > 1 \end{cases}$$

We exclude the case  $x = 1$  because  $g(x)$  is not differentiable in  $x = 1$ .

Obviously,  $x_{e_3}$  is stable for all  $\lambda > 0$  and unstable for  $\lambda < 0$ . The stability of  $x_{e_{1,2}}$  is determined by the following calculations.

$$\begin{aligned} \frac{dg}{dx}(x_{e_{1,2}}) &= 2C \left[ \frac{\lambda}{2C} + \frac{1}{2} \left( \pm \sqrt{\frac{1}{4} \left( \frac{\lambda}{C} + 1 \right)^2 - \frac{Q_i}{C}} \right) \right] - \lambda - C \\ &= \lambda + C \left( \pm \sqrt{C^2 \left( \frac{\lambda}{C} + 1 \right)^2 - 4CQ_i} \right) - \lambda - C \\ &= \pm \sqrt{C^2 \left( \frac{\lambda}{C} + 1 \right)^2 - 4CQ_i} \end{aligned} \quad (3.10)$$

Thus  $x_{e_1}$  is stable and  $x_{e_2}$  unstable for all  $C$  with the exception when

$$\begin{aligned} C^2 \left( \frac{\lambda}{C} + 1 \right)^2 - 4CQ_i &= 0 \\ \Leftrightarrow C^2 + C(2\lambda - 4Q_i) + \lambda^2 &= 0 \end{aligned}$$

Thus, for  $C_{1,2} = 2Q_i - \lambda \pm 2\sqrt{Q_i(Q_i - \lambda)}$   $x_{e_{1,2}} = 0$ . In example 3.2.2  $x_{e_2} = \frac{3}{4}$  is unstable and the first,  $x_{e_1} = \frac{1}{2}$ , and third equilibrium points,  $x_{e_3} = \frac{3}{2}$ , are stable.

**Definition 3.2.6. (hyperbolic equilibrium point)** [Hale (1991)]

An equilibrium point  $x_e$  of  $\frac{dx}{dt} = g(x)$  is called a hyperbolic equilibrium point if  $\frac{dg}{dx}(x_e) \neq 0$ . Otherwise it is called non-hyperbolic.

For  $\lambda \neq 0$   $x_{e_3}$  is hyperbolic and for  $C \neq C_{1,2}$   $x_{e_{1,2}}$  are also hyperbolic.

Equilibrium points can be created or destroyed, or their stability can change depending on the used parameters. These qualitative changes in the dynamics are called bifurcations, and the parameter values at which they occur are called bifurcation points. In the simplified model 3.1  $C := \rho c_p c_D U$ , where  $U$  is the wind speed. Considering the wind speed  $U$  or  $C$  as a bifurcation parameter is equivalent. In the following discussion we use  $C$  as the bifurcation parameter. Changing  $C$ , the bifurcation parameter of the simplified model, leads to a change of the stability of an equilibrium.

In the next part we want to find all bifurcation points that, depending on  $Q_i, \lambda$  and  $C$ , occur. Bifurcations can occur when one of the following three cases is given.

1.  $x_{e_1} = x_{e_2}$
2.  $x_{e_1} = x_{e_3}$
3.  $x_{e_2} = x_{e_3}$

Reminder:  $x_e$  are the equilibrium points of the simplified model.

Note that the conditions  $x_{e_{1,2}} < 1$  and  $x_{e_3} > 1$  must be met and that  $x_{e_{1,2,3}}$  have to be real. Again we exclude the case when  $x_{e_{1,2}} = 1$ . First we look at  $x_{e_1} = x_{e_2}$

$$\begin{aligned}
& \frac{\frac{\lambda}{C} + 1}{2} - \sqrt{\left(\frac{1}{2} \left[\frac{\lambda}{C} + 1\right]\right)^2 - \frac{Q_i}{C}} = \frac{\frac{\lambda}{C} + 1}{2} + \sqrt{\left(\frac{1}{2} \left[\frac{\lambda}{C} + 1\right]\right)^2 - \frac{Q_i}{C}} \\
\Leftrightarrow & \quad -\sqrt{\left(\frac{1}{2} \left[\frac{\lambda}{C} + 1\right]\right)^2 - \frac{Q_i}{C}} = \sqrt{\left(\frac{1}{2} \left[\frac{\lambda}{C} + 1\right]\right)^2 - \frac{Q_i}{C}} \\
\Leftrightarrow & \quad 0 = 2\sqrt{\left(\frac{1}{2} \left[\frac{\lambda}{C} + 1\right]\right)^2 - \frac{Q_i}{C}} \\
\Leftrightarrow & \quad 0 = \left(\frac{1}{2} \left[\frac{\lambda}{C} + 1\right]\right)^2 - \frac{Q_i}{C} \\
\Leftrightarrow & \quad 0 = \lambda^2 + 2\lambda C + C^2 - 4Q_i C \\
\Leftrightarrow & \quad 0 = C^2 + C(2\lambda - 4Q_i) + \lambda^2 \\
\Rightarrow & \quad C_{1,2} = -\frac{2\lambda - 4Q_i}{2} \pm \sqrt{\left(\frac{2\lambda - 4Q_i}{2}\right)^2 - \lambda^2} \\
& \quad = 2Q_i - \lambda \pm 2\sqrt{Q_i(Q_i - \lambda)}
\end{aligned}$$

Hence, for

$$C_1 = 2Q_i - \lambda - 2\sqrt{Q_i(Q_i - \lambda)}$$

and

$$C_2 = 2Q_i - \lambda + 2\sqrt{Q_i(Q_i - \lambda)} \tag{3.11}$$

it holds that  $x_{e_1}$  and  $x_{e_2}$  are equal.

Moreover,  $x_{e_1} = x_{e_3}$  and  $x_{e_2} = x_{e_3}$  can never hold because  $x_{e_{1,2}} < 1$  and  $x_{e_3} > 1$ . Furthermore, we have shown earlier that  $x_{e_{1,2}}$  do not change their stability for  $C \neq C_{1,2}$ . In conclusion, it is sufficient to restrict our analysis to the case when  $x_{e_1} = x_{e_2}$ . Additionally, we make the following assumptions

1.  $C_{1,2} \in \mathbb{R}$
2.  $C_{1,2} > 0$
3.  $0 \leq x_{e_{1,2}}(C_1) < 1$
4.  $0 \leq x_{e_{1,2}}(C_2) < 1$

Note: For physical reasons  $\lambda \geq 0$ . Nonetheless, for the sake of completeness we allow negative  $\lambda$  in the following calculations. The reasoning for the assumptions 1-4 is explained in the following paragraphs. Given the fact that the calculations consist mostly of basic, but lengthy, transformations they are omitted to shorten this section.

To ensure that  $C_1$  and  $C_2$  are real it has to hold that

$$Q_i(Q_i - \lambda) \geq 0.$$

If we choose

- $\lambda = 0$  and  $Q_i \in \mathbb{R}$  or
- $\lambda < 0$  and  $Q_i \geq 0$  or
- $\lambda > 0$  and  $Q_i \geq \lambda$
- $\lambda < 0$  and  $Q_i \leq \lambda$  or
- $\lambda > 0$  and  $Q_i \leq 0$  or

assumption 1 is given.

We set  $C_{1,2} \geq 0$  because  $C := \rho c_p c_D U$  where  $\rho, c_p$  and  $c_D \in \mathbb{R}_{\geq 0}$  and we assume that the wind comes from the same direction which means that  $U \in \mathbb{R}_{\geq 0}$ . But for  $C_{1,2} = 0$   $x_{e_{1,2}}(C_{1,2})$  is equal to infinity. To account for this we only allow  $C_{1,2} > 0$ . By calculating  $C_1$  and  $C_2$  for each case, in which assumption 1 is fulfilled, and checking if  $C_{1,2} > 0$  we get

- $\lambda < 0$  and  $Q_i = 0$  or
- $\lambda > 0$  and  $Q_i = \lambda$  or
- $\lambda < 0$  and  $Q_i > 0$  or
- $\lambda > 0$  and  $Q_i > \lambda$

and for  $C_2$  only

- $\lambda = 0$  and  $Q_i > 0$

In addition, we want  $x_{e_{1,2}}(C_1) \vee x_{e_{1,2}}(C_2) \in [0, 1)$ . Even though, mathematically there is no reason not to allow negative equilibrium points it disagrees with the context of the model by Van de Wiel.  $x$  represents  $\Delta T$  which is the difference of the temperature at reference height and the one at the surface. If this is negative the temperature at the surface is higher than the one at reference height. But this model was specifically defined for a temperature inversion, i.e. the exact opposite. Furthermore, we do not allow  $x_{e_{1,2}}(C_{1,2}) = 1$  because in this case  $g(x)$  is not differentiable. We only consider the four plus one cases in which  $C_{1,2} \in \mathbb{R}_{>0}$ . This gives us that a bifurcation point exists only when  $C = C_1$ ,  $\lambda < 0$  and  $Q_i = 0$  or when  $C = C_2$  and

- (i)  $\lambda < 0$  and  $Q_i = 0$  or (ii)  $\lambda < 0$  and  $Q_i > 0$  or  
 (iii)  $\lambda > 0$  and  $Q_i > \lambda$  or (iv)  $\lambda = 0$  and  $Q_i > 0$ .

When  $\lambda < 0$  and  $Q_i = 0$  it follows that  $C_1 = C_2$ .

In conclusion, we have shown that there can only exist maximal one bifurcation point for fixed  $\lambda$  and  $Q_i$ . Moreover,  $\lambda$  and  $Q_i$  have to be chosen according to the before made assumptions for the bifurcation point to exist.

After finding the bifurcation points we want to classify the bifurcation that occurs.

**Definition 3.2.7. (bifurcation diagram)**[Seydel (1994)]

A diagram depicting  $x(t)$  versus  $C$ , where  $(x, C)$  solves equation  $g(x) = 0$  (Note:  $g(x)$  depends on  $C$ .) is called a bifurcation diagram.

Equilibria may form continuous curves in a bifurcation diagram. These curves are called branches. A branch or a part of a branch is called stable (unstable) if all its solutions are stable (unstable). By convention, stable branches are denoted by a solid line and unstable ones by a dotted one.

By looking at the bifurcation diagram (figure 3.5) with  $\lambda$  and  $Q_i$  according to ii)-iv) we can see that for  $C > C_2$  there are two equilibria and at  $C = C_2$  the stable and unstable equilibrium collide and annihilate each other, i.e. there are no equilibria for  $C < C_2$ . This type of bifurcation is called a saddle-node bifurcation (see the book by Seydel (1994)).

To make the behaviour of the equilibria in case i) more obvious we plot the bifurcation diagram for  $x_e \in \mathbb{R}$  and  $C \in \mathbb{R} \setminus 0$  (figure 3.5), meaning we allow negative equilibria and  $C$ .

Doing this we can see that we have one unstable branch, which is equal to zero, for  $C < C_1$  and one stable branch. At  $C = C_1$  they collide and the former unstable branch turns into a stable one and conversely. This is called a transcritical bifurcation (see Seydel (1994)).

Note: The only physical relevant cases are the two plots at the bottom of figure 3.5.

We proceed by looking at  $C \rightarrow \infty$  and  $Q_i$  and  $\lambda \in \mathbb{R}$ . Again, the equilibrium solutions are:

$$x_e = \begin{cases} \frac{\lambda+1}{2} \pm \sqrt{\left(\frac{1}{2}[\lambda+1]\right)^2 - \frac{Q_i}{C}} & \text{for } x \leq 1 \\ \frac{Q_i}{\lambda} & \text{for } x > 1 \end{cases} \quad (3.12)$$

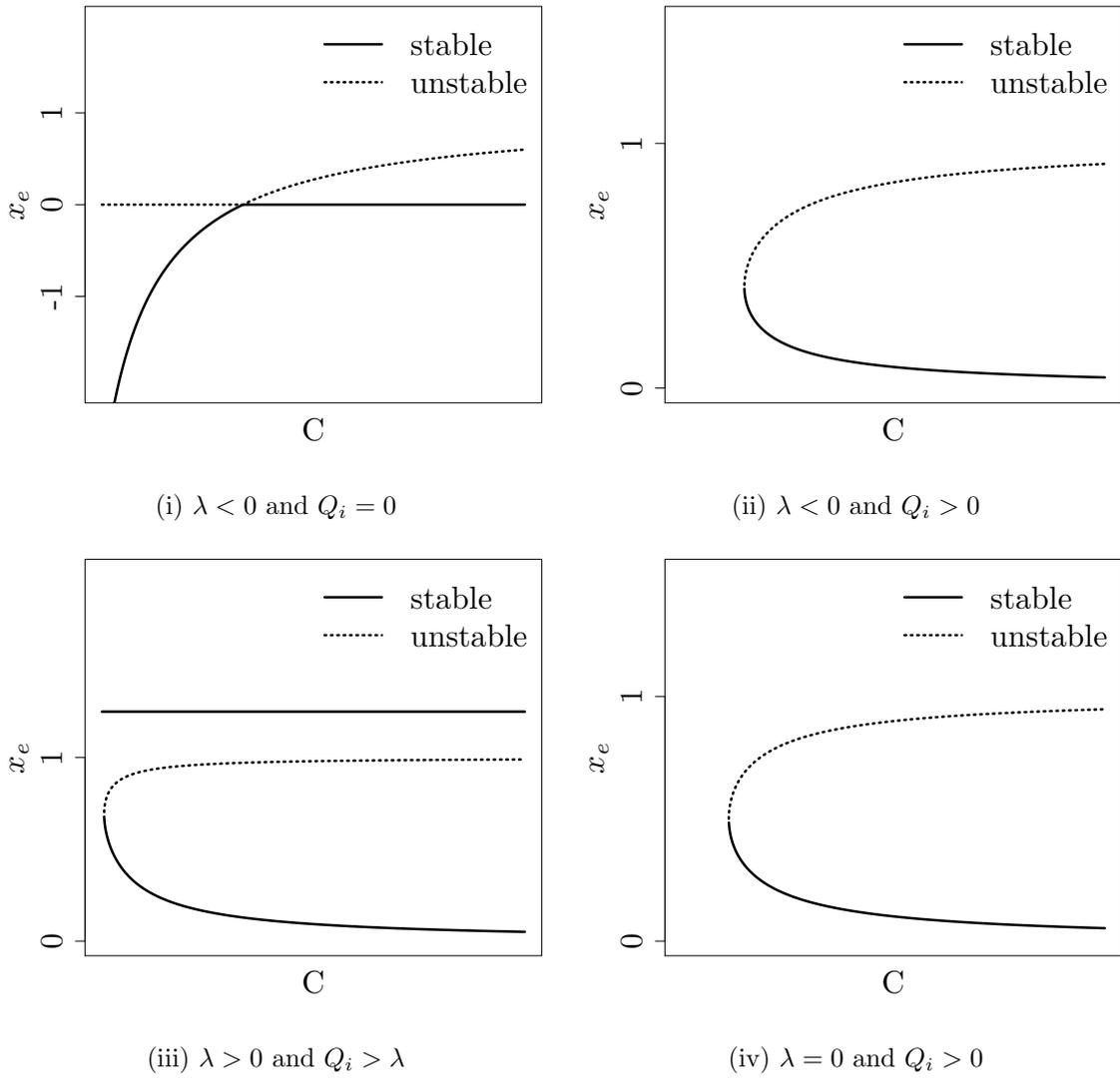


Figure 3.5.: bifurcation plot of 3.12 with  $\lambda$  and  $Q_i$  according to i) - iv)

Thus, as long as  $\lambda \neq 0$  and  $\frac{Q_i}{\lambda} > 1$   $x_{e3}$  exists and is constant. For  $x < 1, C \rightarrow \infty$  and  $Q_i, \lambda \in \mathbb{R}$ , but fixed, we get

$$\begin{aligned} \lim_{C \rightarrow \infty} x_e &= \frac{1}{2} \pm \sqrt{\left(\frac{1}{2}\right)^2} \\ &= \frac{1}{2} \pm \frac{1}{2} \end{aligned}$$

Consequently, for  $C \rightarrow \infty, \lambda \neq 0$  and  $\frac{Q_i}{\lambda} > 1$  we have three equilibrium solutions  $x_{e1} = 0, x_{e2} = 1$  and  $x_{e3} = \frac{Q_i}{\lambda}$ .

To conclude this chapter we examine the behaviour after a perturbation. For this we first calculate the solution of equation 3.9 ( $\dot{\bar{x}} = g'(x_e)\bar{x}$  with  $\bar{x} = x - x_e$ ) which is the linearisation of the Taylor series expansion of  $\frac{dx}{dt} = g(x)$  at an equilibrium  $x_e$ . The solution is given by

$$\bar{x} = \bar{x}_0 \exp(g'(x_e)t). \quad (3.13)$$

With the initial value  $\bar{x}_0 = \bar{x}(t_0)$ . Rearranging equation 3.13 gives

$$\begin{aligned} \exp(g'(x_e)t) &= \frac{\bar{x}}{\bar{x}_0} \\ \Leftrightarrow g'(x_e)t &= \ln\left(\frac{\bar{x}}{\bar{x}_0}\right) \\ \Leftrightarrow t &= \frac{1}{g'(x_e)} \ln\left(\frac{\bar{x}}{\bar{x}_0}\right). \end{aligned}$$

This  $t$  is called the recovery time scale. It gives us an estimate of the amount of time it takes the systems to return to equilibrium after a perturbation. For  $x$  close to  $x_e$ , i.e.  $\|x - x_e\| < \epsilon$  for  $\epsilon > 0$ , it follows that  $\bar{x}$  is nearly 0 and therefore  $t \rightarrow \pm\infty$  depending on the sign of  $g'(x_e)$  and for  $x$  further away from  $x_e$ ,  $\bar{x}$  gets big and therefore  $\ln\left(\frac{\bar{x}}{\bar{x}_0}\right) \rightarrow 1$  which gives  $t \rightarrow \frac{1}{g'(x_e)}$ .

Note that  $g'(x_e)$  depends on  $C$  (see equation 3.10) for the first two equilibria in the simplified model. Therefore, when  $C$  is approaching the bifurcation point it follows that  $g'(x_e(C))$  approaches 0 and therefore  $\frac{1}{g'(x_e)} \rightarrow \pm\infty$ . Which means that it takes an infinite amount of time to damp an initial perturbation. This is called critical slowing down.

The concepts of return time closely connects with ARMA(p,q) modelling. Intuitively  $p$  and  $q$  are related to the memory lag of the process. The longer the system takes to return to the equilibrium after a perturbation, the more memory we expect to observe in the process. Far from the transition (i.e shift from one dynamical regime to another), the time series of a generic physical observable can be described by an ARMA(p,q) model with a reasonably low number of  $p, q$  parameters and coefficients. On the other hand, close to a transition the statistical properties (such as the shape and/or the persistence of the autocorrelation function) of the system change, leading to an increase of the value  $p + q$  [Faranda et al. (2014)]. Thus, we use ARMA(p,q) models in the following chapter to analyse the stability of a dynamical system.



## 4. Dynamical Stability of Near-Surface Temperature Inversions

In the last chapter we analysed the simplified model by [van de Wiel et al. \(2017\)](#) which was developed to understand regime transitions in near-surface temperature inversions. Moreover, a hypothesis to explain the existence of two metastable equilibria of the temperature inversion was given. In this chapter the goal is to statistically detect critical transitions between the two possible equilibria predicted by the model. For the detection we apply an indicator for the dynamical stability (i.e. the resilience to perturbations) which was defined by [Nevo et al. \(2017\)](#). It uses different methods from dynamical systems and statistical modelling. Especially, ARMA(p,q) processes play a key role (see chapter 2). Additionally, in chapter 5 we apply the indicator to a controlled data set and observation data to test its reliability.

### 4.1. The Stability Indicator

The idea of the stability indicator is to use ARMA(p,q) models to detect local stability in a time series. For this we slice the time series with a moving time window of length  $\tau$ . In other words, we obtain subsequences  $\{x_1, \dots, x_\tau\}, \{x_2, \dots, x_{\tau+1}\}, \dots, \{x_{t-\tau+1}, \dots, x_t\}$  of the original time series that overlap. We assume that for this time scale  $\tau$  of interest, the time series represents a stationary phenomenon and we then fit every ARMA(p,q) model with  $p \leq p_{max}$  and  $q \leq q_{max}$  to these subsequences, where  $p_{max}$  and  $q_{max}$  are predefined thresholds. Afterwards we choose the best fitting ARMA(p,q) model by choosing the one with the minimal BIC. By slicing the original time series we obtain shorter time series which satisfy the requirements for ARMA modelling. That is, we assume that the subsequences are realizations of linear processes with Gaussian white noise which then implies that the process is stationary (see Appendix C).

For the definition of the stability indicator it is necessary that the dynamics near a metastable state can be modelled by an AR(1)-process. To prove this statement we proceed by

1. defining a metastable system and a potential,
2. showing that the potential near a stable state can be approximated by a quadratic function,
3. proving that the system near a stable state can be modelled by a Langevin equation with a quadratic potential and

4. showing that the Langevin equation from step 3 is an AR(1)-process.

Proof:

**Step 1:** A metastable system is one with at least two stable states. We call the stable states in a metastable system metastable. Generally, for  $C > C_2$  and  $\lambda > 0, Q_i > \lambda$  the simplified model by [van de Wiel et al. \(2017\)](#) is a metastable system.

The dynamics of a first order system  $\frac{dx}{dt} = g(x)$  can be visualized based on the physical idea of potential energy [[Strogatz \(1994\)](#)]. For this we picture a particle sliding down the walls of a potential well, where the potential  $V(x)$  is defined as

$$g(x) = -\frac{dV}{dx}.$$

Equilibrium points  $x_e$  correspond to extrema of the potential, i.e.

$$\frac{dx}{dt}(x_e) = -\frac{dV}{dx} = 0.$$

The potential of the simplified model [3.1](#) has the form

$$V(x) = \begin{cases} \frac{1}{2}x^2(\lambda + C) - \frac{C}{3}x^3 - Q_i x & \text{for } x \leq 1 \\ \frac{1}{2}\lambda x^2 - Q_i x + \frac{1}{6}C & \text{for } x > 1 \end{cases} \quad (4.1)$$

Figure [4.1](#) shows the plot of a potential with the parameters we use in chapter [5](#), i.e.  $\lambda = 2, Q_i = 2.5, C = 6.4$ . The red line is the linearisation of the potential around  $x_{e1}$  (see step 2) with  $V_0$  being a minimum of the potential.

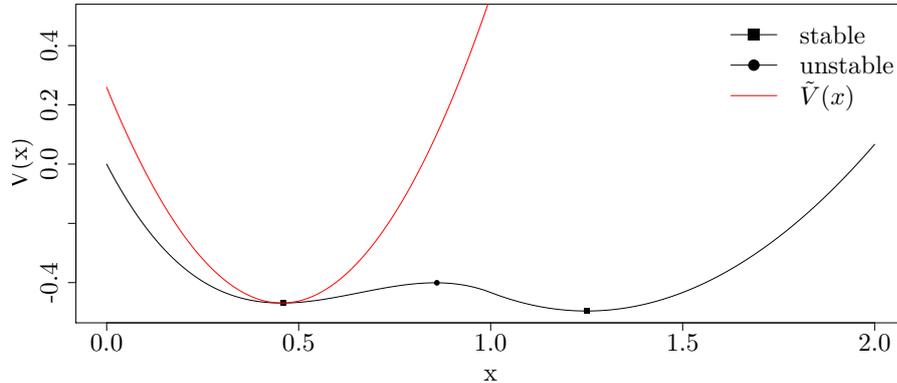


Figure 4.1.: example of a potential  $V(x)$  ([4.1](#)) and the linearisation  $\tilde{V}(x)$  ([4.2](#))

We can clearly see that minima of  $V$  correspond to the stable equilibria  $x_{e1} = 0.46, x_{e3} = 1.25$  the maximum to the unstable one  $x_{e2} = 0.86$ .

**Step 2:** To show that the potential  $V(x)$  is quadratic near a stable equilibrium we use the Taylor series expansion of  $V$  at  $x = \tilde{x}$

$$V(x) = V(\tilde{x}) + \frac{dV}{dx}|_{\tilde{x}}(x - \tilde{x}) + \frac{1}{2} \frac{d^2V}{dx^2}|_{\tilde{x}}(x - \tilde{x})^2 + \dots$$

where  $\frac{dV}{dx}|_{\tilde{x}}$  is the first derivative of  $V$  evaluated at  $x = \tilde{x}$ .

Let  $\tilde{x}$  be an equilibrium point. Then  $\frac{dV}{dx}|_{\tilde{x}} = 0$ . We can set  $V(\tilde{x}) = V_0$ ,  $V_0 \in \mathbb{R}$  as  $V(\tilde{x})$  is constant. For  $x$  close to a stable equilibrium  $x - \tilde{x}$  is small and thus  $\frac{d^2V}{dx^2}$  is the dominant term in the expansion and we do not need to include derivatives of higher order. We define  $k := \frac{d^2V}{dx^2}|_{\tilde{x}}$ ,  $k \in \mathbb{R}$ . Therefore, we obtain the following approximate potential

$$\tilde{V}(x) = \frac{1}{2}k(x - \tilde{x})^2 + V_0. \quad (4.2)$$

Consequently we can approximate the curve of the potential, close to a stable equilibrium, with a parabola.

**Step 3:** The Langevin equation is a stochastic differential equation describing the velocity of a microscopic particle undergoing Brownian motion.

$$\begin{aligned} \frac{dx}{dt} &= u \\ \frac{du}{dt} &= -\frac{\gamma}{m}u - \frac{1}{m} \frac{dV}{du} + \frac{1}{m} \frac{dB}{dt} \\ x(t_0) &= x_0 \\ \frac{dx(t_0)}{dt} &= x_1 \end{aligned} \quad (4.3)$$

where

- $x_0$  and  $x_1$  are initial distributions independent of Brownian motion,
- $m$  is the mass of the particle,
- $\gamma$  is the frictional force,
- $V$  is the potential,
- $x$  is the position and  $u$  the velocity of the particle and
- $B$  is Brownian motion.

**Definition 4.1.1. (Brownian motion)**

A Brownian motion, also called a Wiener process,  $B = (B_t)_{t \geq 0}$  is a stochastic process taking values in  $\mathbb{R}$  such that

- (i)  $B_0(\omega) = 0$  for almost all  $\omega \in \Omega$ ,
- (ii)  $B_{t_n} - B_{t_{n-1}}, \dots, B_{t_1} - B_{t_0}$  are independent for all  $n \geq 0, 0 = t_0 \leq t_1 < t_2 < \dots < t_n < \infty$ ,
- (iii)  $B_t - B_s \sim B_{t+h} - B_{s+h}$  for all  $0 \leq s < t, h \geq -s$ ,
- (iv)  $B_t - B_s \sim \mathcal{N}(0, t - s)$  and
- (v)  $t \rightarrow B_t(\omega)$  is continuous for all  $\omega \in \Omega$ .

Near a stable state we can use the approximate potential  $\tilde{V}(x)$  from step 2. This gives

$$\begin{aligned} \frac{dx}{dt} &= u(t) \\ \frac{du}{dt} &= -\frac{\gamma}{m}u - \frac{1}{m}ku + \frac{1}{m} \frac{dB}{dt} \end{aligned} \quad (4.4)$$

**Step 4:** It is important to note that for almost all  $\omega \in \Omega$  the sample path  $t \rightarrow B_t(\omega)$  is in fact differentiable for no  $t \geq 0$ . Thus  $\frac{dB}{dt}$  does not really exist. But we can discretize it.

We discretize 4.4 with the Euler-Maryuma scheme. As a grid we use

$$\{0 = t_0 < t_1 < \dots < t_N = T\}$$

with  $\Delta t = t_n - t_{n-1}$ . Then,

$$\begin{aligned} u_t - u_{t-1} &= -\frac{\gamma}{m} \Delta t u_{t-1} - \frac{1}{m} \Delta t k u_{t-1} + \frac{1}{m} (B_t - B_{t-1}) \\ \Leftrightarrow u_t &= u_{t-1} \left(1 - \frac{\gamma}{m} \Delta t - \frac{1}{m} \Delta t k\right) + \frac{1}{m} (B_t - B_{t-1}) \end{aligned} \quad (4.5)$$

By defining  $\phi := 1 - \frac{\gamma}{m} \Delta t - \frac{1}{m} \Delta t k$  and  $w_t := \frac{1}{m} (B_t - B_{t-1})$  we get

$$u_t = \phi u_{t-1} - w_t.$$

The last step is to verify that  $u_t$  is stationary and  $w_t$  is white noise.

- expected value:  $E[w_t] = E \left[ \frac{1}{m} (B_t - B_{t-1}) \right]$   
 $= \frac{1}{m} E[B_t - B_{t-1}]$   
 $\stackrel{(iv)}{=} 0$

- variance:  $var[w_t] = \left(\frac{1}{m}\right)^2 var(B_t - B_{t-1})$ 

$$\stackrel{(iv)}{=} \left(\frac{1}{m}\right)^2 (t - t + 1)$$

$$= \left(\frac{1}{m}\right)^2$$

$$= \sigma_w^2$$

$$< \infty$$

Moreover,  $w_t$  is independent of  $w_{\hat{t}}$  for all  $t \neq \hat{t}$  and normally distributed due to the definition of Brownian motion. Therefore,  $w_t$  is Gaussian white noise.

When we assume,  $|\phi| < 1$  and  $sup_t var(u_t) < \infty$  we get, according to example 2.1.6, that  $\{u_t\}$  is stationary. Thereby,  $\{u_t\}$  is an AR(1) process. □

We have thus seen that, in the vicinity of a stable equilibrium, the dynamics can be approximately described by an AR(1) process. In the following, we will study deviations from AR(1) processes in the space of ARMA(p,q) processes to quantify the dynamical stability of our time series. For that, we assume that ARMA(p,q) processes, except AR(1), can be used to model the dynamics when the system is not close to a stable state. The idea behind this assumption is that ARMA processes are an important parametric family of stationary time series [Brockwell (2016)]. Their importance is due to their flexibility and their capacity to describe almost all features of stationary time series. Thereby, as we have to make some Ansatz, choosing ARMA(p,q) processes, except AR(1), for modelling the dynamics near an unstable state is a good start.

As a summary when the system is close to a metastable state it can be modelled by an AR(1) process and if it is close to an unstable state higher order ARMA processes are needed. Therefore, the stability indicator is defined as

$$\Upsilon(p, q; \tau) = 1 - \exp\left(\frac{-|BIC(p, q) - BIC(1, 0)|}{\tau}\right).$$

For a stable state  $\Upsilon = 0$  and the limit  $\Upsilon \rightarrow 1$  corresponds to a very unstable point. To simplify the notation we drop the dependence of  $\Upsilon$  on  $p, q$  and  $\tau$ .

## 4.2. Dependence of $\Upsilon$ on the window length $\tau$

As a reminder:  $\tau$  is the length of the window which slices the time series in subsequences to assure stationarity. The reliability of  $\Upsilon$  highly depends on the choice of  $\tau$ . This can easily be seen in figure 4.2. This is a plot of a process with two stable equilibria (solid red lines) and an unstable one (dotted red line). The colours ranging from dark blue to yellow represent the stability of the points measured by  $\Upsilon$ . It is important to note that we colour the last point of the subsequence, i.e.  $\Upsilon_{k-\tau+1}$  corresponds to the subsequence  $x_{k-\tau}, \dots, x_k$  and  $x_k$  is coloured according to the value of  $\Upsilon_{k-\tau+1}$ . We choose this type of colour coding because an ARMA(p,q) process is defined by its past values, i.e. for an AR(1) process  $x_t$  depends on  $x_{t-1}$ . In the tests we use the `auto.arima()` function from the "forecast" R package [Hyndman et al. (2019)]. Grey dots have one of the following two meanings:

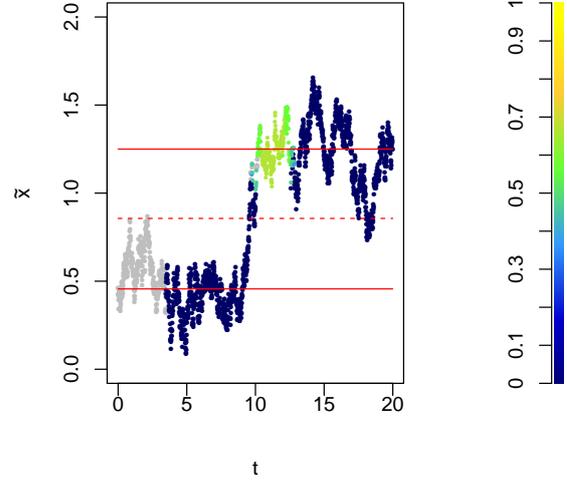


Figure 4.2.

- The `auto.arima()` function was not able to find a stationary ARMA(p,q) with  $p \leq p_{max}$  and  $q \leq q_{max}$  that fits the corresponding subsequence.
- $t \leq \tau - 1$ . This is also due to the fact that we always colour the last element of the subsequence.

In figure 4.2 we clearly see that  $\Upsilon$  correctly recognizes the transition but stable points get marked as unstable. This is due to the fact that in this case  $\tau$  is chosen too big (i.e.  $\tau = 350$ ). The points which were marked as unstable actually are endpoints of subsequences with mainly elements around the unstable equilibrium. At this point we want to remark that one has to be aware of this characteristic of  $\Upsilon$ . Even if we were able to find the optimal  $\tau$  the subsequence has to consist of several points around the unstable equilibrium (We can not apply  $\Upsilon$  to only one point.) and consequently  $\Upsilon$  always marks points after the unstable equilibrium as unstable. Nonetheless, it is important to find a good approach for choosing  $\tau$ . We suggest two approaches. The first one is based on the idea that  $\tau$  should be smaller than the amount of points covering the transition from one stable state to another and the second one takes into account that we need subsequences which are sampled from a normal distribution to be able to find the model with the minimal Bayesian Information Criterion. Both approaches are applicable when the theoretical model which was used to simulate the data is unknown. This is important in cases where data showing signs of metastability are available, but an underlying model is unknown.

### 4.2.1. Approach 1: K-Means

We suggest using the K-means algorithm to select a window length for the analysis. The idea is to cluster the data in different categories for data around each stable fixed point, and data near the unstable fixed point, or in other words, data covering transition periods between two metastable states. By that, the goal is to estimate the time needed by the system to transition between two metastable states. Generally, the aim of the K-means algorithm by [Hartigan & Wong \(1979\)](#) is to divide  $M$  points into  $K$  clusters so that the sum of squares in the clusters is minimized. The clustering is done in the following steps:

- Input:  $K = \#$  number of clusters, set of points  $x_{i-\tau+1}, \dots, x_i$
- Place centroids  $c_1, \dots, c_k$  at random locations.
- Repeat until none of the cluster assignments change:
  - for each point  $x_i$  find nearest centroid  $c_j$  and assign  $x_i$  to cluster  $j$
  - for each cluster  $j = 1, \dots, k$  calculate new centroid  $c_j = \text{mean of all points } x_i \text{ assigned to cluster } j$  in previous step.

We choose  $\tau$  such that it is smaller than the minimal mean time spend in one cluster.

When the data points, more precisely the equilibria, are equally distributed over the  $k$  clusters this algorithm gives a good estimation for the amount of points covering the transition. Otherwise, the approximation of the amount of points covering the transition might be imprecise. Another limitation of the algorithm is that the number of clusters has to be chosen before applying the algorithm.

### 4.2.2. Approach 2: Anderson-Darling Normality Test

Approach 1 does not take into account the statistical properties of the process. But as mentioned in chapter 2 we need to assume a distribution for the underlying process to be able to calculate the Bayesian information criterion. We assume that the subsequences are sampled from a normal distribution. Therefore, we need to choose a  $\tau$  such that this assumption holds. Hence, in this approach we use a normality test to find the biggest window length for which we can conclude that most of the subsequences are in fact sampled from a normal distribution. The normality test we apply is the Anderson-Darling Test [[Anderson & Darling \(1952\)](#)], abbreviated AD test, as it is, for example, more stable than the Kolmogorov-Smirnov test. The AD test statistic is based on the squared difference between the empirical distribution function estimated based on the sample,  $F_n(x)$ , and the normal distribution  $F^*(x)$ . The statistic for this test is,

$$W_n^2 = n \int_{-\infty}^{\infty} [F_n(x) - F^*(x)]^2 \psi(F^*(x)) dF^*(x)$$

where  $\psi$  is a non-negative weight function. We use the modified AD statistic given by D'Agostino and Stephens (1986) which takes into accounts the sample size  $n$

$$W_n^{2*} = W_n^2(1 + 0.75/n + 2.25/n^2).$$

The null hypothesis of this normality test is that the data are sampled from a normal distribution. When the p-value is greater than the predetermined critical value ( $\alpha = 0.05$ ), the null hypothesis is not rejected and thus we conclude that the data is normally distributed.

### 4.3. Alternative Definitions for a Stability Indicator

The accuracy of the stability indicator highly depends on the accuracy of the order selection algorithm. Thus, we want to verify if the BIC is indeed the appropriate choice for the stability indicator.

We compare the results for the AIC, AICc and the BIC (see chapter 2). For the comparison we simulate ARMA( $p,q$ ) processes for each combination of  $p$  and  $q$ , where  $p, q \in [0, 5]$ , except for the ARMA(0,0) as this is not defined. Furthermore, we use 20 randomly chosen white noise variances with each  $\sigma_w \in (0, 10]$  and time-series length from 10 to 1000. Hence, for each time-series length we test 700 time series. As coefficients we use the ones given in tables 4.1 and 4.2. It was ensured that  $\phi(z)$  and  $\theta(z)$  have no common factors.

$\phi_i$	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)
$\phi_1$	0.2	0.2	0.11	0.11	0.109
$\phi_2$	0	0.3	0.13	0.13	0.113
$\phi_3$	0	0	0.17	0.17	0.127
$\phi_4$	0	0	0	0.19	0.131
$\phi_5$	0	0	0	0	0.137

Table 4.1.: autoregressive coefficients

$\theta_i$	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
$\theta_1$	0.149	0.149	0.149	0.149	0.1031
$\theta_2$	0	0.151	0.151	0.151	0.1033
$\theta_3$	0	0	0.157	0.157	0.1039
$\theta_4$	0	0	0	0.163	0.1049
$\theta_5$	0	0	0	0	0.1051

Table 4.2.: moving average coefficients

As an estimation algorithm for the model order we, again, use the `auto.arima()` function of the R "forecast" package (Hyndman et al. (2019)). This function assumes that the

analysed data is a casual and invertible time series. To ensure causality and invertibility, we chose  $\phi(z)$  and  $\theta(z)$  so that they do not have roots for  $|z| < 1$ . When the AR-polynomial is close to being non-stationary or when the MA-polynomial is close to being non-invertible, then the model is rejected by setting an infinite value for the AIC, AICc and BIC related to that model.

The average amount of correctly estimated p and q is for the BIC 6%, AIC 7% and AICc 7%. But if the model was an ARMA(1,0) it was correctly estimated by the BIC in 50% of the cases while the AIC and AICc only estimated 31% model orders correctly. The figures A.1, A.2 and A.3 (Appendix A) show the results of the tests for  $\tau \in seq(from = 10, to = 1000, by = 50)$ . The values in between are omitted in the figures for clarity purposes. These figures show that the AIC, AICc and the BIC all depend highly on the time-series length  $n$ . Only the BIC estimated 50% or more of the ARMA(1,0) models correctly as long as  $n \geq 430$ . For the AIC and AICc no threshold for the time-series length was detected from which on the model was, in more than 50% of the cases, correctly estimated.

Even though, the AIC and AICc are slightly better in estimating general ARMA(p,q) process the choice of Nevo et al. (2017) to use the BIC in their method appears to be valid due to the fact that the BIC is significantly better in the estimation of the order of an ARMA(1,0) process which is essential for the stability indicator. As the stability indicator is generally applied to shorter time series, in future research other model order selection criteria should be taken under consideration.



## 5. Testing of $\Upsilon$ on Controlled and Observational Data

In this chapter we quantify the reliability of the stability indicator introduced in chapter 4. We start by testing it on a controlled dataset generated by the simplified model for near-surface temperature inversion (see chapter 3) and then proceed by applying  $\Upsilon$ , the stability indicator, to observational data. In the tests we use the `auto.arima()` function from the "forecast" R package [Hyndman et al. (2019)]. The `auto.arima()` function calculates all BIC for all ARMA(p,q) models with  $p \leq p_{max}$  and  $q \leq q_{max}$ , where  $p_{max}$  and  $q_{max}$  are the before mentioned thresholds, and then it chooses the ARMA model with the minimal Bayesian Information Criterion.

### 5.1. Simplified Model by van de Wiel with Added Noise

We start by using the simplified model by Van de Wiel et. al.,

$$\begin{aligned} \frac{dx}{dt} &= \begin{cases} Q_i - \lambda x - Cx(1-x) & \text{for } x \leq 1 \\ Q_i - \lambda x & \text{for } x > 1 \end{cases} \\ x(t_0) &= x_0 \end{aligned} \quad (5.1)$$

which is discussed in detail in chapter 3, to simulate a data set. The "particle"  $x_t$  needs to acquire a sufficient amount of energy from noise so that it can surmount the potential barrier  $\Delta V := \max(V(x)) - \min(V(x))$ , where  $V(x)$  is the potential, and escape from one of the metastable states. The escape time depends on the strength of the noise. When the noise is weak, the particle spends a long time at the metastable state, before being able to escape from it. Therefore, we use the simplified model with additive noise to quantify the recognition of transitions:

$$\begin{aligned} \widetilde{\frac{dx}{dt}} &= \begin{cases} Q_i - \lambda x - Cx(1-x) + w(\omega) & \text{for } x \leq 1 \\ Q_i - \lambda x + w_t(\omega) & \text{for } x > 1 \end{cases} \\ x(t_0) &= x_0 \\ w(\omega) &\sim \text{Gaussian wn}(0, \sigma_w^2), \omega \in \Omega \end{aligned} \quad (5.2)$$

The addition of noise is justified by the fact that observational data contains noise and we aim to apply  $\Upsilon$  to observational data. Moreover, due to the additive noise we can apply a Maximum Likelihood based algorithm (BIC) to identify the best fitting ARMA order which is essential when applying  $\Upsilon$ . For the simulation we fix  $\omega \in \Omega$ . Thereby, we

can use the standard Runge Kutta 4 method to solve this ODE.

For the simulation we need to choose  $\lambda, C \geq C_2$  and  $Q_i$  such that we have two stable points and an unstable one (see figure 3.5).

To be able to quantify the accuracy of the stability indicator we solve  $\widetilde{\frac{dx}{dt}}$  for several  $C$ , starting with a  $C$  close to  $C_2$ , and apply  $\Upsilon$  to the solution  $\tilde{x}(t)$ . The initial parameters are set to  $t_0 = 0$  and  $x(t_0) = \min\{x_{e_i} | i = 1, 2, 3\}$ . To generate the controlled data set the following parameters are used:

$$\begin{aligned}\lambda &= 2, \\ Q_i &= 2.5, \\ C &\in \text{seq}(\min = 5.3, \max = 7.2, by = 0.1), \\ w &\sim \text{Gaussian wn}(0, 3.5^2), \\ n &= 2000 \text{ and} \\ h &= 0.01.\end{aligned}$$

The standard deviation  $\sigma_w = 3.5$  for the white noise process is chosen such that we obtain several time series which transition. Moreover, by choosing  $\sigma_w$  this way the resulting time series resemble the observational data. This range for  $C$  is chosen because for these values the series transitions from one stable state to another. To choose the window length  $\tau$  we apply both the K-Means Algorithm (4.2.1) and the Anderson Darling Test (4.2.2) to the dataset. A detailed description of both algorithms/tests is given in chapter 4.

### K-Means Algorithm

The K-Means algorithm can be used to estimate the amount of points covering the transition time. We set the cluster number to three as we have three equilibria. The results of the clustering algorithm are exemplarily shown in figure 5.1 for  $C = 6.4$ . Note:  $t = [0, n \cdot h] = [0, 20]$ .

In this case the equilibria are  $x_{e_1} = 0.46$  (stable),  $x_{e_2} = 0.97$  (unstable) and  $x_{e_3} = 1.25$  (stable). The cluster centers are 0.46, 0.97 and 1.31 which are a close approximation of the equilibria. Therefore, we expect a good estimation for the amount of points covering the transition. The mean times spend in each cluster are:

- $mean(T_1) = 112.2$
- $mean(T_2) = 94$
- $mean(T_3) = 286.67$

where  $T_i$  is the time spend in cluster  $i \in \{1, 2, 3\}$ . The window length  $\tau$  is chosen such that it is smaller than the minimal mean time spent in one cluster, i.e. for  $C = 6.4$  we choose  $\tau < \tau_{KMeans} := \min\{mean(T_i) | i = 1, 2, 3\} = 94$ . To simplify the implementation

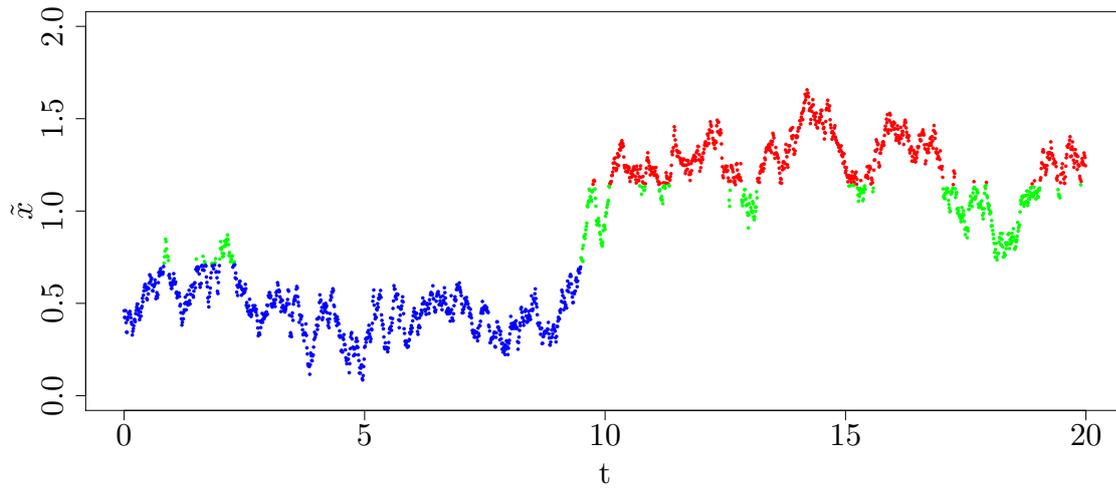
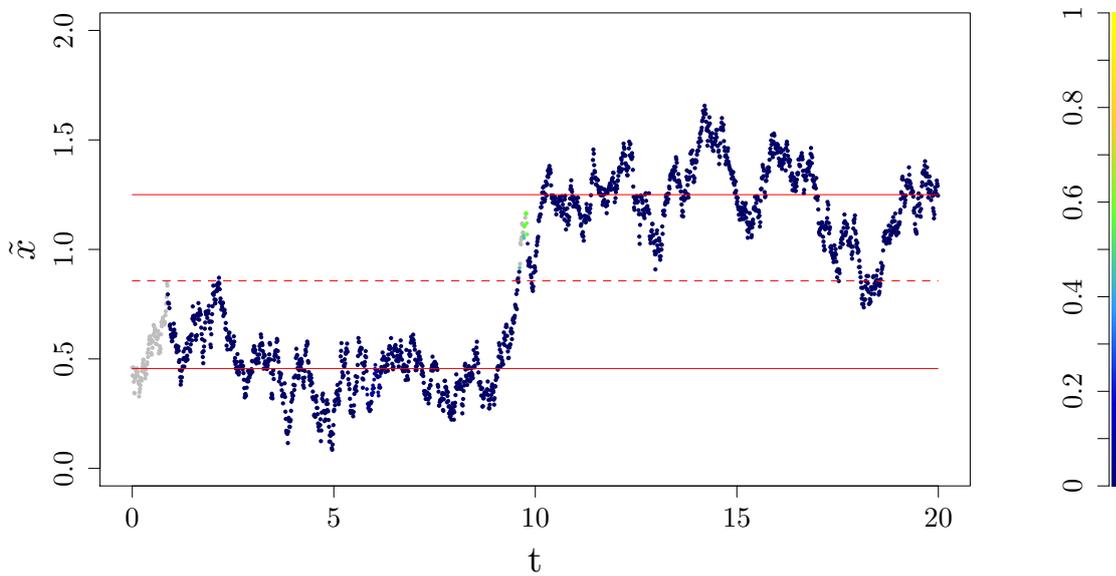


Figure 5.1.: Clustered time series

we choose  $\tau = \min\{mean(T_i)|i = 1, 2, 3\} - 5$  for all tested  $C$ . By applying  $\Upsilon$  to the data generated by the simplified model with 6.4 we get the results which are shown in figure 5.2. As a reminder: The solid red lines correspond to the stable equilibria and the dotted red line to the unstable one. The colours ranging from dark blue to yellow represent the stability of the points measured by  $\Upsilon$  and we always colour the last point of the subsequence (see chapter 4). The first transition through the unstable equilibrium

Figure 5.2.:  $\Upsilon$  for simplified model with  $C = 6.4$  and added noise ( $\tau = 84$ ).

is well recognised (green dots after the dotted red line) but the second transition is not. Considering that  $\tau$  has to be chosen in advance and is fixed for all subsequences this is

expected. This is clearly one limitation of the stability indicator as the amount of points covering the transition may vary in one time series, like in this case. The estimated amounts of points covering the transition for all values for  $C$  are given in table 5.1.

$C$	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0	6.1	6.2
$\tau_{KMeans}$	59	67	65	46	57	57	68	77	91	70
$C$	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0	7.1	7.2
$\tau_{KMeans}$	94	94	85	87	88	88	142	114	99	148

Table 5.1.:  $\tau$  given by K-Means algorithm

Figure 5.3 is a bifurcation diagram for  $C \in \text{seq}(\min = 5.3, \max = 7.2, \text{by} = 0.1)$  and  $\tau = \min\{\text{mean}(T_i) | i = 1, 2, 3\} - 10$ . For this and the following bifurcation plots we use a different type of colour coding than in the plots for one time series. The darker the dots the higher is the  $\Upsilon$  value. Most of the transitions are well recognised and almost all of the stable states are recognised as stable. Nonetheless, there are some transitions which were recognised significantly too late, e.g. for  $C = 5.4$ . In future research it could be of interest to apply a more sophisticated clustering algorithm and to research methods to adapt the window length dynamically.

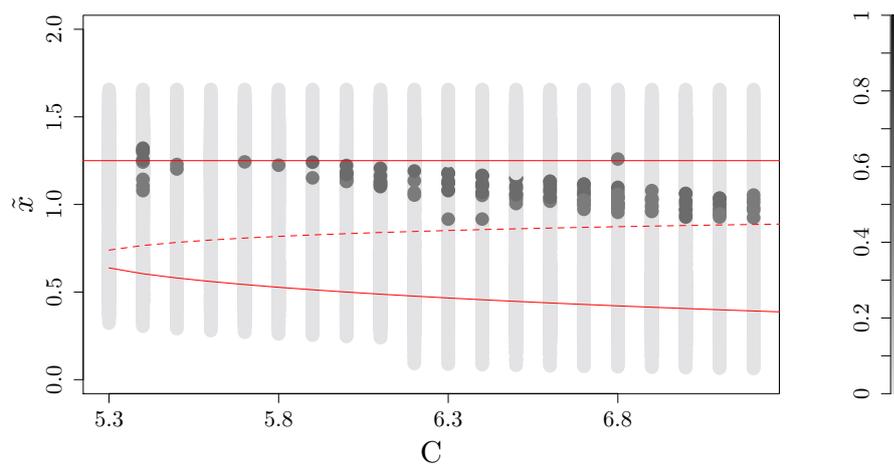


Figure 5.3.: Bifurcation Diagram with  $\tau = \min\{\text{mean}(T_i) | i = 1, 2, 3\} - 5$ .

### Anderson Darling Test

The Anderson Darling Test can be used to find the biggest  $\tau$  for which we can assume that most of the subsequences are sampled from a normal distribution. For example, for  $C = 6.4$  the Anderson Darling Test yields that for  $\tau = 65$  the median of the p values for all subsequences is greater than the significance level 0.05. This can be seen in figure 5.4. The solid line in the grey boxes is the median of the p values for a fixed  $\tau$  while the upper and lower border of the grey boxes refer to the upper and lower quartile of the p values. The dotted horizontal line is the significance level. The values of  $\tau$  for all values

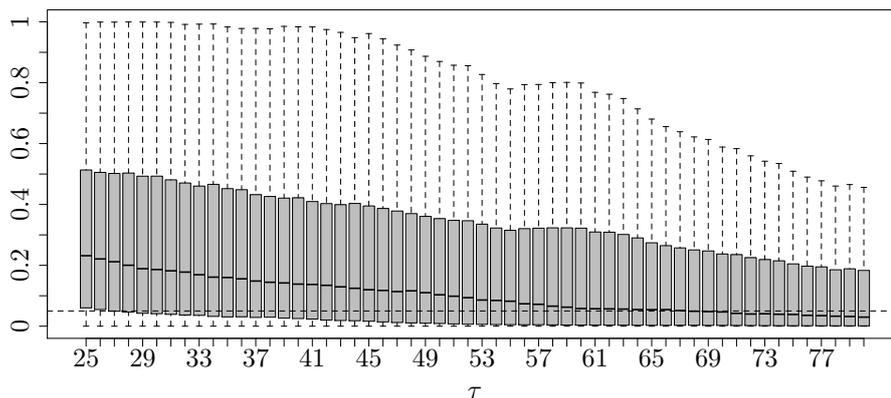


Figure 5.4.: Boxplot of the p values from the Anderson Darling Test for  $C = 6.4$ .

for  $C$ , given by the Anderson Darling Test, are stated in table 5.2.

$C$	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0	6.1	6.2
$\tau_{AD}$	60	62	64	63	63	63	62	62	65	64
$C$	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0	7.1	7.2
$\tau_{AD}$	68	67	65	64	65	67	66	65	69	70

Table 5.2.:  $\tau$  given by Anderson Darling Test

Figure 5.5 and 5.6 show the results for  $\Upsilon$ . They are very similar to the ones where we chose  $\tau$  according to the K-means algorithm except that for  $C \leq 5.6$  there are several points marked as unstable even though they are stable. This is due to the fact that for these  $C$ 's the  $\tau$ 's chosen by the Anderson Darling Test are bigger than the ones estimated by the K-Means algorithm. Therefore, they are most likely bigger than the amount of points covering the transition. For  $C \geq 5.9$  the  $\tau$ 's given by the AD Test are smaller than the ones of the K-Means algorithm. In these cases  $\Upsilon$  gives a good indication for the stability. The summary for all  $\tau$  is given in table 5.3 (note: 1 = TRUE) and figure 5.7 is a bifurcation plot with  $\tau_{AD} < \tau_{KMeans} - 5$ . In this case we see that  $\Upsilon$  is capable of recognizing the location of unstable equilibria for all  $C$  and also the one of stable equilibria. Therefore, we are confident that we can apply  $\Upsilon$  to observational data.

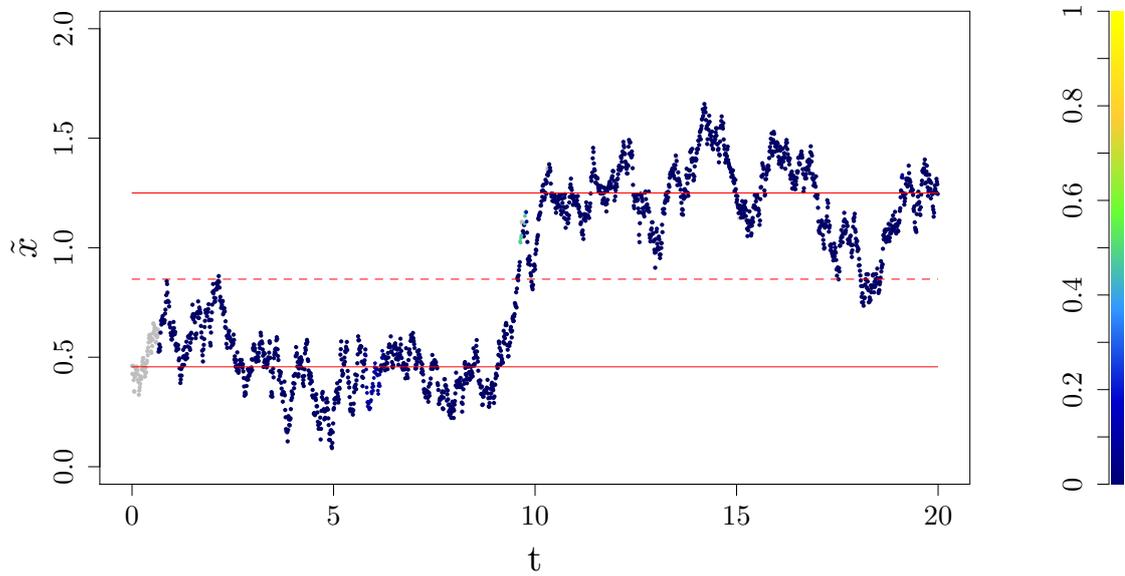


Figure 5.5.:  $\Upsilon$  for simplified model with  $C = 6.4$  and added noise ( $\tau = 67$ ).

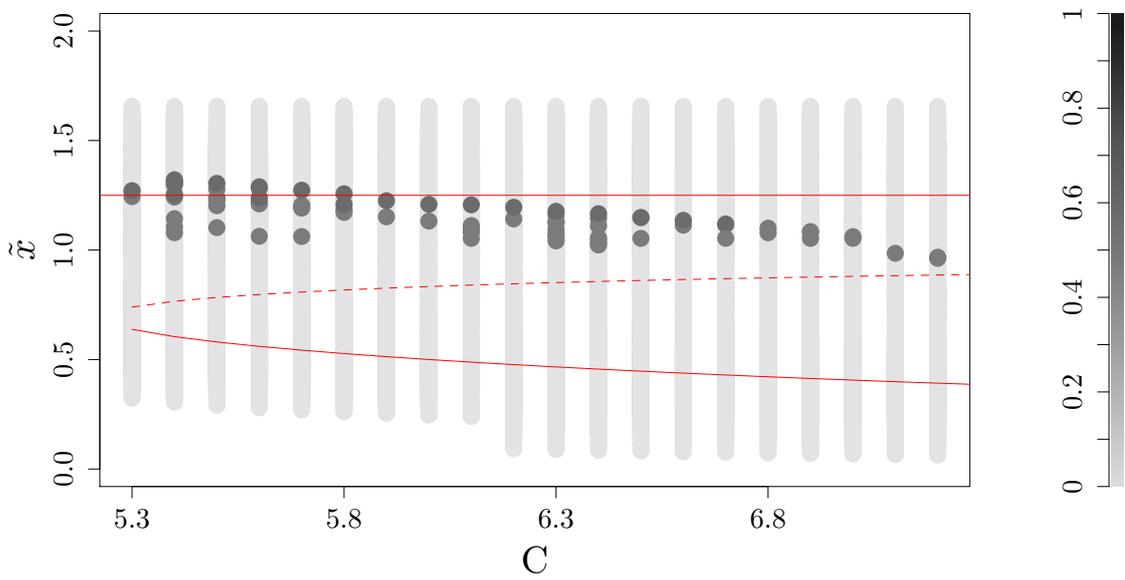
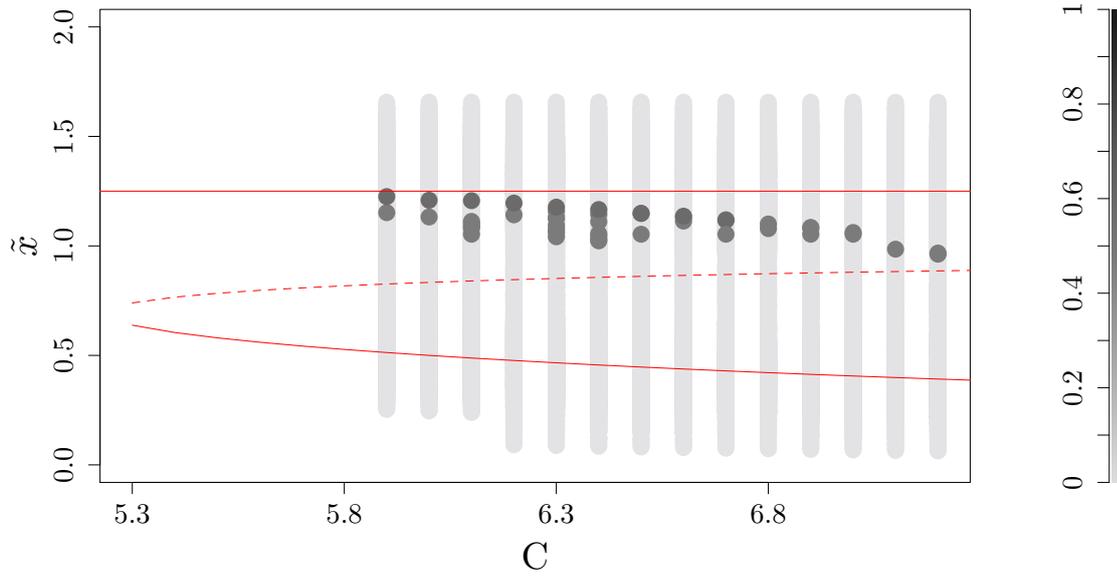


Figure 5.6.: Bifurcation Diagram with  $\tau$  chosen by Anderson Darling Test (see table 5.2).

Figure 5.7.: Bifurcation Diagram with  $\tau_{AD} < \tau_{KMeans} - 5$  (see table 5.3).

Index	$C$	$\tau_{AD}$	$\tau_{KMeans}$	$\tau_{AD} < \tau_{KMeans} - 5$
1	5.3	60	59	0
2	5.4	62	67	0
3	5.5	64	65	0
4	5.6	63	46	0
5	5.7	63	57	0
6	5.8	63	57	0
7	5.9	62	68	1
8	6.0	62	77	1
9	6.1	65	91	1
10	6.2	64	70	1
11	6.3	68	94	1
12	6.4	67	94	1
13	6.5	65	85	1
14	6.6	64	87	1
15	6.7	65	88	1
16	6.8	67	88	1
17	6.9	66	142	1
18	7.0	65	114	1
19	7.1	69	99	1
20	7.2	70	148	1

Table 5.3.: Summary of the results from the Anderson Darling Test and K-Means algorithm.

## 5.2. Dumosa

The first observational data we use to test the stability indicator are temperature measurements from a site in Dumosa, Victoria, Australia. The temperature measurements



Figure 5.8.: Dumosa measurement set up [provided by Danijel Belušić].

were made at height 6 and 3m and the wind measurements at 6m. The frequency of measurements is 1 minute. As we want to use data where we can expect temperature inversions to take place we exclusively use evening and night time data from March until June 2013 (89 days). We apply  $\Upsilon$  to two nights, one in May and one in April, as these two nights exhibit the S-shape dependence between the temperature difference  $\Delta T = T_{6m} - T_{3m}$  and the wind speed  $U_{6m}$  mentioned in chapter 3. Again we use the K-Means algorithm and the Anderson Darling Test to choose the window length  $\tau$ .

### Night in May

Figure 5.9 shows the results for the Anderson Darling Normality Test and the K-means clustering algorithm. According to these tests the maximal  $\tau$  for which we can assume normality is 27 and the  $\tau$  which is smaller than the amount of points covering the transition is 22. Unlike in the previous section we apply  $\Upsilon$  to subsequences of length  $\tau_{KMeans}$ , not  $\tau_{KMeans} - 5$ , because  $\tau_{KMeans}$  is already very small and making the subsequences shorter can increase the risk of fitting the wrong model. Moreover,  $\tau_{AD}$  is not smaller than  $\tau_{KMeans}$ . Nonetheless, there are hardly any differences in the results with the two different  $\tau$ 's and this shows that  $\Upsilon$  is capable of recognising the stability of the points. The results for  $\Upsilon$  with both  $\tau$ 's are given in figures 5.10 - 5.13. Especially, the time periods with small fluctuations and the transition at time point 800 are well recognised.

In figures 5.12 and 5.13 the regime shift from a very stable boundary layer (high  $\Delta T$  values, small  $U$  values) to a weakly stable boundary layer (small  $\Delta T$  values, high  $U$  values) with an intermittent regime in between is clearly detected by  $\Upsilon$ . However, the very stable regime is not very pronounced. We have two hypothesis for that:

1. The observed night is not long enough to form a distinct very stable regime.
2. The surface cooling is not strong enough to reach a very strong temperature inversion.

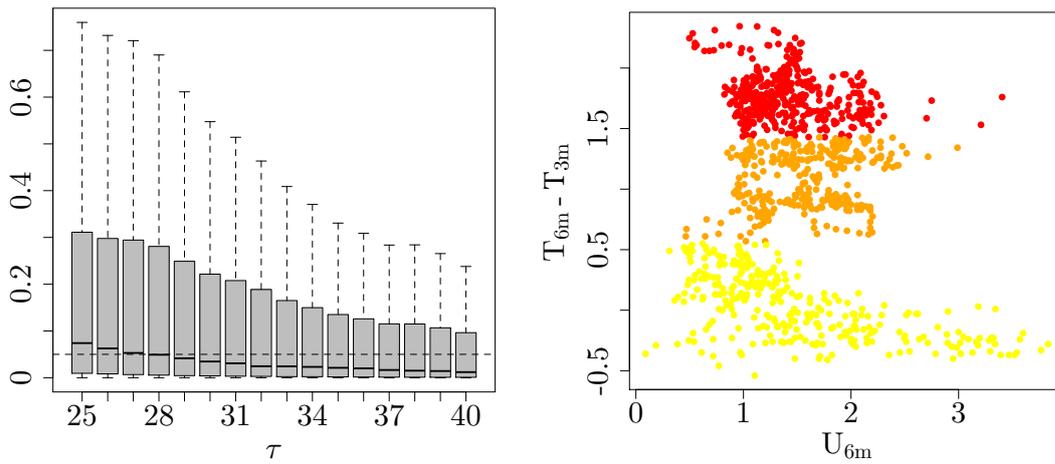


Figure 5.9.: Left: Boxplot of the p values from the Anderson Darling test.  
Right: Clustered data with K-Means.

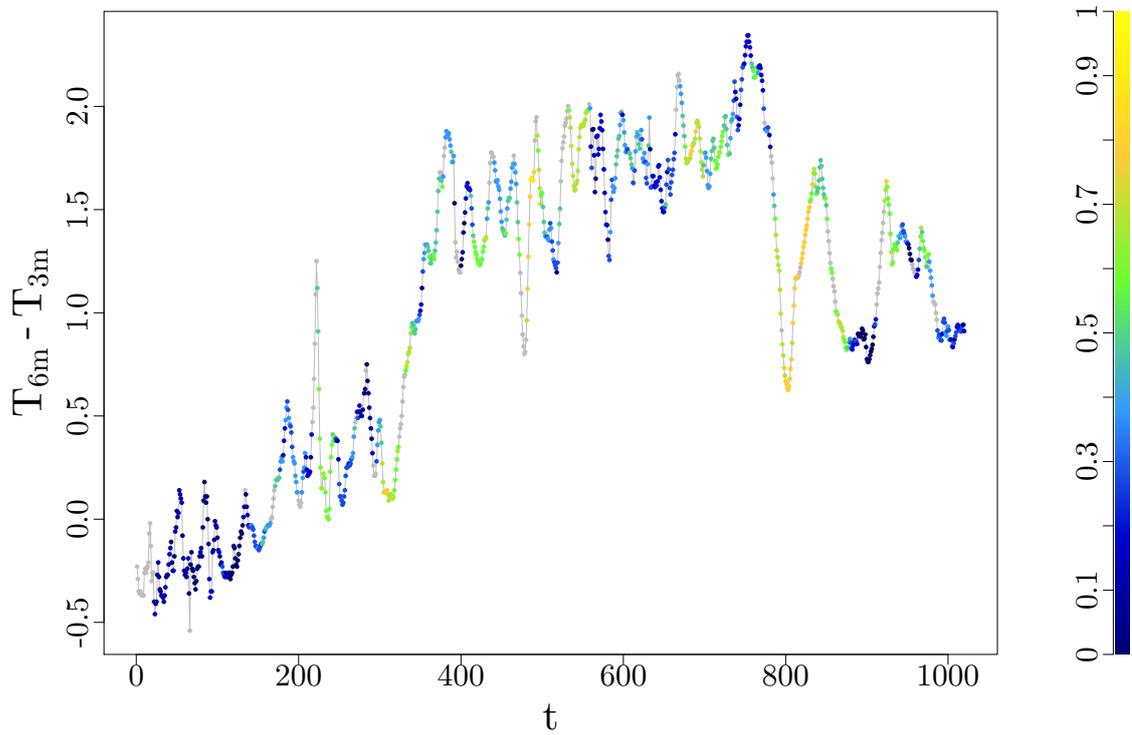


Figure 5.10.: Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in May 2013.  $\tau = 22$  (K-Means)

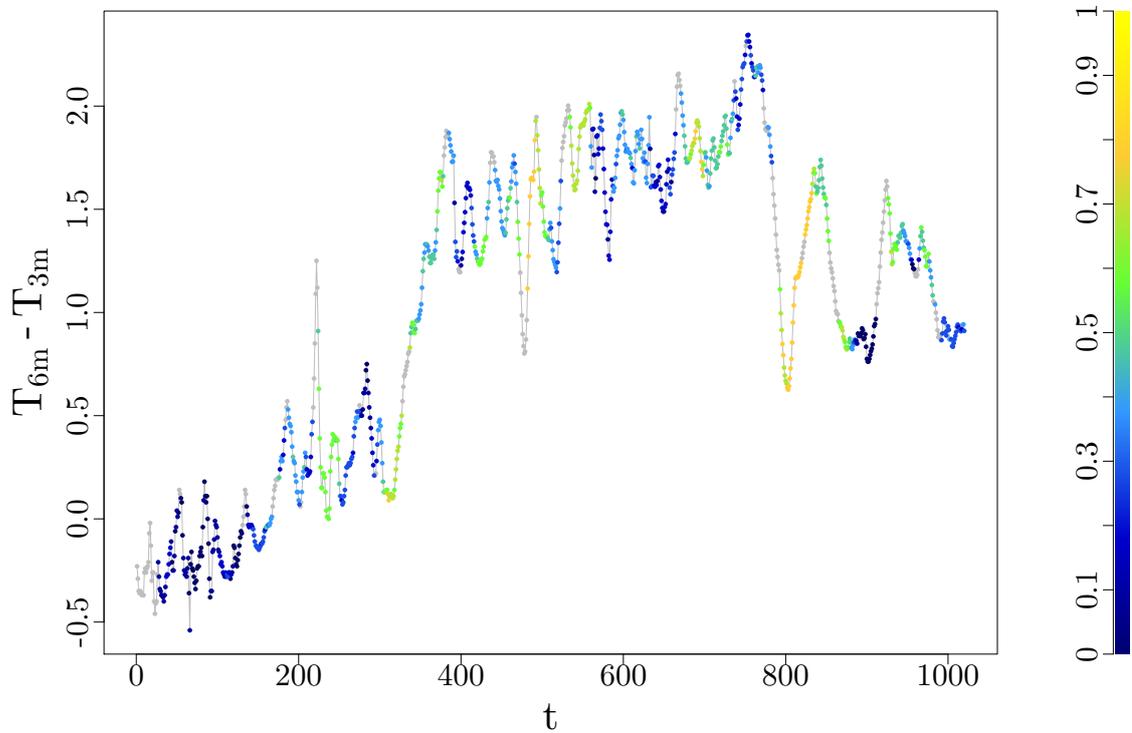


Figure 5.11.: Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in May 2013.  $\tau = 27$  (AD Test)

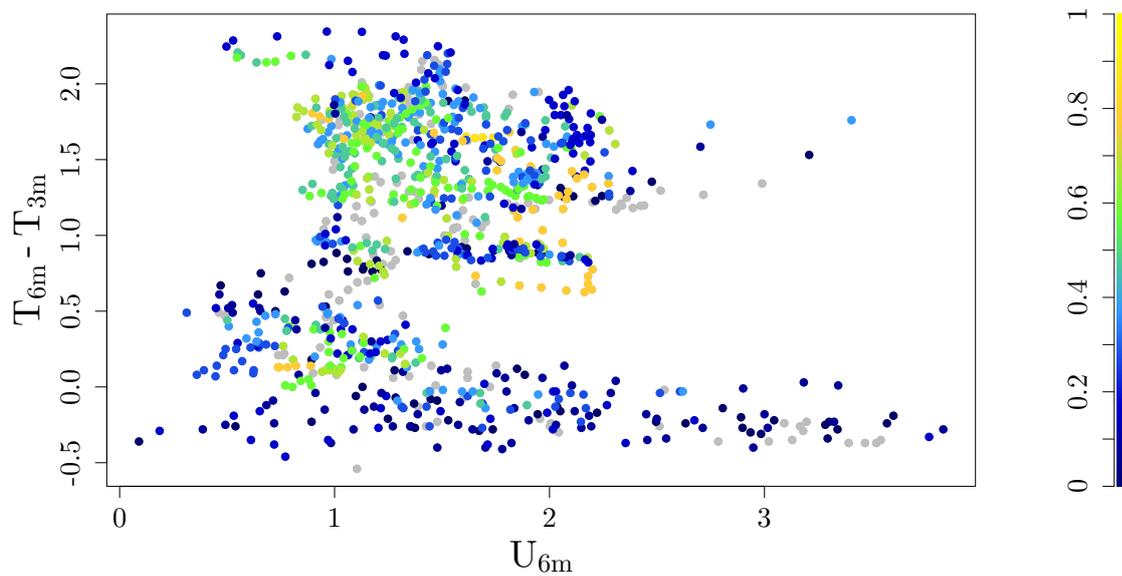


Figure 5.12.: Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in May 2013.  $\tau = 22$  (K-Means)

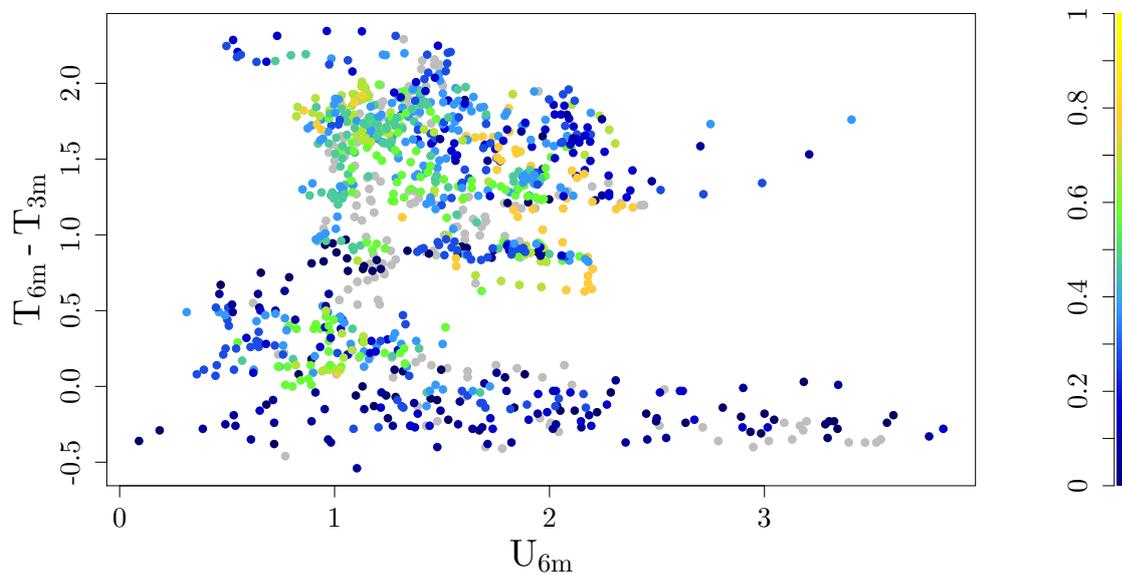


Figure 5.13.: Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in May 2013.  $\tau = 27$  (AD Test)

### Night in April

Figure 5.14 shows the results for the Anderson Darling Normality Test and the K-means clustering algorithm. According to these tests the maximal  $\tau$  for which we can assume normality is 30 and the  $\tau$  which is smaller than the amount of points covering the transition is 23. Again  $\tau_{AD}$  is not smaller than  $\tau_{KMeans}$ . But, they are still close. The results for  $\Upsilon$  with both  $\tau$ 's are given in figures 5.10 - 5.13. Again the results with both

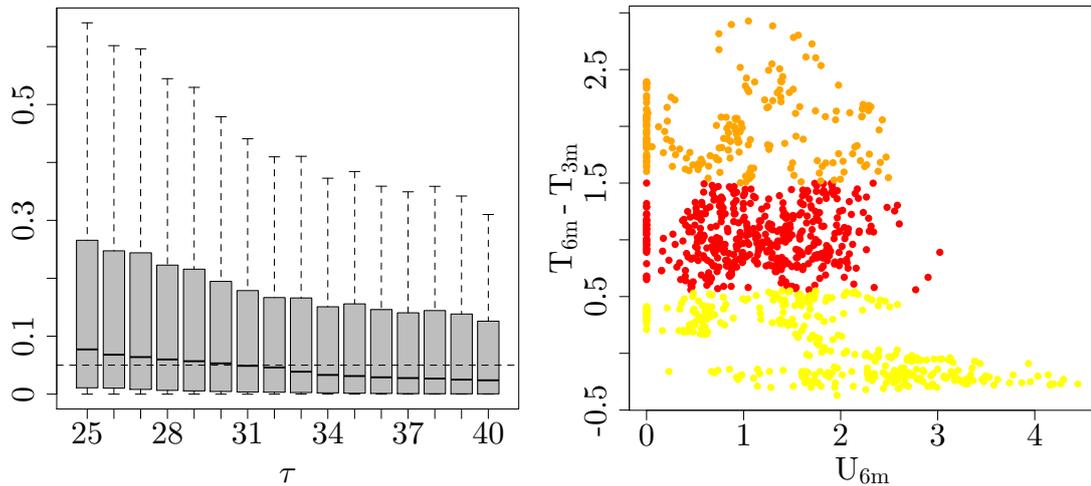


Figure 5.14.: Left: Boxplot of the p values from the Anderson Darling test.  
Right: Clustered data with K-Means.

$\tau$ 's look very similar but the transitions are not as well recognised as in the data set from May. Moreover, in the bifurcation plot there is no distinct separation into two stable and one unstable cluster like in the time series from May. In fact, the results of  $\Upsilon$  suggest one stable regime for small  $\Delta T = T_{6m} - T_{3m}$  and one intermittent for higher  $\Delta T$ . Therefore, we assume that, possibly due to the same but more pronounced reasons mentioned before, the temperature inversion was less notable than during the measurements in May.

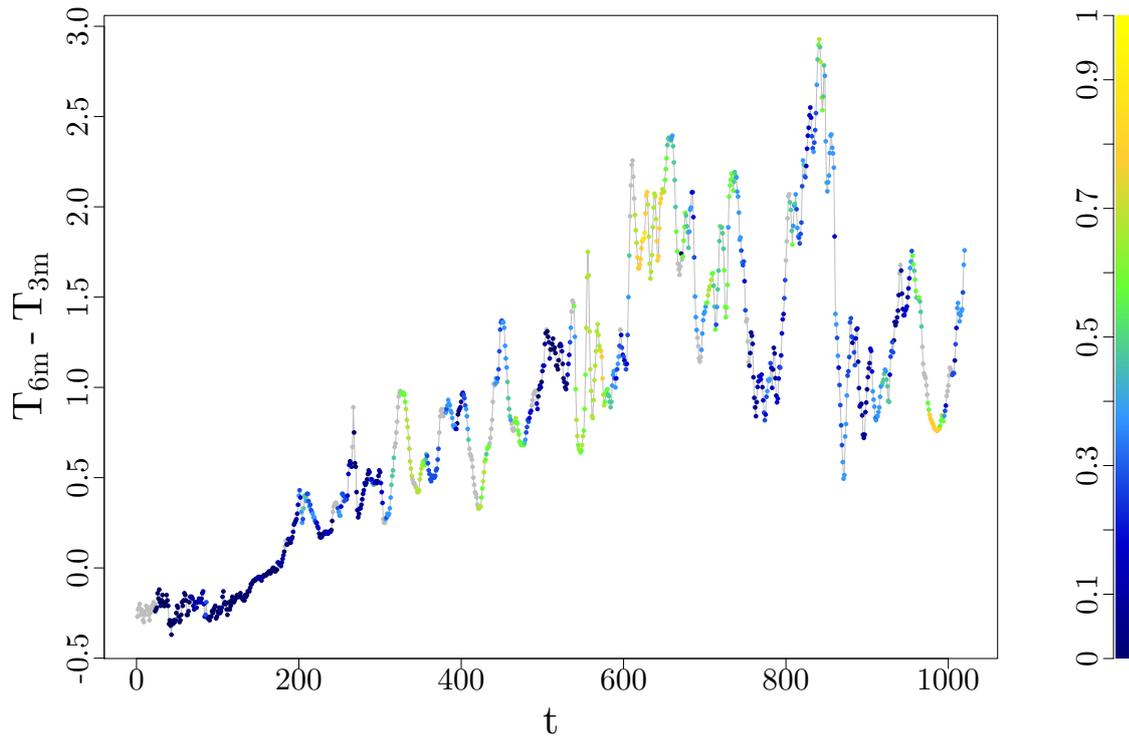


Figure 5.15.: Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in April 2013.  $\tau = 23$  (K-Means)

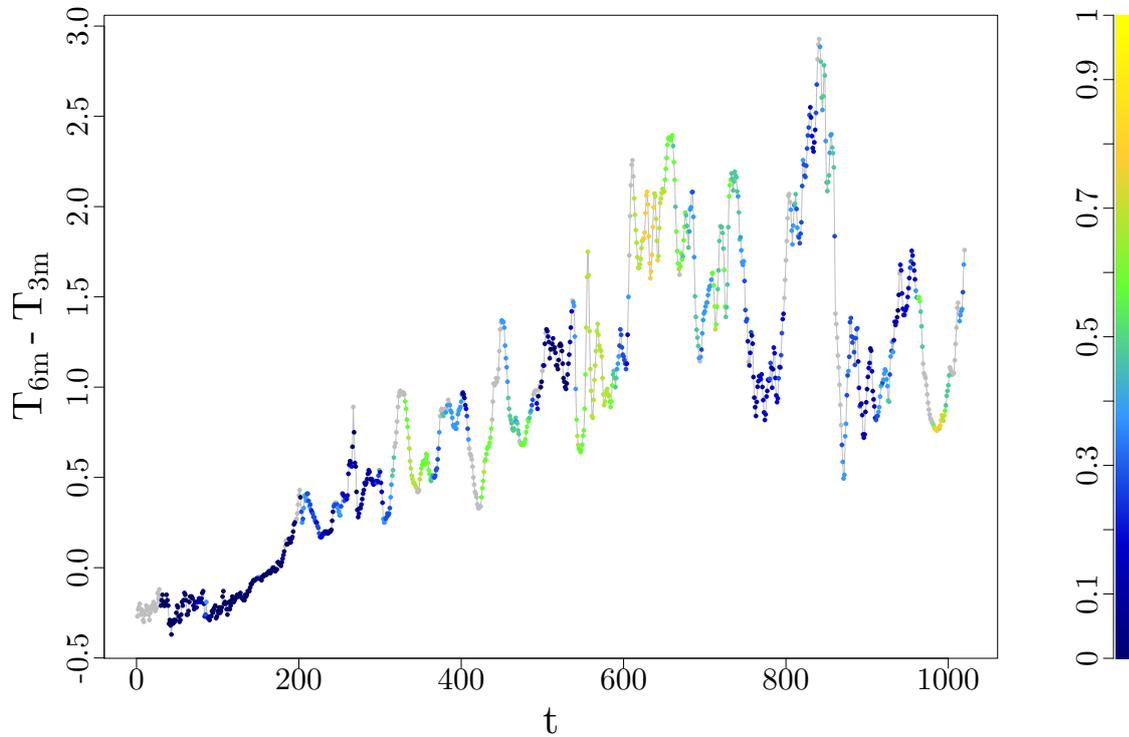


Figure 5.16.: Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in April 2013.  $\tau = 30$  (AD Test)

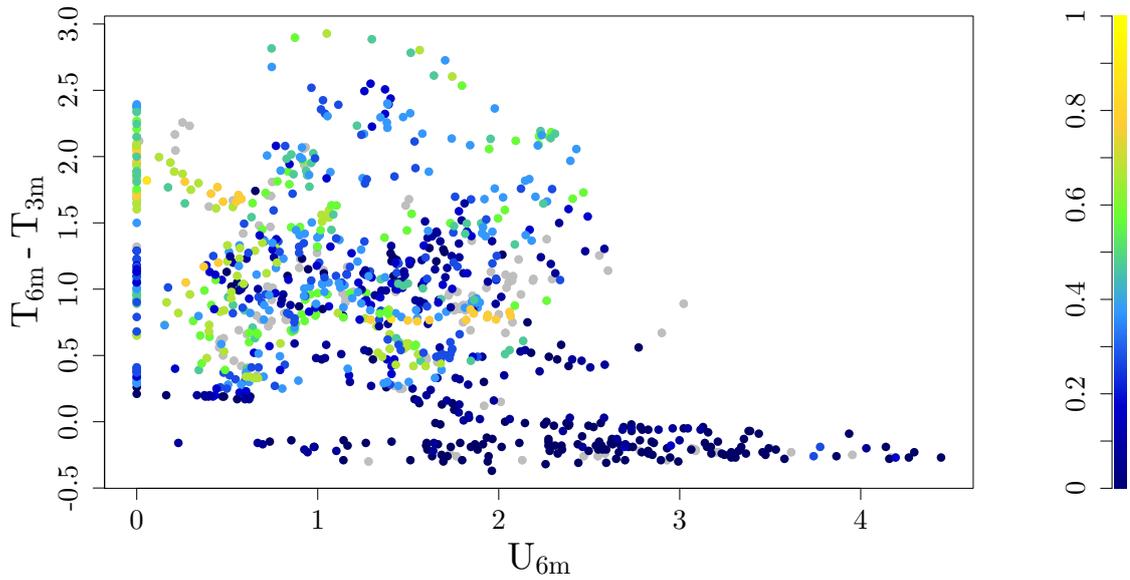


Figure 5.17.: Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in April 2013.  $\tau = 23$  (K-Means)

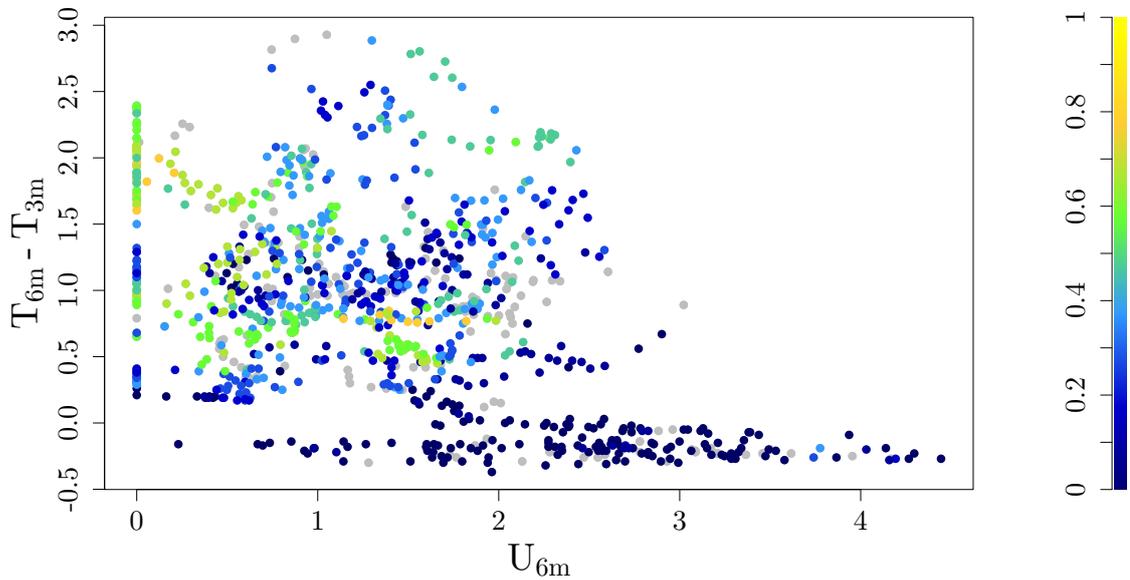


Figure 5.18.: Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in April 2013.  $\tau = 30$  (AD Test)

### 5 min Block Averaged Data for Night in May

Tests of  $\Upsilon$  on other data sets (see section 5.3) suggested that the data frequency is crucial for the reliability of the results. Observational is often stored in block averages, e.g. measurements for 10 min are averaged into 1 data point. The issue with this can be that the data frequency can be too low for the observed transition. In more detail, if the time taken by the system to transition from one metastable state to the next represents is less than about 25 discrete measurement points (the minimum needed to have relevant statistical results), then the approach fails. Therefore, we need to find a criteria to decide if the data frequency is high enough to give reliable results for  $\Upsilon$ . We start by block averaging the temperature measurements for the night in May, we discussed earlier, such that we have 5 min data instead of 1 min. Thereby, we reduce the length of the time series from 1020 to 204 data points. Again we choose  $\tau$  with the Anderson Darling Test and the K-Means algorithm (see figure 5.19). The results of both

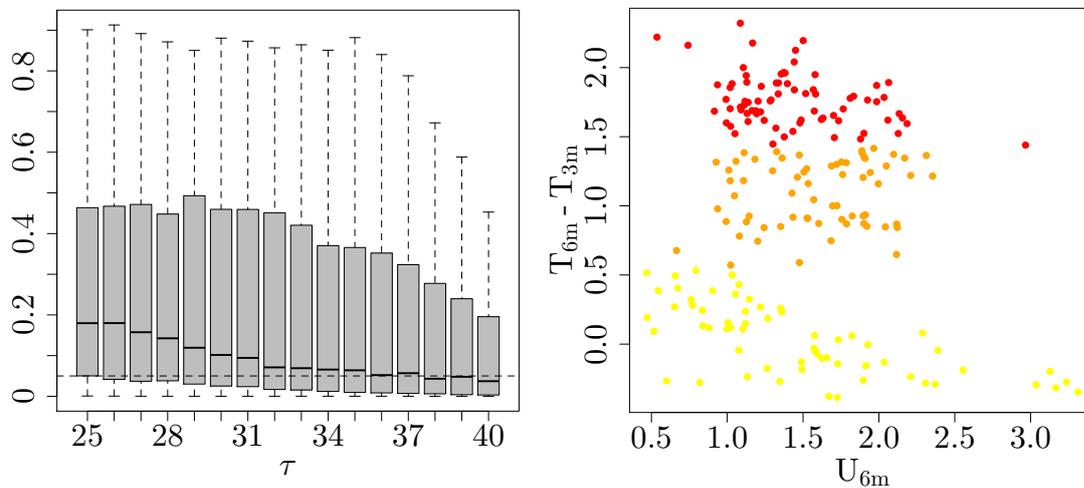


Figure 5.19.: Left: Boxplot of the p values from the Anderson Darling test.  
Right: Clustered data with K-Means.

tests/algorithms are compared to the ones for the 1 min data in table 5.4.

	$\tau_{KMeans}$	$\tau_{AD}$
1 min	23	30
5 min	6	37

Table 5.4.:  $\tau$  given by K-Means algorithm and Anderson Darling Test

There is a clear distinction between the  $\tau$  estimated by the Anderson Darling Test and the one given by the K-Means algorithm for the 5 min data contrary to the 1 min data. The small value for  $\tau$  given by the K-Means algorithm suggests that there is only a small amount of points covering the transition time and we can not fit an ARMA(p,q) model

properly to subsequences this short. This is shown in figures 5.20 and 5.22. Moreover, as the value for  $\tau$  given by the Anderson Darling Test is much bigger than the amount of points covering the transition time we do not expect reliable results for  $\Upsilon$  with this  $\tau$ . Figures 5.21 and 5.23 confirm this hypothesis. In fact, hardly any transitions are recognised.

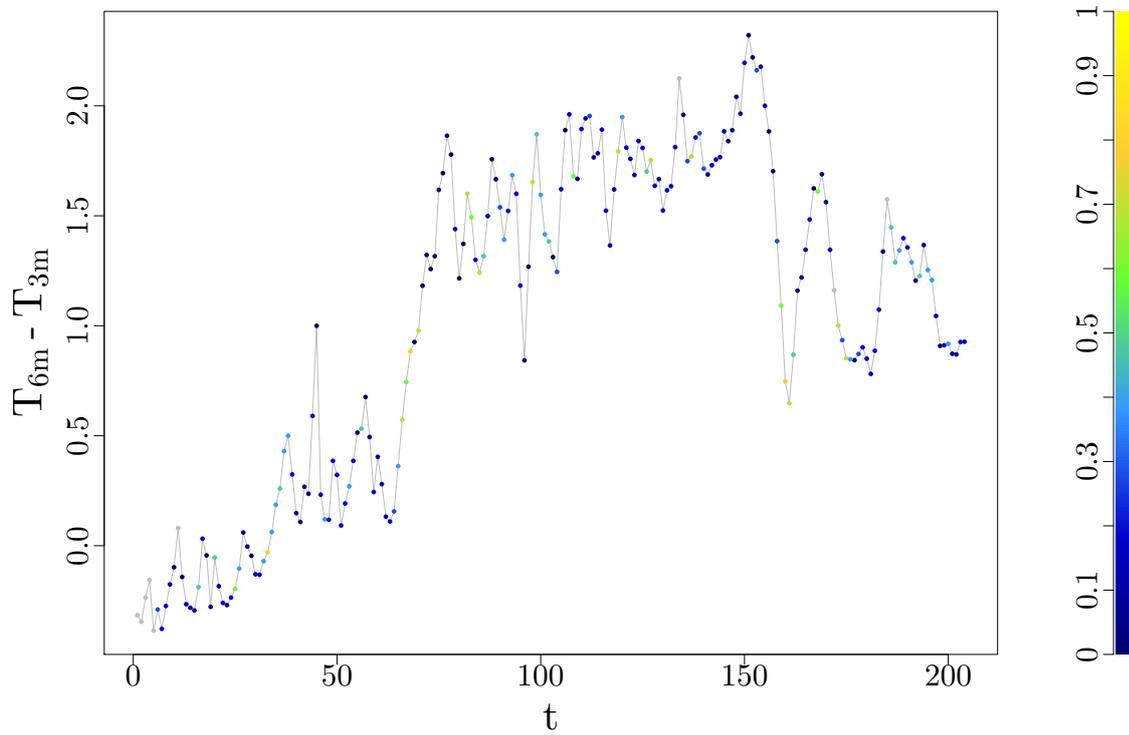


Figure 5.20.: Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in May 2013.  $\tau = 6$  (K-Means)

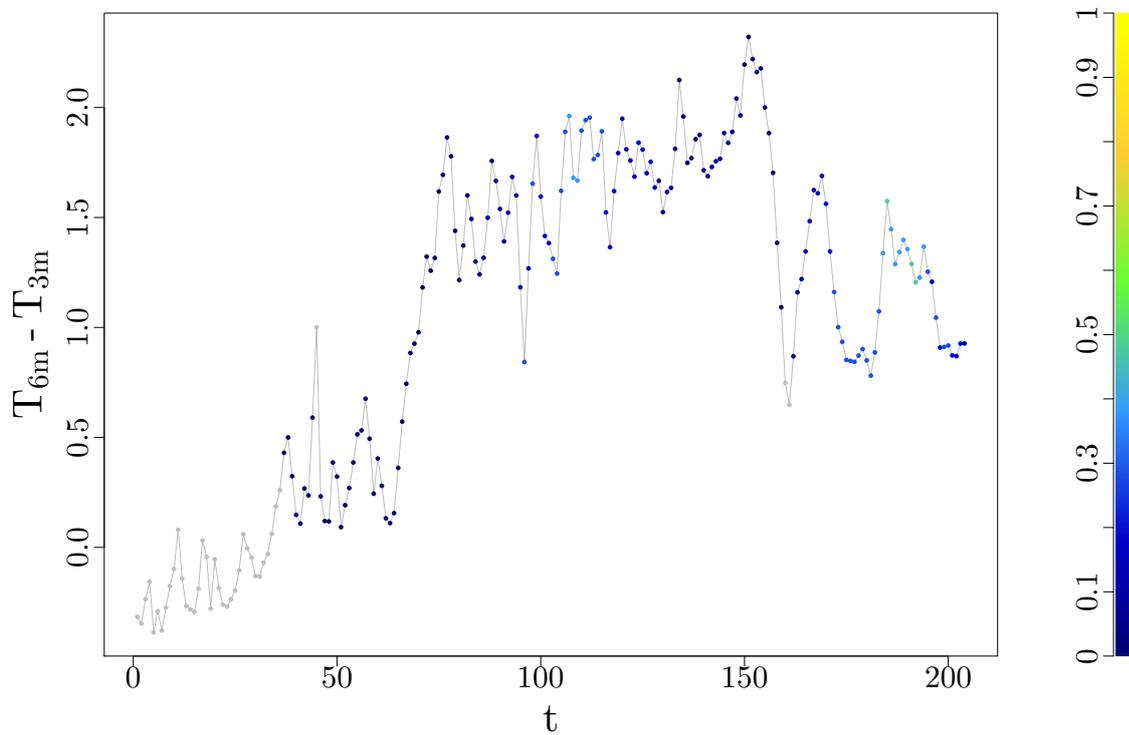


Figure 5.21.: Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in May 2013.  $\tau = 37$  (AD Test)

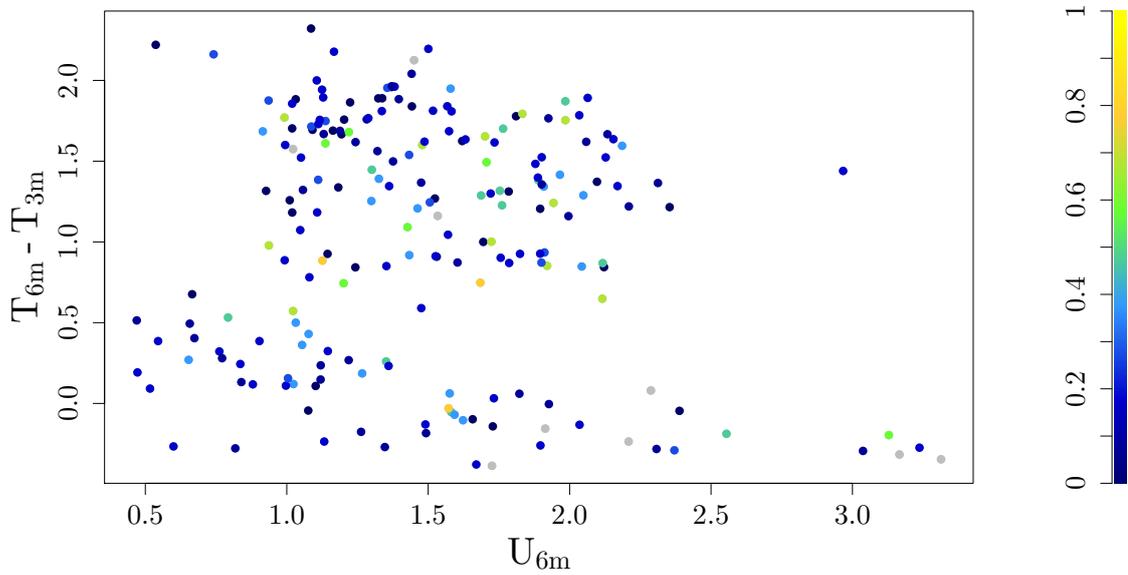


Figure 5.22.: Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in May 2013.  $\tau = 6$  (K-Means)

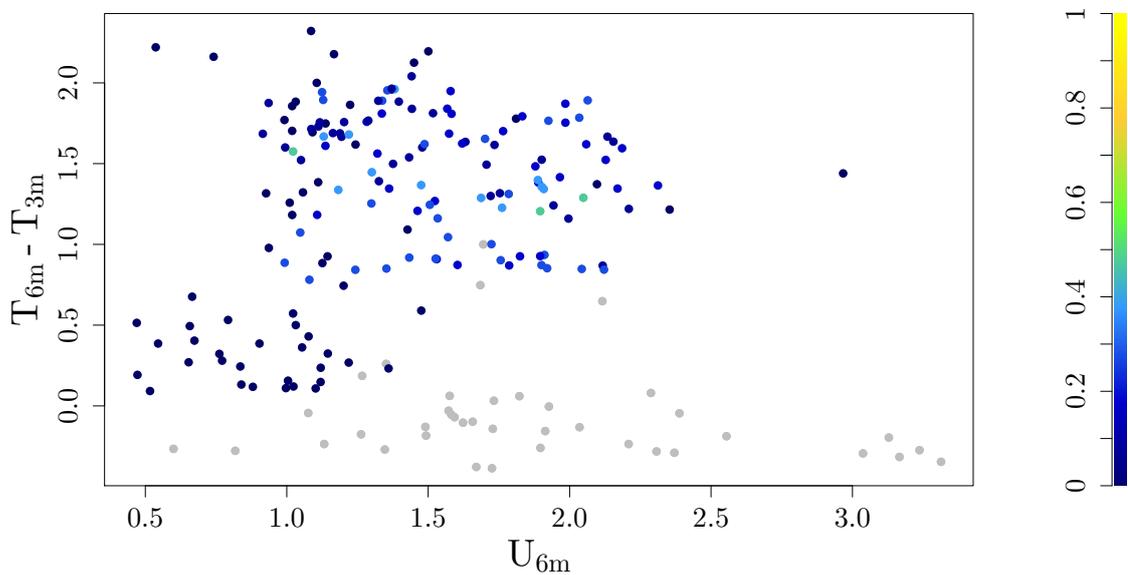


Figure 5.23.: Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in May 2013.  $\tau = 37$  (AD Test)

### 5.3. Dome C Data

The Dome C data was measured at the Concordia Research Station which is located on the Antarctica Plateau. It is a French-Italian research facility that was built 3233 m above sea level. The Dome C dataset contains 10 min mean meteorological data from 2017. Important for our analysis are measurements of the temperature (Kelvin) at height 9.4 m and surface, the wind speed (m/s) at height 8 m and the radiation made in the polar night which is from March to September. We focus on the polar night as we want to test  $\Upsilon$  on a data set with several temperature inversions taking place. The height of the instruments was measured on 26<sup>th</sup> of December. Like in the paper by [van de Wiel et al. \(2017\)](#) the data is classified into two subcategories of radiative forcing being the sum of net shortwave and incoming longwave radiation:  $R_+ = K^\downarrow - K^\uparrow + L^\downarrow$ . When plotting  $\Delta T = T_{9.4m} - T_s$  over the wind speed  $U_{8m}$  we see the before mentioned S-shape for the measurements with  $R^+ < 80Wm^{-2}$ . Therefore, we focus on the case when  $R^+ < 80Wm^{-2}$ .

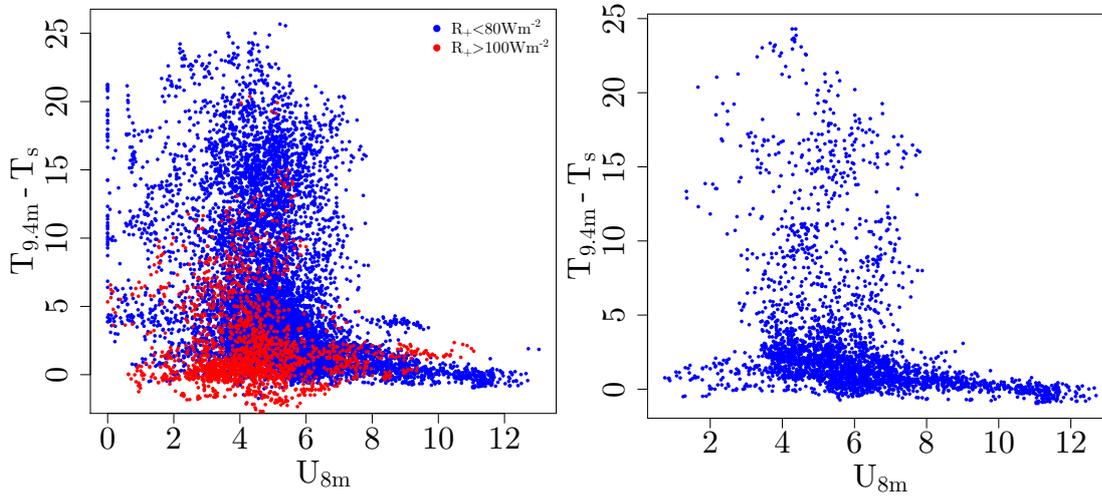


Figure 5.24.: Temperature inversion between 9.4 m and the surface, as a function of wind speed as observed at Dome C in polar winter 2017. Left: All data points. Right: Longest consecutive time series with  $R^+ < 80Wm^{-2}$ .

We apply  $\Upsilon$  to the longest consecutive time series with  $R^+ < 80Wm^{-2}$  which is from 2017-08-03 10:50 to 2017-08-24 21:50 (see right plot in figure 5.24), i.e. 3091 data points. Again we choose  $\tau$  with the Anderson Darling Test and the K-Means algorithm (see figure 5.25 and 5.26). Similar to the results from the 5 min averaged Dumosa data the value for  $\tau$  given by the Anderson Darling Test  $\tau_{AD} = 43$  is much bigger than the one given by the K-Means algorithm ( $\tau_{KMeans} = 10$ ), which is in fact too small to expect a good fit for the ARMA(p,q) model and again we see that hardly any transitions are recognized by  $\Upsilon$  with  $\tau_{AD}$  but with  $\tau_{KMeans}$  some transitions are noted. In fact, the bifurcation plot with  $\tau_{KMeans}$  indicates that there was one stable regime and an intermittent one during the measurement period.

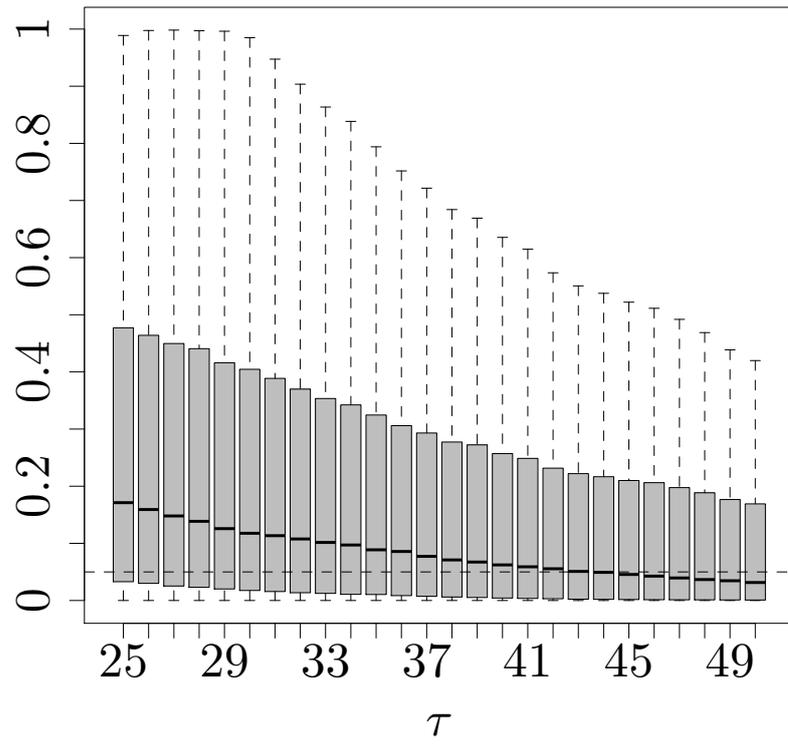


Figure 5.25.: Boxplot of the p values from the Anderson Darling test.

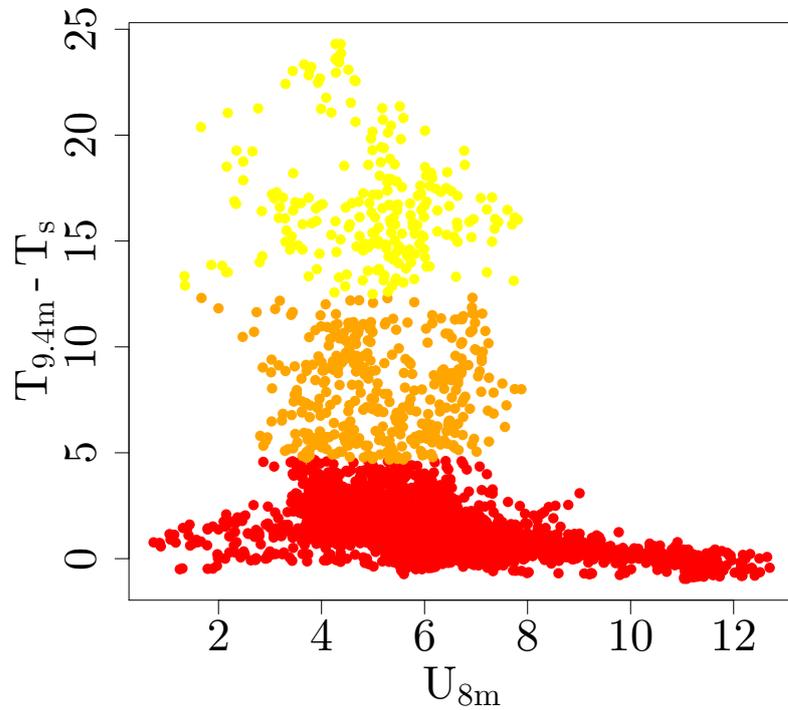


Figure 5.26.: Clustered data with K-Means.

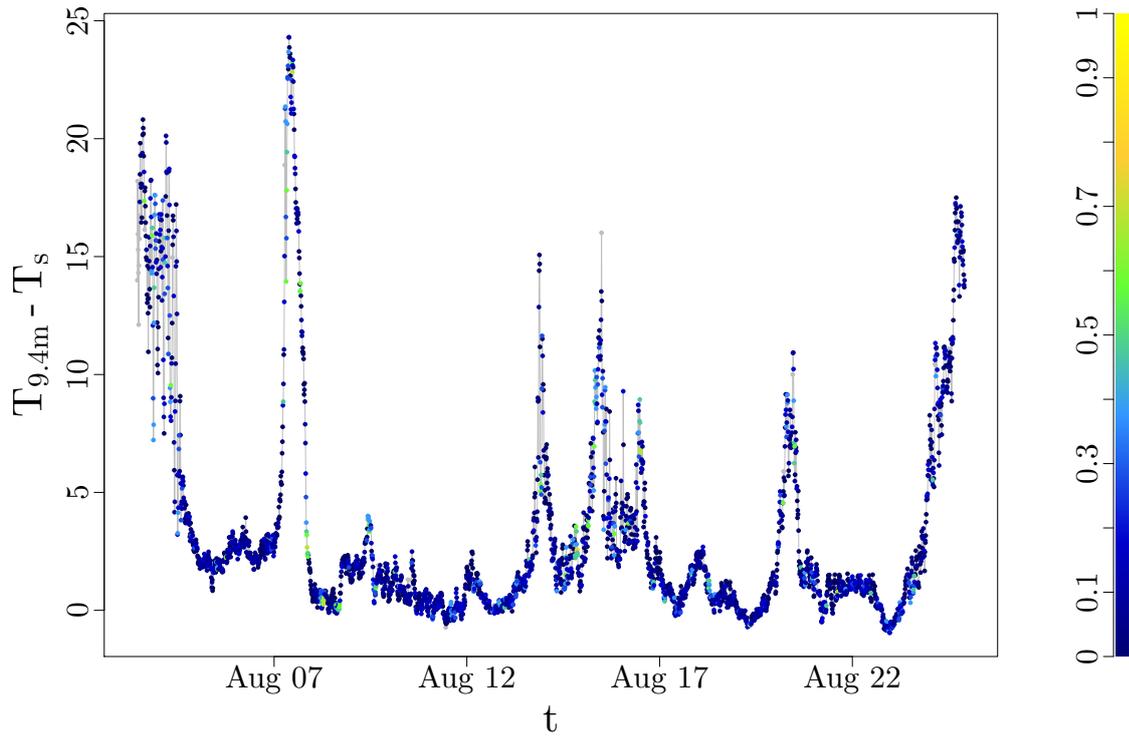


Figure 5.27.: Temperature inversion between 9.4 m and the surface, as a function of time as observed at Dome C.  $\tau = 10$  (K-Means)

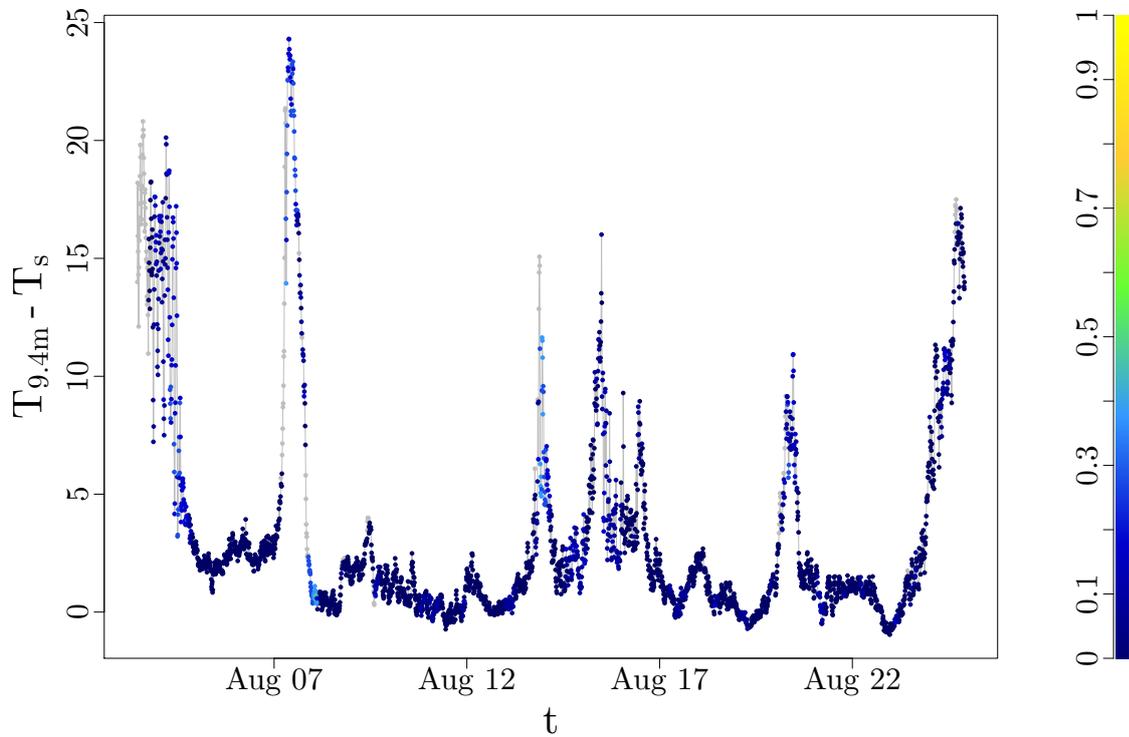


Figure 5.28.: Temperature inversion between 9.4 m and the surface, as a function of time as observed at Dome C.  $\tau = 43$  (AD Test)

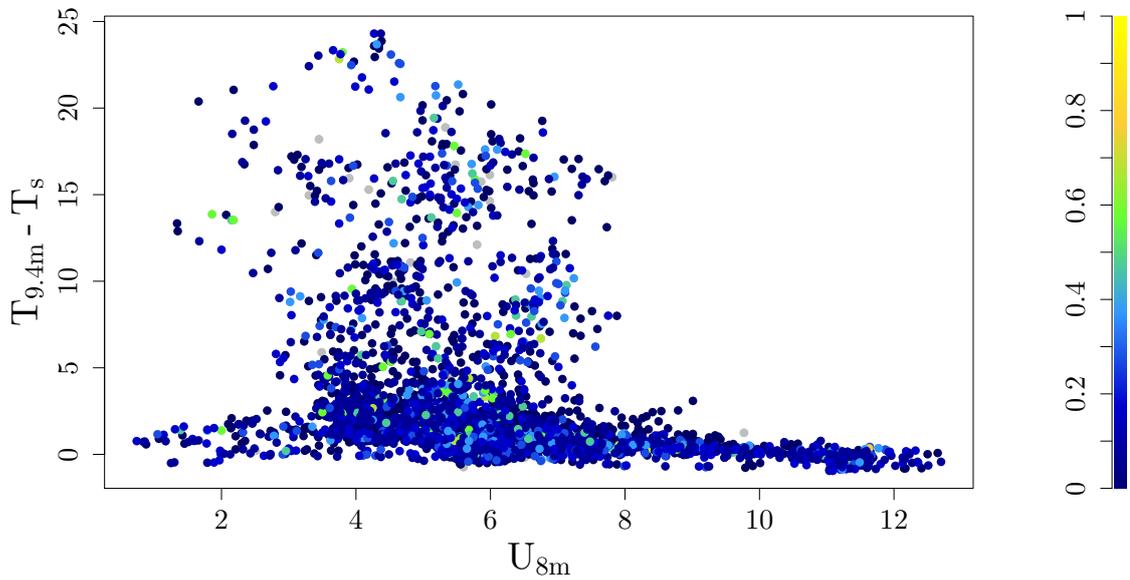


Figure 5.29.: Temperature inversion between 9.4 m and the surface, as a function of wind speed at 8m as observed at Dome C.  $\tau = 10$  (K-Means)

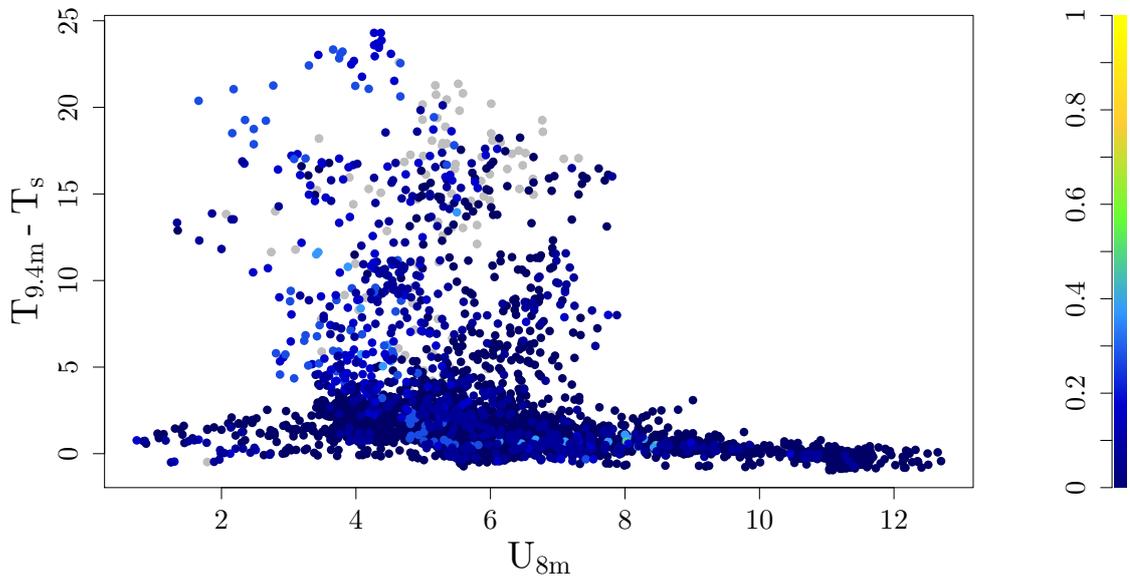


Figure 5.30.: Temperature inversion between 9.4 m and the surface, as a function of wind speed at 8m as observed at Dome C.  $\tau = 43$  (AD Test)

## 6. Outlook

In chapter 5 we estimated the amount of points covering the transition from one stable state to another by using the K-Means Clustering algorithm. Even though this algorithm gives a good indication it has its limitations. Therefore, it is desirable to calculate the exact transition time. In order to do this we can calculate the time it takes on average for the diffusion process  $X_t$  to escape from one of the local minima of the potential  $V(x)$ , or more generally, the time it takes on average for a diffusion process to escape from a metastable state. This mean exit time from a metastable state is an example of a mean first passage time (MFPT): We want to calculate how long it takes on average for a diffusion process to reach the boundary of a domain  $D \in \mathbb{R}$  for the first time if it started at  $x \in D$ . When the domain is the basin of attraction of one of the local minima of the potential, the mean first passage time gives us the average time it takes for the diffusing particle to reach the local maximum of the potential and thereby the transition time [Pavliotis (2014)]. To calculate the mean first passage time we need to use a different model than the one used in chapter 5 to simulate the controlled data. In fact, we have to use a stochastic differential equation

$$dx(t) = -V'(x(t))dt + h(x(t))dB(t), x(0) = x.$$

Further, let  $D$  be a bounded set in  $\mathbb{R}$  with a smooth boundary and  $V$  a confining double well potential. We define the first exit time as

$$\tau_D(x) := \inf\{t \geq 0 | x(t) \notin D\}.$$

This is an example of a stopping time: the information that we have about the stochastic process up to time  $t$  is sufficient to determine whether the event  $\tau \leq t$  has occurred. The average of this time is called the mean exit time

$$\tau(x) := E(\tau_D(x)) = E(\inf\{t \geq 0 | x(t) \notin D\} | x(0) = x).$$

We define  $D := [a, b]$  with  $a < x_{e_1}$  and  $b = x_{e_2}$  ( $x_{e_1}$  first stable equilibrium,  $x_{e_2}$  unstable equilibrium). The assumption is that the left boundary of  $D$  is reflecting and the right one absorbing, i.e.  $\tau'(a) = 0$  and  $\tau(b) = 0$ . The MFPT  $\tau(x)$  can be written as

$$\tau(x) = -2 \int_a^x \int_a^y \frac{\exp[\psi(z) - \psi(y)]}{h(z)^2} dz dy + c_1 \int_a^x \exp[-\psi(y)] dy + c_0$$

[Pavliotis (2014), Krumscheid et al. (2015)] where  $\psi(x) := 2 \int_a^x h^{-2}(z)(-V'(z))dz$  and  $c_0$  and  $c_1$  are constants given by the boundary conditions. With the before mentioned

boundary conditions we obtain

$$\begin{aligned} c_1 &= 0 \text{ and} \\ c_0 &= 2 \int_a^b \exp[-\psi(y)] \int_a^y \frac{\exp[\psi(z)]}{h(z)^2} dz dy \\ \Rightarrow \quad \tau(x) &= 2 \int_x^b \exp[-\psi(y)] \int_a^y \frac{\exp[\psi(z)]}{h(z)^2} dz dy. \end{aligned}$$

Due to time constraints we only briefly discuss this approach. The aim, is to compare the results with time windows for the stability indicator chosen by the K-Means clustering algorithm, Anderson-Darling Normality test and the mean first passage time. As a potential we use the one from the simplified model by [van de Wiel et al. \(2017\)](#), i.e.

$$V(x) = \begin{cases} \frac{1}{2}x^2(\lambda + C) - \frac{C}{3}x^3 - Q_i x & \text{for } x \leq 1 \\ \frac{1}{2}\lambda x^2 - Q_i x + \frac{1}{6}C & \text{for } x > 1 \end{cases} \quad (6.1)$$

This leads to

$$V'(x) = \begin{cases} x(\lambda + C) - Cx^2 - Q_i & \text{for } x \leq 1 \\ \lambda x - Q_i & \text{for } x > 1 \end{cases} \quad (6.2)$$

As we are interested to estimate the time it takes the leave the domain  $D := [a, x_{e_2}]$  where  $a < x_{e_1}$  we are only interested in the case when  $x \leq 1$  (Note:  $x_{e_2} \leq 1$ ). Therefore,

$$\begin{aligned} dx &= (-x(\lambda + C) + Cx^2 + Q_i)dt + \sigma^2 dB \\ x(0) &= x. \end{aligned} \quad (6.3)$$

As parameters we use,

$$\begin{aligned} \lambda &= 2, \\ Q_i &= 2.5, \\ C &= 5.7 \text{ and } 7, \\ \sigma^2 &= 0.6^2, \\ n &= 2000 \text{ and} \\ h &= 0.01. \end{aligned}$$

As a starting point  $x$  we choose  $x_{e_1}$  and as the right boundary  $a$  we choose -2. For  $C = 5.7$   $C$  is close to the bifurcation point and for the other  $C$  it is further away. In fact, this has huge impact on the estimated MFPT. When the system is close to the bifurcation point the estimated MFPT is very small otherwise it is similar to the estimate given by the K-Means algorithm. This is not surprising as  $x_{e_1}$  and  $x_{e_2}$  are very close together close to the bifurcation point while further away from the bifurcation point they differ much stronger. Hence, in the mean it takes less time leaving  $[a, x_{e_2}]$  when the starting

$C$	$\tau_{AD}$	$\tau_{KMeans}$	$\tau_{MFPT}$
5.7	52	69	10
7.0	62	74	52

Table 6.1.:  $\tau$  given by the Anderson-Darling Normality Test, the K-Means clustering algorithm and the mean first passage time.

point  $x_{e_1}$  is close to the right boundary of the domain. The estimated window length by the Anderson-Darling Normality Test, the K-Means clustering algorithm and the mean first passage time are given in table 6.1. Note that all  $\tau$ 's are rounded as we need integers for the window length and we use, like in the previous chapter,  $\tau = \tau_{KMeans} - 5$ .

Figure 6.1 - 6.3 show the results for  $\Upsilon$  with  $\tau = \tau_{AD}, \tau_{KMeans}$  and  $\tau_{MFPT}$  and  $C = 5.7$  and figure 6.4 - 6.6 for  $C = 7$ . We see that for the  $C$  close to the bifurcation point the series fluctuates strongly. This could be a reason why  $\Upsilon$  does not pick up the transitions very well. We clearly see that when  $\tau$  equals the MFPT  $\Upsilon$  does not give reliable results. This is due to the fact that in this case  $\tau$  is too small to expect a good fit for the ARMA models. But when  $C$  is further away from the bifurcation point the transition is noticed with all three  $\tau$ 's and  $\tau = \tau_{MFPT}$  gives the best results. This indicates that using the MFPT to choose  $\tau$  is an approach worth pursuing in the future. But, we need to point out that this approach is only suitable when the underlying theoretical model is known.

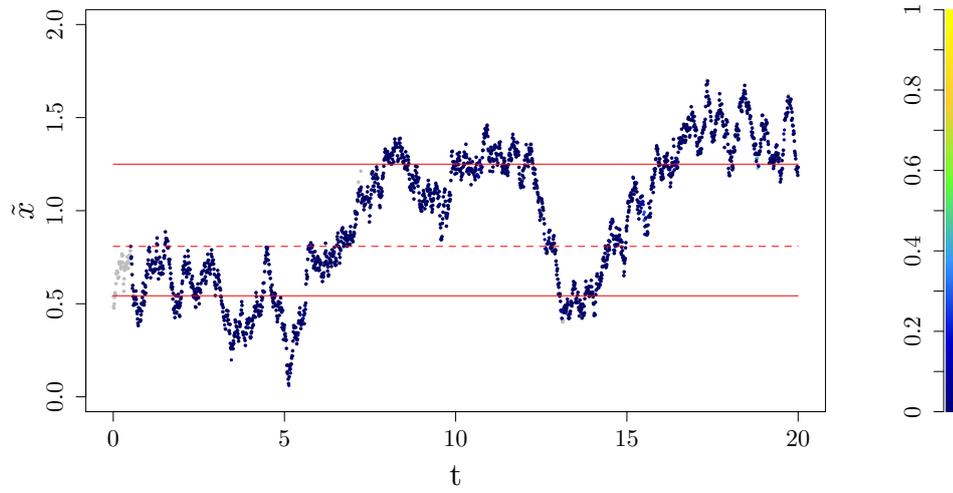


Figure 6.1.:  $\Upsilon$  for  $C = 5.7$  and  $\tau = 52$  (AD).

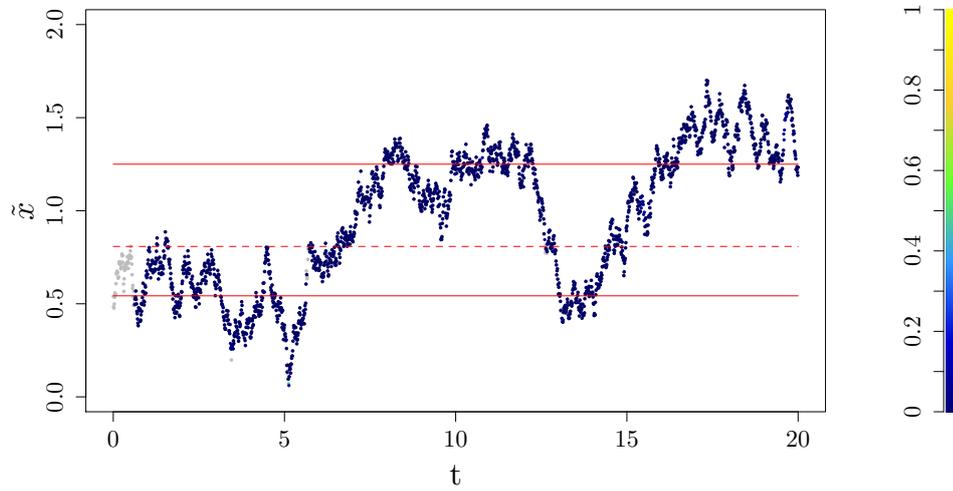


Figure 6.2.:  $\Upsilon$  for  $C = 5.7$  and  $\tau = 64$  (K-Means).

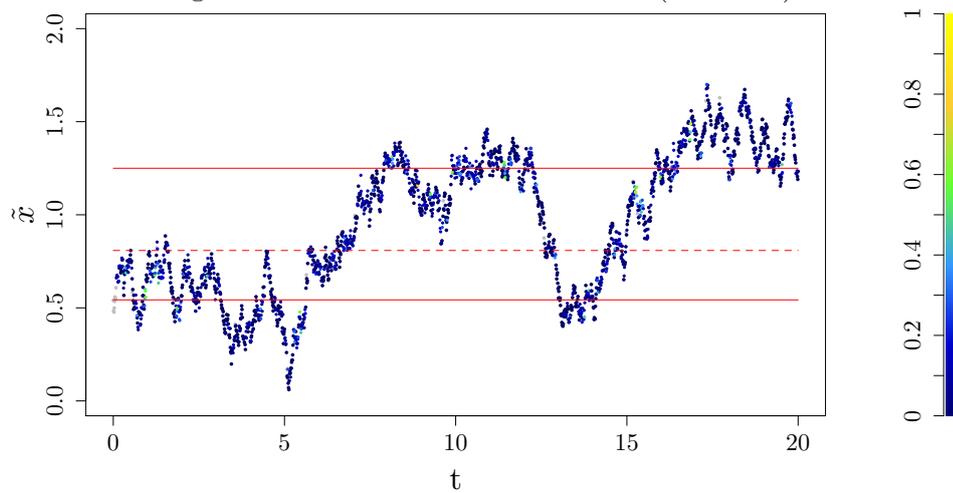


Figure 6.3.:  $\Upsilon$  for  $C = 5.7$  and  $\tau = 10$  (MFPT).

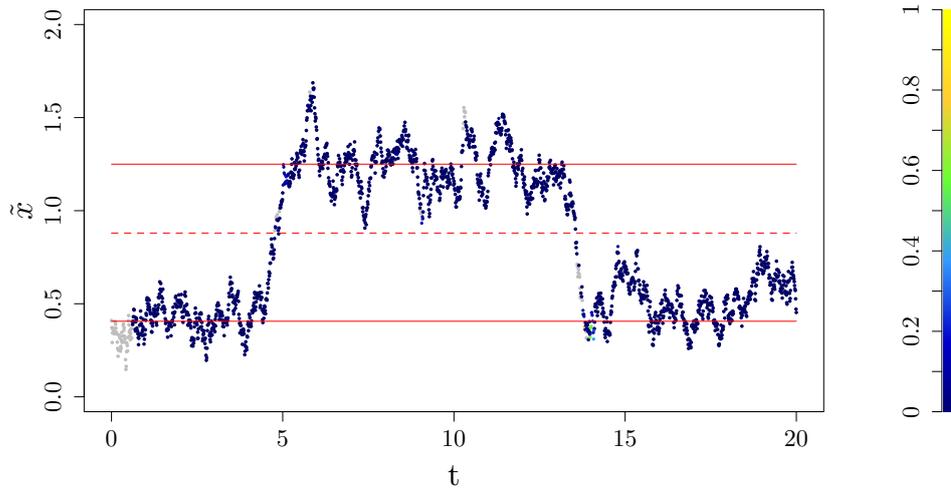


Figure 6.4.:  $\Upsilon$  for  $C = 7$  and  $\tau = 62$  (AD).

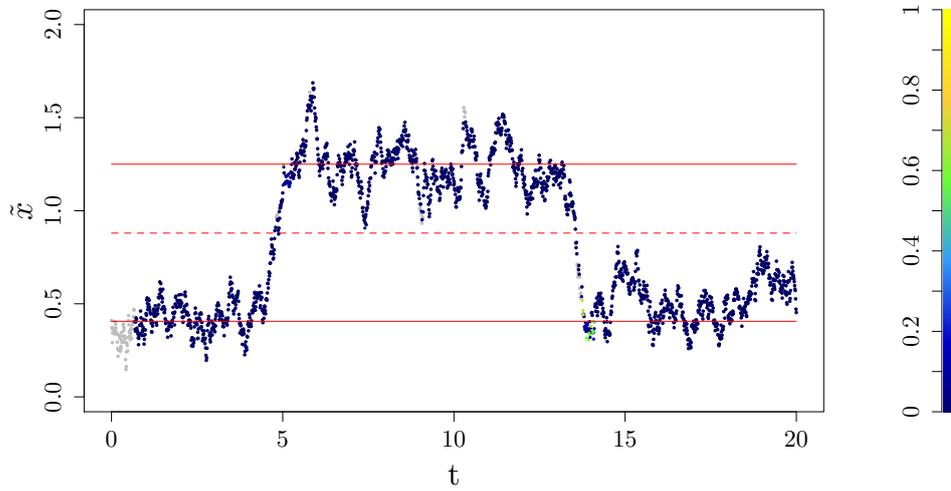


Figure 6.5.:  $\Upsilon$  for  $C = 7$  and  $\tau = 69$  (K-Means).

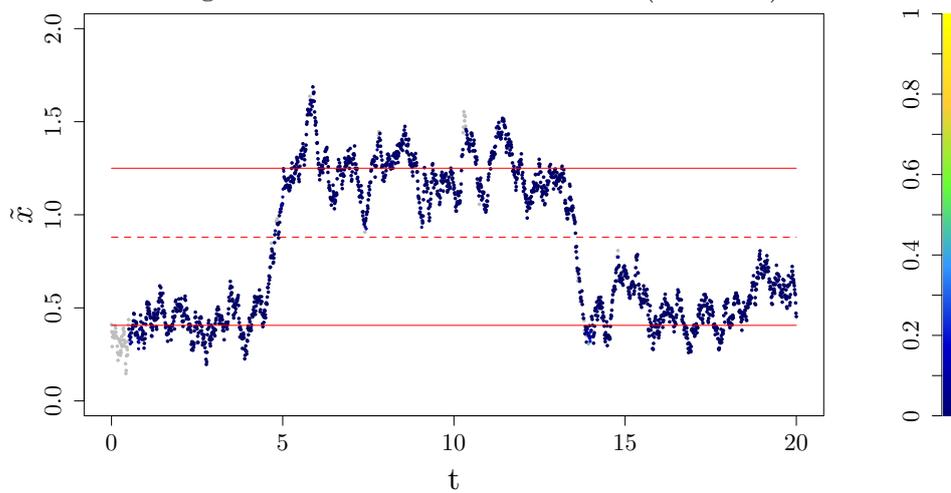


Figure 6.6.:  $\Upsilon$  for  $C = 7$  and  $\tau = 52$  (MFPT).



## 7. Conclusion

We have analysed an indicator  $\Upsilon$  that is able to quantify the dynamical stability of a system observed through a time series of some dynamical quantity. Moreover, we studied a conceptual model developed to understand regime transitions in near-surface temperature inversions in the nocturnal and polar atmospheric boundary layer. The stability indicator was first applied to a time series of controlled data produced by the model for near-surface temperature inversions and then to two data sets of observational meteorological data. The results for the controlled data set showed that  $\Upsilon$  is capable of determining the location of metastable states. This also holds for the nocturnal data. Additionally, we looked at a block averaged time series of one of the nocturnal data sets. From this, we concluded that the measurements frequency is of high importance for  $\Upsilon$ . This finding was also supported by the results of  $\Upsilon$  for the polar data set.

The main advantages of using  $\Upsilon$  is the local nature of the stability indicator and the low requirements in terms of data [Nevo et al. (2017)]. By local we mean that  $\Upsilon$  is applied to subsequences of the original time series. We chose the length of the subsequences with a clustering algorithm [Hartigan & Wong (1979)] and a normality test [Anderson & Darling (1952)]. When both approaches gave similar results  $\Upsilon$  correctly detected the change in the stability of the system. If not we could conclude that the data frequency was too low.

We also suggested using the mean first passage time [Pavliotis (2014), Krumscheid et al. (2015)] to determine the length of the subsequences when the underlying theoretical model is known. Due to time constraints this approach was only briefly discussed.

In further research we plan on having a closer look at the data frequency dependence of  $\Upsilon$  and the mean first passage time to choose the length of the subsequences. It could also be insightful to apply a more sophisticated clustering algorithm than K-Means. Moreover, it would also be interesting to consider a modified version of the stability indicator which allows the usage of different window lengths for subsequences of the same time series. This would potentially improve the recognition of transitions of different lengths.



## 8. Acknowledgement

I thank Étienne Vignon (LTE, EPFL, Lausanne, Switzerland) and Christophe Genton (Laboratoire de Météorologie Dynamique Université Paris Sorbonne / Pierre et Marie Curie) for providing the Dome C data and Danijel Belušić (SMHI, Norrköping, Sweden) for the Dumosa data. Furthermore, I thank Nikki Vercauteren, Wilhelm Stannat, Davide Faranda and Sebastian Krumscheid for helpful discussions and valuable input. I am grateful to two anonymous reviewers for their comments that helped to improve the thesis substantially.





## A. BIC vs AIC vs AICc

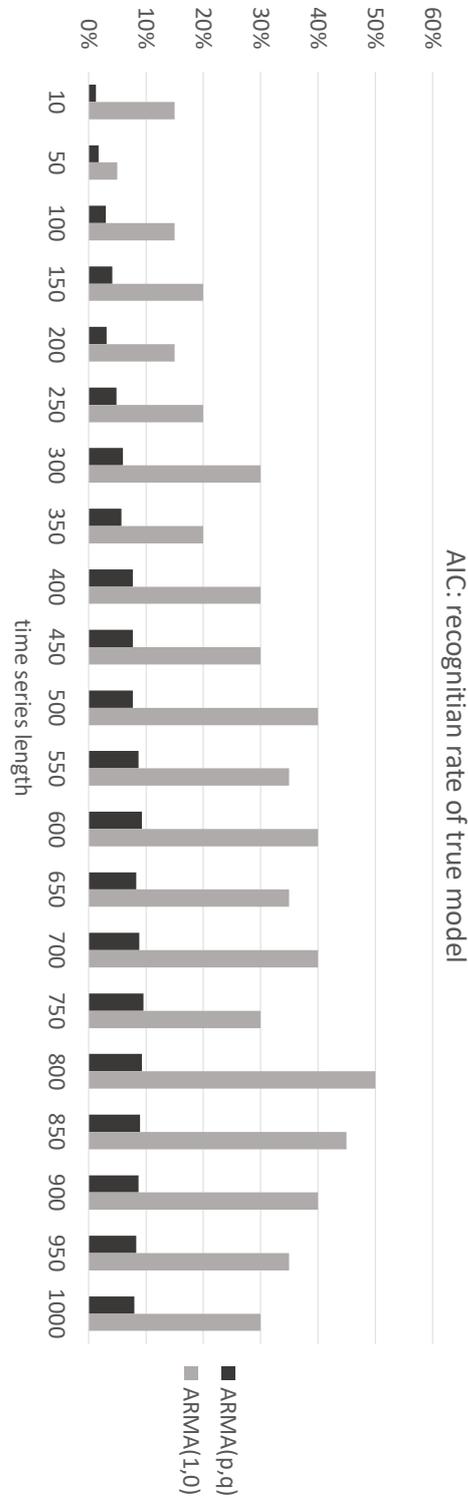


Figure A.1.: percentage of correctly estimated models by the AIC

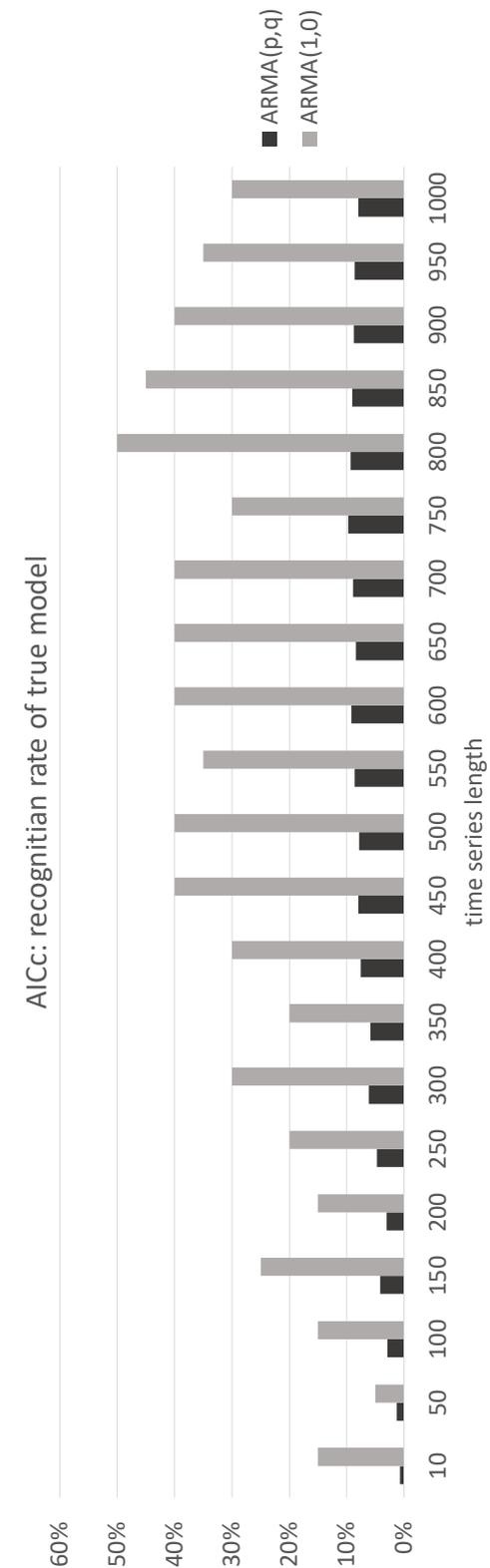


Figure A.2.: percentage of correctly estimated models by the AICc

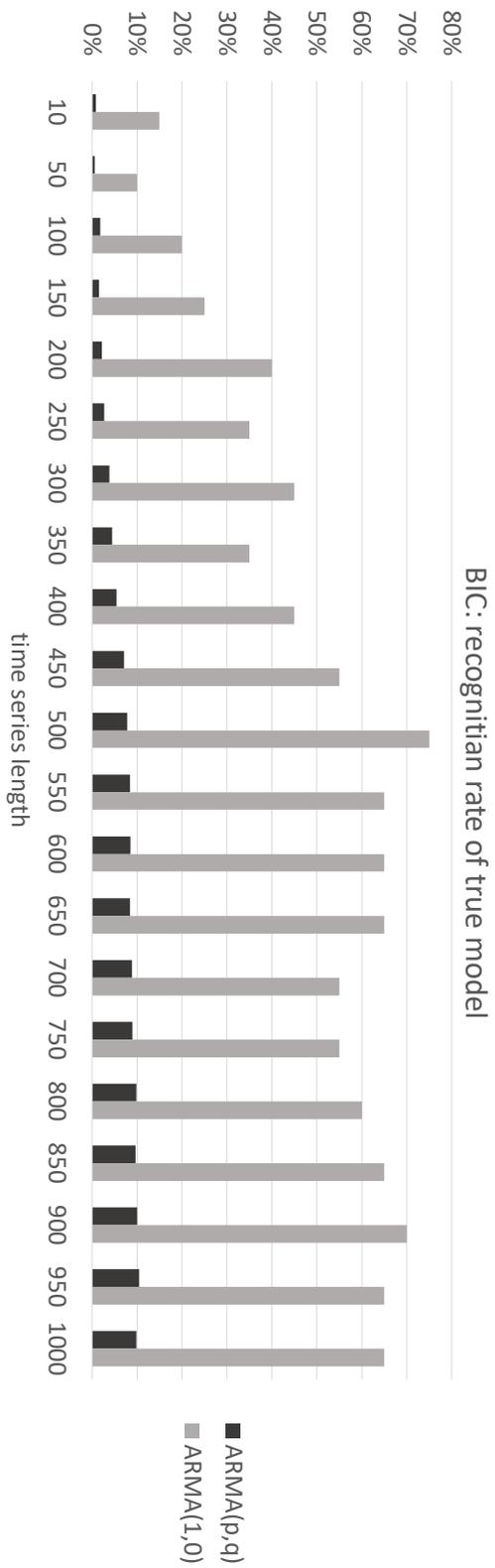


Figure A.3.: percentage of correctly estimated models by the BIC

## B. Properties of Equation 3.1

### B.1. Continuity

We want to verify if

$$\begin{aligned}g(x) &= \frac{dx(t)}{dt} \\ &= \begin{cases} Q_i - \lambda x(t) - Cx(t)(1 - x(t)) & \text{for } x(t) \leq 1 \\ Q_i - \lambda x(t) & \text{for } x(t) > 1 \end{cases} \\ x(t_0) &= x_0\end{aligned}$$

is continuous for all  $x \in \mathbb{R}$ .

First of all, for  $x < 1$  and  $x > 1$   $g(x)$  is a sum of continuous functions and therefore continuous itself. But for  $x = 1$  we have to check whether the left-handed and right-handed limes exist and coincide.

$$\lim_{x \rightarrow 1^-} Q_i - \lambda x - Cx(1 - x) = Q_i - \lambda$$

and

$$\lim_{x \rightarrow 1^+} Q_i - \lambda x = Q_i - \lambda$$

Thus,  $g(x)$  is continuous for all  $x \in \mathbb{R}$ .

### B.2. Differentiability

Now we want to verify if  $g(x)$  is also differentiable for all  $x \in \mathbb{R}$ .

**case 1:**  $x < 1$

$$\begin{aligned}\lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h} &= \lim_{h \rightarrow 0} \frac{1}{h} (Q_i - \lambda(x+h) - C(x+h)(1-x-h) - Q_i + \lambda x + Cx(1-x)) \\ &= \lim_{h \rightarrow 0} C(2x - 1 + h) - \lambda \\ &= C(2x - 1) - \lambda\end{aligned}$$

**case 2:**  $x > 1$

$$\begin{aligned}\lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h} &= \lim_{h \rightarrow 0} \frac{1}{h} (Q_i - \lambda(x+h) - Q_i + \lambda x) \\ &= -\lambda\end{aligned}$$

**case 3:**  $x = 1$

For  $g(x)$  to be differentiable in  $x = 1$  the limit from above and below have to be equal.

$$\lim_{h \rightarrow 0^-} \frac{g(1+h) - g(1)}{h} = C - \lambda$$
$$\lim_{h \rightarrow 0^+} \frac{g(1+h) - g(1)}{h} = -\lambda$$

For  $C \neq 0$  the limits are not the same. Hence,  $g(x)$  is not differentiable in  $x = 1$  for  $C \neq 0$  (but everywhere else). Moreover  $\frac{dg}{dx}$  is not continuous in  $x = 1$ .

## C. Proof: Linear Process with Gaussian White Noise is Stationary and Gaussian Itself

In chapter 4 we assume that the subsequences are realizations of a linear process with Gaussian white noise. In this chapter we justify the choice of these assumptions and analyse their implications.

**Proposition C.0.1.** *A linear process  $\{x_t\}$  is stationary.*

Proof: As  $\{x_t\}$  is linear it can be written as

$$x_t = \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}$$

with

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$

For stationarity we have to show that

- (i)  $E[x_t]$  is constant,
- (ii)  $\text{var}(x_t) < \infty$  and
- (iii)  $\text{cov}(x_s, x_t)$  only depends on  $s$  and  $t$  through their difference.

Obviously, (i) follows by the definition of a white noise process. Moreover,

$$\begin{aligned} \text{var}(x_t) &= \text{var} \left( \sum_{j=-\infty}^{\infty} \psi_j w_{t-j} \right) \\ &= \sum_{j=-\infty}^{\infty} \psi_j^2 \text{var}(w_{t-j}) \\ &= \sigma_w^2 \sum_{j=-\infty}^{\infty} \psi_j^2 \\ &< \infty \end{aligned}$$

and

$$\begin{aligned}
 \text{cov}(x_t, x_{t+h}) &= \text{cov} \left( \sum_{j=-\infty}^{\infty} \psi_j w_{t+h-j}, \sum_{k=-\infty}^{\infty} \psi_k w_{t-k} \right) \\
 &= E \left[ \sum_{j=-\infty}^{\infty} \psi_j w_{t+h-j} \cdot \sum_{k=-\infty}^{\infty} \psi_k w_{t-k} \right] \\
 &= E \left[ \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+h} (w_{t+h-j})^2 \right] \\
 &= \sigma_w^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+h}
 \end{aligned}$$

□

**Theorem C.0.2.** *A linear process  $\{x_t\}$  with Gaussian white noise  $\{w_t\}$  is Gaussian itself.*

Proof: First we note that  $\{w_t\}$  is independent of itself at different times as it is uncorrelated per definition and Gaussian per assumption.

We fix  $t \in \mathbb{R}$  and then define

$$x_t^n := \sum_{j=0}^{\infty} \psi_j w_{t-j}$$

for some  $n \in \mathbb{N}$ . Per assumption

$$x_t = \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}$$

with

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$

Let  $m < n$ ,  $m$  and  $n \in \mathbb{N}$  and  $L^2(Pr)$  be the space of all twice differentiable functions then

$$\left\| \sum_{j=0}^n \psi_j w_{t-j} - \sum_{j=0}^m \psi_j w_{t-j} \right\|_{L^2}^2 = \left\| \sum_{j=m+1}^n \psi_j w_{t-j} \right\|_{L^2}^2$$

Using the Bienyamé formula and noting that  $\|w_t\|_{L^2}^2 = \text{var}(w_t)$  we get

$$\begin{aligned}
 \left\| \sum_{j=0}^n \psi_j w_{t-j} - \sum_{j=0}^m \psi_j w_{t-j} \right\|_{L^2}^2 &= \sum_{j=m+1}^n (\psi_j)^2 \sigma_w^2 \\
 &\rightarrow 0 \text{ for } n, m \rightarrow \infty.
 \end{aligned}$$

Thus,  $\{\sum_{j=0}^n \psi_j w_{t-j}\}_{n \in \mathbb{N}}$  is a Cauchy sequence in  $L^2(P_r)$ . Consequently,

$$\lim_{n \rightarrow \infty} \sum_{j=0}^n \psi_j w_{t-j} = \sum_{j=0}^{\infty} \psi_j w_{t-j}.$$

We can use the same argumentation to show that  $x_t^{-m} := \sum_{j=-m}^{-1} \psi_j w_{t-j}$  converges pointwise to  $\sum_{j=-\infty}^{-1} \psi_j w_{t-j}$ . Therefore,

$$\begin{aligned} \lim_{m,n \rightarrow \infty} (x_t^n + x_t^{-m}) &= \lim_{m,n \rightarrow \infty} \left( \sum_{j=0}^n \psi_j w_{t-j} + \sum_{j=-m}^{-1} \psi_j w_{t-j} \right) \\ &= \lim_{m,n \rightarrow \infty} \sum_{j=-m}^n \psi_j w_{t-j} \\ &= \sum_{j=-\infty}^{\infty} \psi_j w_{t-j} \\ &= x_t \end{aligned}$$

exists.

We know  $x_t^{-m,n} := \sum_{j=-m}^n \psi_j w_{t-j}$  is Gaussian because  $\{w_t\}$  is Gaussian. The characteristic function of  $x_t^{-m,n}$  is

$$\varphi_{x_t^{-m,n}} = \exp\left[-\frac{1}{2} \sigma^2 t^2\right]$$

where  $\sigma^2 = \text{var}(x_t^{-m,n}) = \sum_{j=-m}^n (\psi_j)^2 \sigma_w^2$ . From this it follows that

$$\begin{aligned} \lim_{m,n \rightarrow \infty} \varphi_{x_t^{-m,n}} &= \exp\left[-\frac{1}{2} t^2 \lim_{m,n \rightarrow \infty} \sum_{j=-m}^n (\psi_j)^2 \sigma_w^2\right] \\ &= \exp\left[-\frac{1}{2} t^2 \sum_{j=-\infty}^{\infty} (\psi_j)^2 \sigma_w^2\right] \end{aligned}$$

which is the characteristic function of  $x_t$ . By Levy's continuity theorem it follows that  $x_t^{-m,n}$  converges in distribution to  $x_t$  and therefore is  $\{x_t\}$  a Gaussian process.  $\square$



# List of Figures

1.1.	diurnal evolution of the BL [on the basis of graphic by Stull (2009)] . . . .	1
3.1.	Temperature inversion between 10m and the surface, as a function of wind speed as observed at Dome C, Antarctica, for the period 2014/2015. Each points corresponds to a 30 min averaged measurement. [van de Wiel et al. (2017)] . . . . .	15
3.2.	Plot of $x(t)$ . . . . .	20
3.3.	comparison of the analytically and numerically derived solution (top) and their relative errors (bottom) ( $Q_i, \lambda, C$ from example 2 in the paper by van de Wiel et al. (2017)). From left to right: $h = 1$ and $h = 0.01$ . . . . .	22
3.4.	plots of $\frac{dx}{dt}$ . . . . .	24
3.5.	bifurcation plot of 3.12 with $\lambda$ and $Q_i$ according to i) - iv) . . . . .	30
4.1.	example of a potential $V(x)$ (4.1) and the linearisation $\tilde{V}(x)$ (4.2) . . . .	34
4.2.	. . . . .	38
5.1.	Clustered time series . . . . .	45
5.2.	$\Upsilon$ for simplified model with $C = 6.4$ and added noise ( $\tau = 84$ ). . . . .	45
5.3.	Bifurcation Diagram with $\tau = \min\{mean(T_i) i = 1, 2, 3\} - 5$ . . . . .	46
5.4.	Boxplot of the p values from the Anderson Darling Test for $C = 6.4$ . . . .	47
5.5.	$\Upsilon$ for simplified model with $C = 6.4$ and added noise ( $\tau = 67$ ). . . . .	48
5.6.	Bifurcation Diagram with $\tau$ chosen by Anderson Darling Test (see table 5.2). . . . .	48
5.7.	Bifurcation Diagram with $\tau_{AD} < \tau_{KMeans} - 5$ (see table 5.3). . . . .	49
5.8.	Dumosa measurement set up [provided by Danijel Belušić]. . . . .	50
5.9.	Left: Boxplot of the p values from the Anderson Darling test. Right: Clustered data with K-Means. . . . .	51
5.10.	Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in May 2013. $\tau = 22$ (K-Means) . .	52
5.11.	Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in May 2013. $\tau = 27$ (AD Test) . .	52
5.12.	Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in May 2013. $\tau = 22$ (K-Means) . . . . .	53
5.13.	Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in May 2013. $\tau = 27$ (AD Test) . . . . .	53

5.14. Left: Boxplot of the p values from the Anderson Darling test. Right: Clustered data with K-Means. . . . .	54
5.15. Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in April 2013. $\tau = 23$ (K-Means) . . . . .	55
5.16. Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in April 2013. $\tau = 30$ (AD Test) . . . . .	55
5.17. Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in April 2013. $\tau = 23$ (K-Means) . . . . .	56
5.18. Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in April 2013. $\tau = 30$ (AD Test) . . . . .	56
5.19. Left: Boxplot of the p values from the Anderson Darling test. Right: Clustered data with K-Means. . . . .	57
5.20. Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in May 2013. $\tau = 6$ (K-Means) . . . . .	59
5.21. Temperature inversion between 6 and 3m, as a function of time as observed at Dumosa for the period of one night in May 2013. $\tau = 37$ (AD Test) . . . . .	59
5.22. Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in May 2013. $\tau = 6$ (K-Means) . . . . .	60
5.23. Temperature inversion between 6 and 3m, as a function of wind speed at 6m as observed at Dumosa for the period of one night in May 2013. $\tau = 37$ (AD Test) . . . . .	60
5.24. Temperature inversion between 9.4 m and the surface, as a function of wind speed as observed at Dome C in polar winter 2017. Left: All data points. Right: Longest consecutive time series with $R^+ < 80Wm^{-2}$ . . . . .	61
5.25. Boxplot of the p values from the Anderson Darling test. . . . .	62
5.26. Clustered data with K-Means. . . . .	62
5.27. Temperature inversion between 9.4 m and the surface, as a function of time as observed at Dome C. $\tau = 10$ (K-Means) . . . . .	63
5.28. Temperature inversion between 9.4 m and the surface, as a function of time as observed at Dome C. $\tau = 43$ (AD Test) . . . . .	63
5.29. Temperature inversion between 9.4 m and the surface, as a function of wind speed at 8m as observed at Dome C. $\tau = 10$ (K-Means) . . . . .	64
5.30. Temperature inversion between 9.4 m and the surface, as a function of wind speed at 8m as observed at Dome C. $\tau = 43$ (AD Test) . . . . .	64
6.1. $\Upsilon$ for $C = 5.7$ and $\tau = 52$ (AD). . . . .	68
6.2. $\Upsilon$ for $C = 5.7$ and $\tau = 64$ (K-Means). . . . .	68
6.3. $\Upsilon$ for $C = 5.7$ and $\tau = 10$ (MFPT). . . . .	68
6.4. $\Upsilon$ for $C = 7$ and $\tau = 62$ (AD). . . . .	69
6.5. $\Upsilon$ for $C = 7$ and $\tau = 69$ (K-Means). . . . .	69

6.6. $\Upsilon$ for $C = 7$ and $\tau = 52$ (MFPT) . . . . .	69
A.1. percentage of correctly estimated models by the AIC . . . . .	76
A.2. percentage of correctly estimated models by the AICC . . . . .	77
A.3. percentage of correctly estimated models by the BIC . . . . .	78



## List of Tables

4.1. autoregressive coefficients . . . . .	40
4.2. moving average coefficients . . . . .	40
5.1. $\tau$ given by K-Means algorithm . . . . .	46
5.2. $\tau$ given by Anderson Darling Test . . . . .	47
5.3. Summary of the results from the Anderson Darling Test and K-Means algorithm. . . . .	49
5.4. $\tau$ given by K-Means algorithm and Anderson Darling Test . . . . .	57
6.1. $\tau$ given by the Anderson-Darling Normality Test, the K-Means clustering algorithm and the mean first passage time. . . . .	67



# Bibliography

Anderson, T. W. & Darling, D. A. (1952), ‘Asymptotic theory of certain ”goodness of fit” criteria based on stochastic processes’, *The Annals of Mathematical Statistics* **23**(2), 193–212.

Brockwell, P. J. (1991), *Time Series : Theory and Methods*, 2. edn.

Brockwell, P. J. (2016), *Introduction to Time Series and Forecasting*, 3. edn.

Corless, R. M. (2013), *A Graduate Introduction to Numerical Methods : From the Viewpoint of Backward Error Analysis*.

Dudewicz, E. J. (1976), *Introduction to Statistics and Probability*.

Faranda, D., Dubrulle, B. & Pons, F. M. E. (2014), ‘Statistical early-warning indicators based on autoregressive moving-average models’, *Journal of Physics A: Mathematical and Theoretical* **47**(25), 252001.

**URL:** <https://doi.org/10.1088/1751-8113/47/25/252001>

Hale, J. K. (1991), *Dynamics and Bifurcations*.

Hartigan, J. A. & Wong, M. A. (1979), ‘Algorithm as 136: A k-means clustering algorithm’, *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **28**(1), 100–108.

Hyndman, R., Athanasopoulos, G., Bergmeir, C. & et. al. (2019), ‘Forecasting functions for time series and linear models’. R package version 8.4.

**URL:** <http://pkg.robjhyndman.com/forecast>

Klenke, A. (2008), *Probability Theory: a Comprehensive Course*.

Krumscheid, S., Pradas, M., Pavliotis, G. A. & Kalliadasis, S. (2015), ‘Data-driven coarse graining in action: Modeling and prediction of complex systems’, *Phys. Rev. E* **92**, 042139.

**URL:** <https://link.aps.org/doi/10.1103/PhysRevE.92.042139>

Mahrt, L. (2014), ‘Stably stratified atmospheric boundary layers’, *Annual Review of Fluid Mechanics* **46**(1), 23–45.

**URL:** <https://doi.org/10.1146/annurev-fluid-010313-141354>

- Neath, A. A. & Cavanaugh, J. E. (2012), ‘The bayesian information criterion: background, derivation, and applications’, *Wiley Interdisciplinary Reviews: Computational Statistics* **4**(2), 199–203.  
**URL:** <https://doi.org/10.1002/wics.199>
- Nevo, G., Vercauteren, N., Kaiser, A., Dubrulle, B. & Faranda, D. (2017), ‘Statistical-mechanical approach to study the hydrodynamic stability of the stably stratified atmospheric boundary layer’, *Phys. Rev. Fluids* **2**(8), 084603.  
**URL:** <https://link.aps.org/doi/10.1103/PhysRevFluids.2.084603>
- Olive, D. J. (2014), *Statistical Theory and Inference*.
- Pavliotis, G. A. (2014), *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*.
- Schwarz, G. (1978), ‘Estimating the dimension of a model’, *The Annals of Statistics* **6**(2), 461–464.
- Seydel, R. (1994), *Practical Bifurcation and Stability Analysis : from Equilibrium to Chaos*, 2. edn.
- Shumway, R. H. (2017), *Time Series Analysis and its Applications : with R Examples*, 4. edn.
- Strogatz, S. (1994), *Nonlinear Dynamics and Chaos : with Applications to Physics, Biology, Chemistry and Engineering*, 1. edn.
- Stull, R. B. (2009), *An Introduction to Boundary Layer Meteorology*, reprinted edn, Springer.
- van de Wiel, B. J. H., Vignon, E., Baas, P., van Hooijdonk, I. G. S., van der Linden, S. J. A., van Hooft, J. A., Bosveld, F. C., de Roode, S. R., Moene, A. F. & Genthon, C. (2017), ‘Regime transitions in near-surface temperature inversions: A conceptual model’, *Journal of the Atmospheric Sciences* **74**(4), 1057–1073.  
**URL:** <https://doi.org/10.1175/JAS-D-16-0180.1>