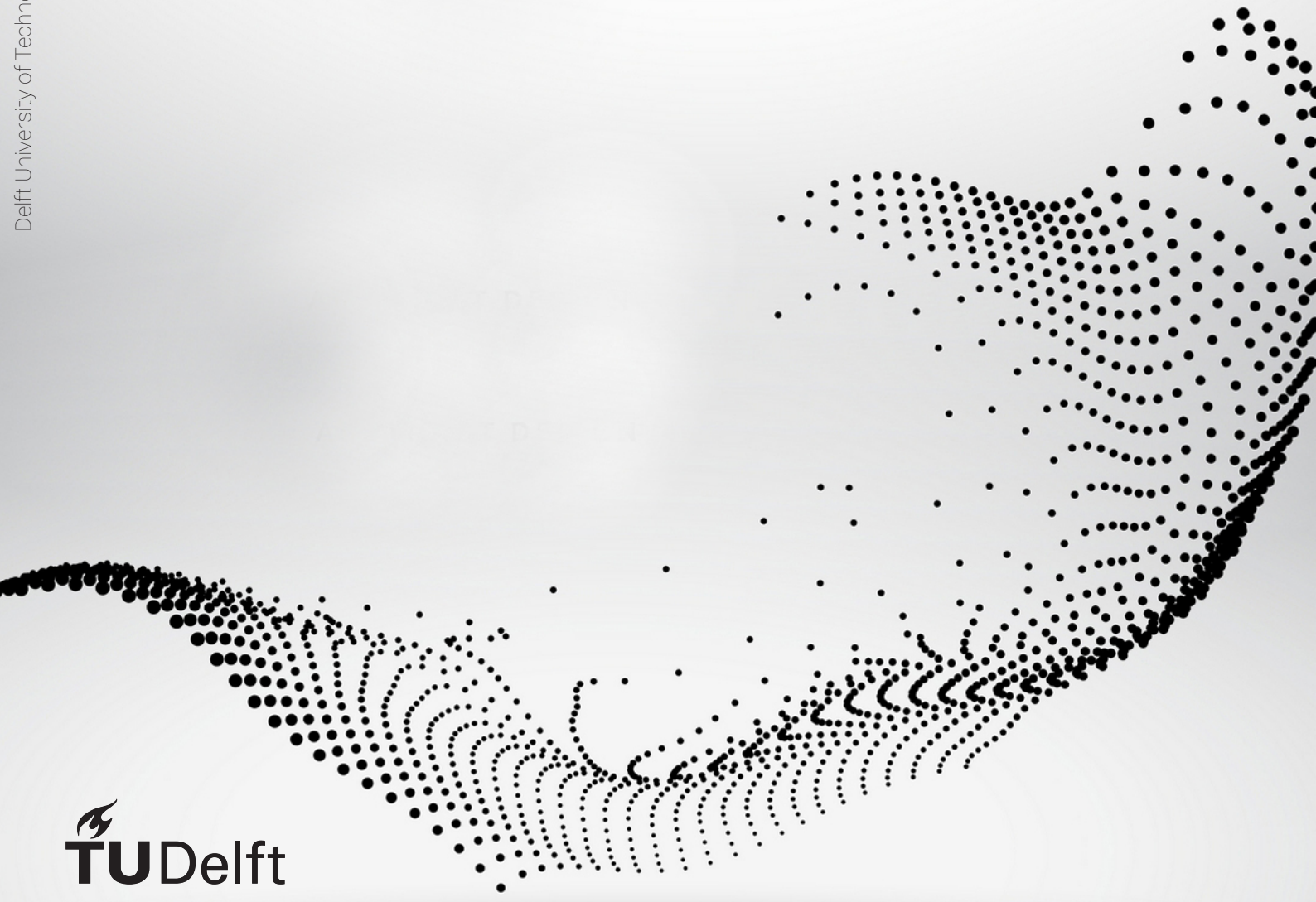
Early Warning Signals for Loss of Control Prediction of a Damaged Quadcopter

Master of Science Thesis

Anthony van der Pluijm

Delft University of Technology



TU Delft

Early Warning Signals for Loss of Control Prediction of a Damaged Quadcopter

Master of Science Thesis

by

Anthony van der Pluijm

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Friday January 24, 2020 at 9:30.

Student number:	4379497	
Project duration:	December 1, 2018 – January 24, 2020	
Thesis committee:	Dr. ir. Q.P. Chu,	TU Delft, chairman
	Dr. ir. C.C. de Visser,	TU Delft, supervisor
	Dr. ir. E. Mooij,	TU Delft, external member

TU Delft

DELFT UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF
CONTROL AND SIMULATION

The undersigned hereby certify that they have read and recommend to the Faculty of Aerospace Engineering for acceptance a thesis entitled **"Early Warning Signals for Loss of Control Prediction of a Damaged Quadcopter"** by **Anthony van der Pluijm** in partial fulfillment of the requirements for the degree of **Master of Science**.

Date : 24 January 2020

Approved:

Dr. ir. Q.P. Chu
Chair of the Department of C&S

Dr. ir. C.C. de Visser
TU Delft, supervisor

External member

Abstract

Loss of Control (LOC), a common form of failure in unmanned aerial systems, has recently gained considerable attention. Indeed, with the projected increase in drone usage, safety has become a critical element. Usually, an on-board Fault-Tolerant Control System (FTCS) attempts to maintain stability and maximize performance by minimizing the effects of external disturbances. Current research focuses on developing elements of such a controller, enabling the quadcopter to overcome specific challenges. However, when a vehicle undergoes an unknown amount of damage that changes its dynamical behavior, few solutions exist. With the recent advances made in data-based methods, exciting options exist that should be considered. Most of the research in this area makes use of machine learning techniques to find potential precursors to failure. These methods often require knowledge about possible faults or about what can be considered as nominal system behavior.

One branch of data-driven techniques is capable of extracting precursors without such prior knowledge. Called Critical Slowing Down (CSD), this method focuses on identifying Early Warning Signals (EWS) to critical transitions in complex systems. The main idea behind this concept lies in measuring a decrease in system resilience as a so-called tipping point is approached. CSD can be observed through rising trends in certain statistical indicators such as variance and autocorrelation. However, CSD has traditionally been applied to ecological and biological systems to, for instance, explain the disappearance of a population. Only recently have the applications been extended to the financial world where stock market crashes are explored.

Expanding on this concept to provide a more general solution for quadcopter safety is not unreasonable. A damaged quadcopter shows similar behavior to other complex systems that undergo a critical transition. As the wind speed is increased, a damaged quadcopter shows increasing oscillations until it eventually enters a LOC situation, which leads to a crash. However, there are some fundamental differences. First, unlike an ecological system, a quadcopter is a closed-loop controlled system where a control input attempts to minimize disturbances. Secondly, the timespans considered in the quadcopter are significantly smaller than the ones in an ecological system that span for years. This thesis work examined whether this method can be applied in such a context and if the resulting EWS are reliable enough to serve as LOC precursors.

To do so, the CSD method was first tested on a less complex system that undergoes a transition. An inverted pendulum on a cart system is equipped with a poorly tuned controller and is subject to a wind force, causing it to tip at a critical value. When the wind speed is slowly increased, EWS are found, but when the rate is increased, fewer signals are detected, and they appear to be less reliable. The system can also tip due to actuators reaching a maximum value and not being able to counter the wind force. In this case, no EWS were found as tipping induced by saturation occurs at wind forces a lot smaller than the critical one associated with a tipping point. Flight data of a quadcopter flying without one rotor was analyzed in a second step. When only observing the quasi-steady state moments, a rise in indicators is noticeable, suggesting the quadcopter is undergoing a form of CSD. However, when transients are also considered, several EWS are found but at unreliable locations. Finally, unlike the inverted pendulum, the quadcopter does not tip over immediately when actuators start to saturate. Instead, the saturation frequency increases over time. Using this aspect, along with certain CSD indicators, results in more reliable EWS.

Concluding, limiting the scope to only CSD indicators to generate EWS is, in retrospect, ironic. More robust controllers are being developed that diminish the effect of disturbances, which is precisely what CSD attempts to quantify when generating an EWS. The increasing frequency of actuators saturating as tipping is approached can be leveraged to generate reliable EWS.

Preface

This report presents the work that was carried out over the last year as part of the masters' thesis assignment at the TU Delft Faculty of Aerospace Engineering. Drone safety and performance is one of the leading research topics of the Control and Simulation department. The main focus of this research is to improve the safety of drones by developing new flight control systems that are capable of preventing loss of control after a failure. This scope includes research on flight envelope prediction, fault detection systems, and prediction of loss of control, to name a few.

This technical report is written from an engineering point of view. When starting this thesis, the author had no prior knowledge regarding the topics of loss of control, critical slowing down, and faults within a quadcopter. These will be explained in the report, but basic knowledge regarding "stability and control" of dynamic systems is required. The research performed in this project is also the first of its kind. Indeed, to the best of the authors' knowledge, this is the first time nature-related contexts are applied to quadcopters to predict failure. A link between these two worlds was challenging to establish as there was no reference point, let alone literature on loss of control prediction with data-based methods.

This work would not have been possible without the help of certain people. Most notably, I would like to thank Coen de Visser for enabling me to perform this project. He had the initial idea of using data-based methods for loss of control prediction, and I am proud to have worked under his supervision for the last year. I would like to thank Sihao Sun, who found the time in his busy schedule to have multiple brainstorming sessions with me and critically discuss some of the results. His enthusiasm and support always encouraged me during the tougher moments of this thesis. I would also like to thank Vasilis Dakos, author of multiple scientific papers that were used in this work, for taking the time to discuss possible research directions at the start of the thesis.

Finally, this project marks the end of my time at the TU Delft. There are several people that I also would like to thank for making this an unforgettable moment. Firstly, I want to thank my girlfriend Marianne, my housemates and good friends Quesada, Jaime, Tino, and Steph. Together, you not only supported me throughout this final project but made sure that every day of this masters' degree was fun, entertaining, and exciting. I also want to thank my friends Pablo, Roberto, and Diego, for the fun times during the year, and Chatham and the IRPdelft team for my amazing experiences abroad in Australia and India. Last but not least, I want to thank my family for the continuous support throughout my degree.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms

AFTCS	Active FTCS
AI	Artificial Intelligence
AR1	Lag-1 Autocorrelation
CSD	Critical Slowing Down
DFA	Detrended Fluctuation Analysis
DFT	Discrete Fourier Transform
EWS	Early Warning Signal
FAA	Federal Aviation Administration
FDD	Fault Detection and Diagnosis
FN	False Negative
FP	False Positive
FTC	Fault Tolerant Controller
FTCS	Fault Tolerant Control Systems
GPS	Global Positioning System
IMU	Inertial Measurement Unit
INDI	Incremental Nonlinear Dynamic Inversion
IQR	InterQuartile Range
KDD	Knowledge Discovery in Databases
KS	Kolmogorov-Smirnov test
LDS	Linear Dynamic System
LE	Lyapunov Exponent
LOC	Loss Of Control
ML	Machine Learning
MSD	Multivariate Shapelet Detection
MTS	Multivariate Time Series
NDI	Nonlinear Dynamic Inversion
NN	Neural Network
ODE	Ordinary Differential Equation
PD	Proportional Derivative [controller]
PFTCS	Passive FTCS
PSD	low-frequency Power Spectrum Density
QLC	Quantitative Loss-of-control Criteria
RF	Random Forest
RL	Reinforcement Learning
RM	Research Module
RWS	Rolling Window Size
RWS2	Rolling Window Size 2
SD	Standard Deviation
SFE	Safe Flight Envelope
SPRT	Sequential Probability Ratio Test
SRF	Single Rotor Failure
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
UAS	Unmanned Aerial Systems
UAV	Unmanned Aerial Vehicle
VRS	Vortex Ring State
WN	White Noise

Greek symbols

α	Angle of attack [rad]
β	Sideslip angle [rad]
δ_ϕ	Roll control
δ_θ	Pitch control
$\dot{\theta}$	Pitch rate [rad/s]
κ	Recovery rate

λ	Autoregressive coefficient
μ	Average or expectation
ω	Angular rate [rad/s]
$\omega_{1/2}$	Frequency associated with half the power [Hz]
ϕ	Bank angle [rad]
ψ	Yaw angle [rad]
σ	Standard deviation
σ_1	Kernel bandwidth
τ	[Kendall's-τ] used to measure trends
τ_1	Correlation time
θ	Pitch angle [rad]
ϑ	Time-varying bifurcation parameter
$\xi(t)$	Zero-mean Gaussian noise

Operators

$G(.)$	Gaussian kernel
$E[.]$	Expectation
$F(.)$	Fourier transform
$f(.)$	Drift function
$g(.)$	Diffusion function

Roman symbols

A_h	Precursors
c	Constant in the AR1 process
$C1 \rightarrow C10$	Indicator combinations
C_m	Moment coefficient
dp	Distance to a critical point
ds	Distance to the unstable state
E_t	Event of interest
ep	Indicator endpoint
F_s	Sampling frequency [Hz]
F_{wind}	Wind force [N]
I_y	Moment of inertia [kg m^2]
idx	Index
K_{d1}	Derivative gain
K_{p1}	Proportional gain
kt	Kendall τ
L	Minimum allowable time for a useful prediction, Pole length [m]
L_A	Critical value threshold
m_c	Cart mass [kg]
m_p	Pendulum mass [kg]
N	Number of samples
n	Load factor
p	Roll rate [rad/s]
P_b	Critical probability
q	Pitch rate [rad/s]
r	Yaw rate [rad/s]
r_1	Bifurcation parameter
t	Time [s]
T_{bif}	Critical bifurcation time
T_{sat}	Critical saturation time
$\bar{u}$	Normalized control input
$\mathbf{u}$, u_k	Control input
∇V	Potential function
V_y	Velocity projection on y [m/s]
$V_{critical}$	Critical wind speed
V_{LOC}	LOC wind speed
V_E	Equivalent airspeed [m/s]
V_x	Velocity projection on x [m/s]
var	Variance
W	Brownian motion
w_i	Rotational velocity of rotor i [RPM]
$w_{i,err}$	Tracking error of rotor i [RPM]
Y_k	DFT of sequence y_n
x_0	Initial state
x(.), x	State vector
Y	Prediction time horizon
y	Residuals
$\hat{z}$	Estimated target parameter
z	Target parameter

Part I

Thesis Introduction

Chapter 1

Introduction

This thesis aims to set a foundation for the usage of data-based methods in the context of loss of control prediction with a particular focus on quadcopters. There are several reasons why this is interesting. First, quadcopters and drones, in general, have seen their possible domains of application increase exponentially over the last years. By interacting more with humans, but at the same time removing them from the loop, concerns regarding the safety of these systems have been raised. For this purpose, controllers that maintain stability and performance have been developed over the last years. Chapter 3 discusses in more detail the current state of the art and emphasizes why safety is so important.

One of the discoveries in this chapter is that when the quadcopter undergoes an unknown amount of damage, using these controllers becomes a challenge. This challenge becomes even more substantial when the physics of the vehicle change drastically. New solutions must be considered that go beyond the ones that make use of the system's equations of motion. This presents an opportunity for data-driven methods, which do not require such prior knowledge. By taking advantage of the recent technological advances in this field, safety-critical moments could be identified in the time series of a quadcopter. Indeed, if potential precursory patterns to loss of control could be identified, a crash can be avoided. The various data-based techniques considered for this purpose are presented in Chapter 4.

A method commonly used in ecological and biological systems is further investigated in this thesis. A research group led by M. Scheffer has developed a method to detect Early Warning Signals (EWS) in the time series data of a system that is close to a tipping point. This thesis researches whether this method is also applicable to identify EWS in the time series of a quadcopter. The methodology followed is described in Chapter 5.

Overall, this thesis is structured into multiple parts. Part I. contains a detailed break down of the research objective and question. In Part II. the relevant background information required to understand the contents and the context of this thesis is provided. Part III. presents the results of the thesis by means of two scientific papers. Finally, Part IV. contains the answers to the research questions. Conclusions and recommendations are provided.

Chapter 2

Thesis Project

This section will elaborate on the thesis proposal. The research motivation and the question will be provided in Section 2.1 and Section 2.2 respectively.

2.1. Research Motivation

With the usage of autonomous vehicles expected to increase significantly over the next couple of years, it is essential to ensure their safety. LOC has been identified as the most important cause of accidents. Attempting to mitigate LOC directly contributes to the safety of these systems. Instead of developing more robust controllers, it could be interesting to determine whether LOC can be predicted without relying on known system characteristics. If it is possible only to use data, the methodology developed in this thesis could be used on a variety of systems where the dynamics might not be fully known.

2.2. Research Questions

Various methods are discussed in the literature study that are capable of extracting knowledge from time series data in order to predict events. Some of the methods explicitly search for patterns in the data, whereas others generate knowledge implicitly in the model. Some of the methods require labeled data, whereas others use a nominal set. To limit the scope of this thesis, one of the methods will be extensively examined: critical slowing down indicators. As this method has been used to identify precursory signals to the tipping of various ecological systems, it is interesting to apply the methods to a quadcopter. This leads to the following research objective:

The objective of this research project is to determine whether quadcopters suffering from single rotor failure show signs of critical slowing down when approaching a loss of control condition. This objective can be achieved by using the traditional critical slowing down indicators to detect phase transitions in the time series data of the quadcopter.

With this objective in mind, the main research question is defined as:

Do the time series of a quadcopter suffering a single rotor failure contain early warning signals that can be used to predict loss of control?

From the main research question, several sub-questions can be posed:

1. How can CSD theory be used to quantify these EWS?

1.1. Which CSD indicators exist in literature, and which ones should be used?
1.2. Are the traditional indicators adequate for this problem?
1.3. Which variables should be considered?

2. How can LOC be identified in a time series?
   2.1. Can the SFE be used?
   2.2. How can actuator saturation be used?
   2.3. What is the role of bifurcation points in the LOC context?

3. Can resulting EWS be used to predict LOC?
   3.1. What is the influence of transients on the EWS?
   3.2. How robust are the EWS?
   3.3. Can actuator saturation be used as an EWS to LOC?

Part II

Preliminary Study

Chapter 3

Quadcopter safety

Unmanned Aerial Systems (UAS) is the more general term used to refer to drones that have a fixed-wing format or use multiple rotors. The first designs of rotatory UAS were helicopters with two rotors: a horizontal one to generate lift and a vertical one in the back to counter the torque created by the first one. However, the many problems of balance caused by the tail rotor led to the development of a solution that employs four rotors to keep the body in the air while maintaining stability: the quadcopter [1]. The four rotors allow this system to move more freely than a classic aircraft or helicopter at the expense of a larger amount of degrees of freedom. The blades are usually fixed in pitch (unlike a helicopter) and are powered by brushless motors. To move around safely, a quadcopter often uses cameras and sensors which usually include an accelerometer to determine in-flight position and orientation, an Inertial Measurement Unit (IMU) or a magnetometer (used in combination with a GPS) to maintain direction and flight path, and a three-axis gyroscope to provide the angular acceleration around the three axes. Additional sensors can include distance ones or barometers to determine the altitude. Figure 3.1 illustrates a typical quadcopter (The Parrot BEBOP 2).

With the large variety of shapes and sizes, quadcopters are now being used for a wide range of tasks. These include but are not limited to aerial photography, express delivery, information gathering for disaster management, safety inspection, crop monitoring in the agriculture world, military usage, and personal/leisure usage. Projected demand in these fields is expected to increase [2], with, for example, the Federal Aviation Administration (FAA) estimating that 2.4 million hobby drones will be used by 2022 [2]. Interaction between quadcopters and humans will also only increase. This poses new requirements on the safety of these machines and, more specifically, on the actuator reliability of quadcopters. Indeed, quadcopters are designed with the requirement of being as lightweight as possible, meaning they usually do not carry redundant rotors.

One way of increasing the safety of these vehicles is through Fault Tolerant Control Systems (FTCS). As the name suggests, the system can operate with satisfactory performance and stability with the presence of potential faults [5]. These FTCS systems come into variants: passive (PFTCS) and active (AFTCS) forms. In the former, faults that could occur are assumed to be known, and a robust control strategy is already mapped out. In the latter, the AFTCS system actively reconfigures itself to deal with unknown faults. By doing so, the system can remain stable while occasionally doing so at the expense of performance.

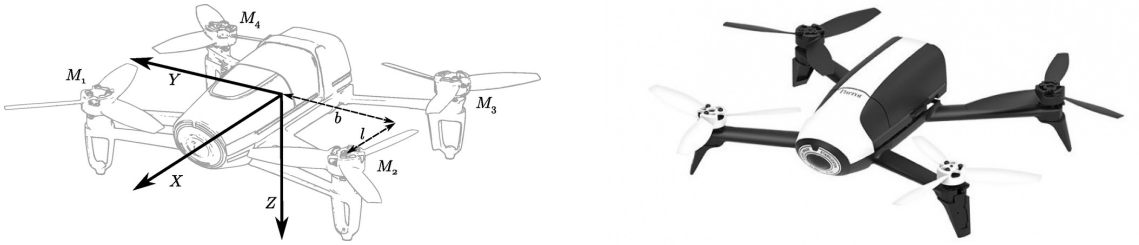


Figure 3.1: A typical quadcopter: the Parrot Bebop 2 [3, 4]

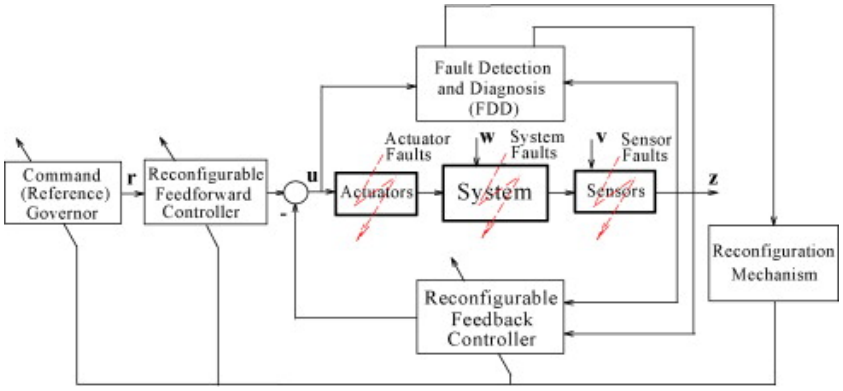


Figure 3.2: General structure of an AFTCS [5]

An AFTCS system is divided into four subsystems: a Fault Tolerant Controller (FTC) (also named reconfigurable controller in literature), a Fault Detection and Diagnosis (FDD) scheme, a controller reconfiguration mechanism, and a command/reference governor [5]. The FDD and the reconfigurable controller are what differentiate an active one from a passive one. The structure of the AFTCS is presented in Figure 3.2. The FDD's goal is to detect and isolate the fault in real-time as quickly as possible. The FTC is designed such that stability, desired dynamics, and steady-state performance are achieved automatically. A reconfigurable feedforward controller is used to ensure that the closed-loop system can track an input trajectory. A command governor is designed to adjust the input trajectory automatically to avoid actuator saturation. Overall, the design objectives of the AFTCS are described by Zhang as follows [5]:

1. to possess an FDD system that can detect faults as soon and precisely as possible;
2. to design a control system that can compensate for the detected fault. This includes the controller parameters but also the structure of the new controller.

3.1. Fault detection and diagnosis

As mentioned above, the purpose of the FDD system is to identify a fault as soon as possible. Techniques to do so can be split up into two categories: model and data-based methods. A classification is presented in Figure 3.3. Model-based methods construct a model of the system from underlying physical principles. Data-driven techniques build a model of the system using historical data from the time series of the system variables.

The most common model-based methods are state estimation ones. They usually compare an estimated state or parameter to its actual value and define it as a fault if very different. In [7, 8], use is made of a Thau observer to detect faults in a quadcopter. Residuals are generated for the fault identification task by estimating the states of the nonlinear system. In [9], a Thau observer was used to detect time-varying faults through an adaptive algorithm. In [10], a combination of a nonlinear Thau observer and a linearized Kalman filter are used to detect actuator faults. A two-stage Kalman filter for actuator FDD of an unmanned quadrotor helicopter was also proposed [11]. The faults are modeled as losses in control effectiveness of one or multiple rotors with varying magnitudes.

However, in the context of fault detection, the need for an accurate model is essential. Doing so for complex systems such as Unmanned Aerial Vehicles (UAV) is increasingly tricky; hence this has opened the door for data-driven methods. With the increased efficiency of onboard processing and sensors, data-driven techniques can be efficient alternatives for model-based ones. In Figure 3.3, data-driven approaches are divided into quantitative and qualitative ones. Qualitative methods can be classified as symbolic Artificial Intelligence (AI) [1]. Quantitative models can be split into knowledge-based and signal-based methods. The former

[1] Symbolic AI reduces the level of abstraction of the problem to a more “symbolic" one making it easier to interpret for a human

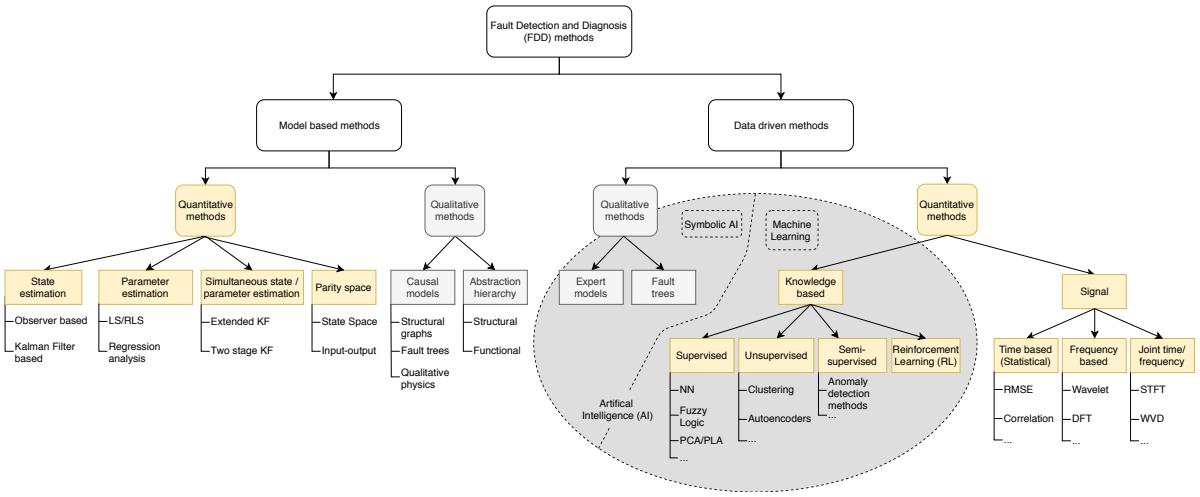


Figure 3.3: Classification of FDD methods (Inspired from [5, 6]). Circled are the methods that are part of the AI domain.

is also known as Machine Learning (ML), and together with symbolic AI, they form the entire AI domain.

Machine learning is traditionally divided into three categories. The first is supervised learning where "the data is first classified and labeled with tags that indicate the system's conditions and symptoms, such as healthy, faulty, and the type of faults. The labels are also known to the algorithm. The objective is to search for patterns and rules representing the information and relationship between data and faults." [6]. Supervised techniques for FDD in quadcopters have been used extensively. For instance, in [12], a fault detection method that uses extreme machine learning is proposed to approximate the dynamics of the actuator of a quadcopter. By monitoring the norm of the output of the network, faults can be detected. In [13], a supervised Support Vector Machine (SVM) uses accelerator and gyro measurements to distinguish between faulty and nominal flight conditions of a UAV. In [14], Zhang makes use of an FDD method composed of a fuzzy system and a Neural Network (NN) to detect sensor faults of the navigation system of a quadcopter. A Kalman filter is used to calculate the residual sequences of sensor faults, and the FDD is used to detect the faults.

Finally, unsupervised techniques do not make use of labels. The algorithm has to determine by itself what the tags are. In order to do so, similarities between pieces of data are discovered in order to group them into clusters. Then, depending on the cluster, a piece of data finds itself in, the algorithm can label it as a fault or not. No direct studies or applications were found in literature where unsupervised techniques are used for FDD in UAVs.

A more in-depth analysis of techniques used to detect faults and extract precursory patterns to failure are presented in Chapter 4. Overall, FDD methods often make use of a physical model of the system that is derived from a physical model or supervised techniques. Indeed to detect a fault, knowledge of what is not a fault is required. This is why few to no applications of unsupervised FDD techniques exist in literature.

3.2. Fault tolerant controller

With the diagnosis of the FDD system, the AFTCS can reconfigure the controller in-flight allowing quadcopters to maneuver with faults being present. Usually, the magnitude and the type of fault are first estimated, after which the necessary control actions are undertaken. Zhang provides in his FTCS review an extensive classification of the techniques available [15]. The most commonly used ones will be presented. Most of the methods that exist are focused on handling small faults or partial actuator failures by modeling them as uncertainties [2] [16, 17]. Examples of FTC controllers include the sliding mode control [18, 19], which focus on reducing the dynamics of the system to a lower order making it more robust against model uncertainties [20], or model reference adaptive control systems [16, 17] which use a reference model that represents the

[2] Uncertainties refer to parameter uncertainties caused by external disturbances or uncertain time delays in processing data

desired behavior of the system in order to adjust accordingly the controller parameters in-flight [21].

Several FTC techniques also deal with the total loss of one rotor. In this case, a rotor is either missing or is completely broken, meaning the quadcopter can only generate lift with the remaining rotors. The first known case where control of a quadcopter was regained after losing one rotor, looked at giving up control about the yaw axis in order to achieve stability [22]. In [23], control about the yaw axis is kept by turning off the rotor across the one that failed. As such, the quadcopter would enter a bi-rotor configuration allowing it to perform a safe landing. In [24], methods such as the linear–quadratic regulator are proposed to allow a quadcopter to hover despite the loss of one or more rotors.

These techniques focus on the actuator loss at low-speed flight. However, failure of a rotor can also occur at high velocities. This is something that has not been studied extensively yet. New research carried out by the TU Delft in 2018, has shown it is possible to continue high-speed cruise flight when a rotor is missing or when an actuator has suffered a complete failure. A cascaded Incremental Nonlinear Dynamic Inversion (INDI) controller was designed, allowing the drone to fly safely up to 9 m/s - or 50% of its nominal maximum velocity [25].

Concluding, a proper design of the FTC and the FDD systems in an AFTCS contributes towards safer and more reliable UAV flight. Through the FDD, faults are detected as soon as possible. This enables the FTC to reconfigure the controller in order to keep the system stable. With models being more complex, a shift in research towards data-driven techniques can be seen. The ability to predict a failure could provide the FTC more time to adjust the controller accordingly. Failure could even be prevented if the controller adopts the adequate actions and if the prediction is made with a large enough time window.

3.3. Loss of Control

A direct consequence of hardware and system failures is Loss of Control (LOC). In small UAS, LOC is responsible for almost 35% of the recorded incidents, more than any of the other reasons for failure (e.g., collisions, loss of navigation...). In this section, the historical background of how LOC was defined and what it exactly is will be provided. Measures beyond the traditional AFTCS will be discussed.

3.3.1. Quantifying LOC

In 2000, the commercial aviation safety team created a joint safety analysis team, whose objective was to analyze aviation safety data. The goal of this team was to reduce the overall fatal accident rate by 80% in the following seven years [26]. They identified LOC as the largest cause of fatal accidents and defined it as “a significant, unintended departure of the aircraft from controlled flight, the operational flight envelope, or usual flight attitudes, including ground events"; with significant referring to an accident or total failure of the aircraft [27, 28]. In order to quantify LOC, measurable characteristics were needed. Five such criteria were defined in a report written by NASA and Boeing [29]. They are listed below:

1. “operating outside of the flight envelope;
2. not predictably altered by pilot control inputs;
3. characterized by strong nonlinear effects such as divergent behavior of the states or a disproportionately large response to small state variable changes;
4. likely to result in large angular rates and displacements;
5. characterized by the inability to maintain heading, altitude, and wings-level flight."

In order to say whether LOC has occurred or not, a set of metrics was needed. These metrics were named collectively as the Quantitative Loss-of-Control Criteria (QLC). They provide a means to quantify LOC. They relate to airplane flight dynamics, aerodynamics, structural integrity, and flight control use. A list of flight

dynamics characteristics that play an important role in LOC was found and was used to develop the QLC. Parameters included: angle of attack, sideslip angle, bank angle, pitch angle (α, β, ϕ, θ); pitch and roll rate (p, q); load factor (n), equivalent airspeed (V_E) and pitch and roll control (δ_θ, δ_ϕ). Five envelopes were defined for the QLC that capture the important relationships between those variables. LOC for fixed-wing aircraft was defined as a maneuver that exceeds three or more of those envelopes [29].

Unlike aircraft, ensuring the safe flight of UAVs has only recently become a topic of interest. NASA has launched a UAS traffic management program to enable low altitude flight of unmanned systems. In order for this to actually take place, a key part of the project focuses on enabling the safety of autonomous vehicles under non-nominal or unexpected conditions. This is, however, a challenging task as there have not been as many reportings of mishaps with quadcopters. Of the ones reported, LOC was identified as the largest cause of accidents [30]. Moreover, there has been little research in developing high fidelity models of these vehicles when operating under non-nominal conditions [31]. Indeed most of the research focuses on making these vehicles safe and reliable in order to precisely avoid possible adverse scenarios. Finally, the models developed for LOC detection/mitigation for aircraft cannot be applied to quadcopters as not only the dynamics (more degrees of freedom) but also the structural composition of the system (fixed-wing vs. rotors) are different.

3.3.2. Preventing LOC

Most of the LOC prevention research has been directed on aircraft. Belcastro is an author that has researched this extensively. In [32], an integrated systems concept for preventing LOC is presented. A holistic “Aircraft Integrated Resilient Safety Assurance and Failsafe Enhancement" system was examined where LOC is prevented by monitoring the vehicle’s health in combination with external hazard conditions. In 2014, technologies for in-flight LOC prediction and prevention were made [33]. The position of the aircraft is consistently measured relatively to the five previously defined envelopes. Another onboard system for LOC prevention is developed in [34], where an autonomous flight envelope estimation is derived using system identification techniques.

Most of these methods rely on defining the flight envelope, which spans the region in which the aircraft can operate safely under nominal conditions. An example of such an envelope is presented in the left illustration of Figure 3.4. When the aircraft does leave the envelope, safety is compromised; hence restrictions are set, preventing it from leaving this space. However, the flight envelope is only defined for stationary and quasi-stationary states meaning that during dynamic maneuvers, there is no reference safe space.

To solve this, attempts started in 2016, where a dynamic flight envelope protection system was developed to ensure that critical parameters remain between predetermined limits [35]. More recently, a database-driven approach to determine the Safe Flight Envelope (SFE) was developed such that an aircraft can recover from abnormal conditions [36]. The SFE is an exciting concept as LOC can be avoided by defining the space of actions for which safe dynamic maneuvers are guaranteed. In 2018, the SFE uncertainty was determined under the influence of turbulence by using probabilistic theories [37].

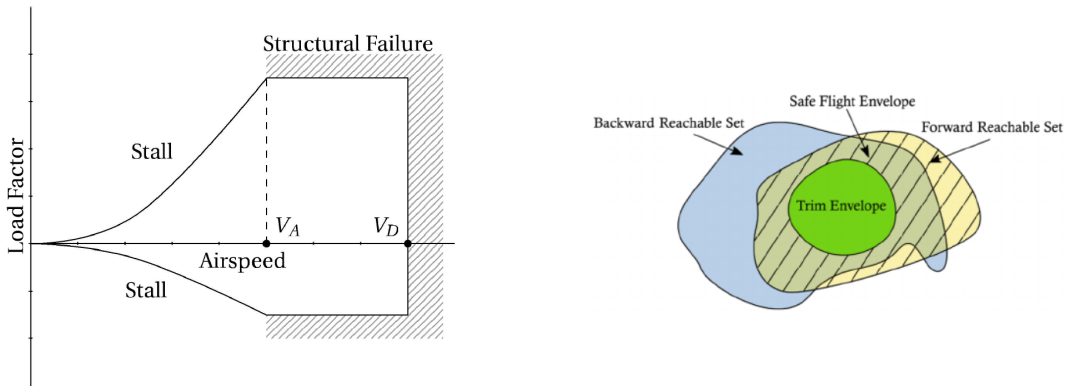


Figure 3.4: Left: Quasi-stationary flight envelope, right: safe flight envelope defined as the intersection between the forward and backward reachable set [38]

The SFE can be defined as: “the part of the state space for which safe operation of the system can be guaranteed and externally posed constraints will not be violated", with externally posed constraints referring to the dynamic, structural, and environmental envelopes of the system [37]. By using the reachable set theory, the SFE can be delimited. This can be done by solving Equation (3.1) and identifying all the states that can be reached over a certain time horizon when starting at an initial condition. Larger time horizons and larger control inputs increase the space of this set.

$$\dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t), \mathbf{u}(t)) \quad \mathbf{x} \in \mathbb{R}^n \tag{3.1}$$

The system in Eq. 3.1 can be solved forward in time. This leads to all the states **x(.)** for which there exists a control input **u**, such that this set can be approached from an initial condition x_0. When solved backward in time, the set of states **x(.)** is found for which a control input **u** exists that the initial condition can be reached. The SFE is defined as the intersection between both sets. For each state in the SFE, there is a path going from an initial condition to a specific state, and back to the initial one, by applying a control sequence **u**, over a certain time horizon.

The SFE theory has been recently applied to quadcopters, but there have been several limitations. Defining this SFE in real-time is quite complicated [18]. Firstly, the computational power needed can be quite significant. As a quadcopter is designed such that each part is optimally used, the required computing power is often not available. Moreover, LOC occurs quite rapidly hence the need for a quick algorithm. One solution that could work is the offline computation of the SFE. As suggested in Zhang's paper [39], by comparing the current vehicle state to the ones registered in a database, the SFE could be determined. Another solution has been proposed in 2019, where a Monte Carlo approach is used to compute the SFE online in high-speed forward flight [40].

Computing the SFE has other challenges. Besides being computationally expensive, the dependence on the time horizon can raise possible question marks about its application. As such, recent attempts have been made to identify the causes of LOC. This was the focus of a MSc thesis in 2018 called “Quantifying Loss of Control of quadcopters" [41], where an attempt to provide a qualitative description was made. Kersbergen performed test flights under nominal and non-nominal conditions, and the resulting failures were linked to LOC. Actuator saturation was determined to be the primary cause of LOC events, whereas other observations such as the Vortex Ring State (VRS) [3] and blade flapping [4] can be seen as likely precursors to LOC [41]. However, no conclusive evidence was found that these parameters can be used to qualify LOC. Instead, they were deemed as likely precursors.

Concluding, LOC is an area of research that has been around for a while in the aircraft world but is still relatively new in the quadcopter world. It is defined as an adverse condition the quadcopter finds itself in, from which it usually cannot recover, therefore often resulting in a crash. It is usually prevented by defining flight envelopes. However, these envelopes have their limitations. This has opened the door to a new research direction into determining possible precursors to LOC of a quadcopter. With visual tests having been performed in Kersbergen's work, the next natural step is to look deeper into the data to define what precisely the quantitative causes of LOC are.

[3] VRS is a condition where air recirculates through the rotor. This is an unstable condition where the produced thrust starts to fluctuate.

[4] Blade flapping occurs in helicopters. The advancing part of the rotor sees a larger effective velocity, hence generates more lift than the retreating side. This causes an asymmetry in the lift distribution. To counter this, helicopters enable the blades to flap - something which quadcopters do not.

Chapter 4

Data-based methods for precursor identification

As previously mentioned, data-based methods should be considered when attempting to predict LOC. Indeed, they can be used to extract useful patterns from the time series without requiring knowledge of the physical model that drives the system. Two common ways this problem is described in literature are present below.

Firstly, “prognosis" is a term commonly used when describing the remaining useful life of industrial components or machinery [42]. It can be defined as “detecting the precursors of failure and predicting how much time remains before a likely failure" [43]. Examples where prognosis has been used, are in [44] where a bearing is monitored or in [45] where a prognostic algorithm is discussed for electro-mechanical actuators. In the setting of the thesis, the system could be monitored with the prognosis algorithm, and when a failure is detected, the time to LOC would be determined.

Event forecasting is another common way in literature to describe such a problem. An event can be defined as “an occurrence over time where significant changes in the time series can be noted" [46]. Event prediction is used in a variety of fields ranging from medicine [47], to natural disaster forecasting (earthquakes, floods...) [48–51] and financial market crashes [52] to name a few. In this thesis, the LOC of a quadcopter can be defined as the event of interest that needs to be forecasted. The event prediction problem can be formulated as “learning a prediction procedure that, given a time series data, correctly predicts whether the target event will occur in the near future" [53]. A minimum allowable time (L) is denoted by which the event should be detected by, in order for the prediction to be useful. A search period can also be posed where the time series data can be used to make the prediction. This is denoted by (Y) in Figure 4.1. The goal is to, as soon as possible, classify the time series as "likely to lead to an event," with a high probability.

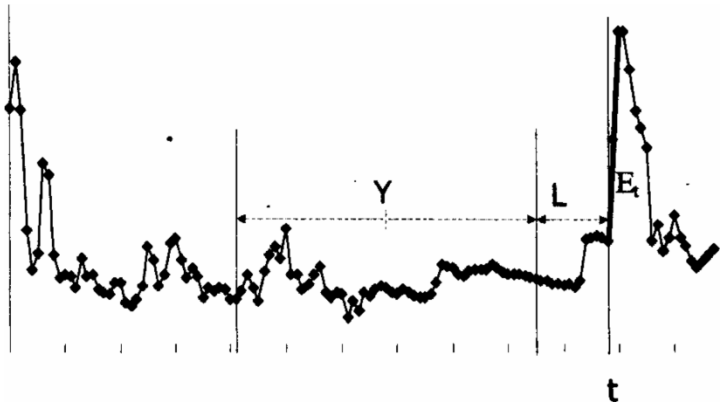


Figure 4.1: Example of an event with its prediction horizon (Y) and minimum time (L) before the failure (E) [53]

In all these ways to look at the problem, the precursors to LOC are somehow extracted. They are either implicit, hidden in the knowledge the model has created for itself, or are explicitly searched. In the setting of

this project, the definition proposed by Dolenko to describe the precursor identification problem will be used [54]. The concept is described below and is illustrated in Figure 4.2.

- A precursor can be defined as “a combination of the input features that are strongly connected and/or correlated to the occurrence of the event".
- The precursor identification problem can be formulated as “finding the most probable set of input variables over a search interval that causes the occurrence of the event and to find the delay between the initiation and occurrence of the event."

A comment must now be made: in literature, the word *precursors* (to an event) is not a commonly used term. Some more recurring terms include anomaly, fault, change point, tipping point, “malicious" pattern and symptom. One must, however, be aware that some of these terms mentioned above can also refer to the event in question - LOC, or to the failure that puts the system in a non-nominal state (e.g., rotor failure). In most articles, this is the case hence the methods found in literature must be approached with caution.

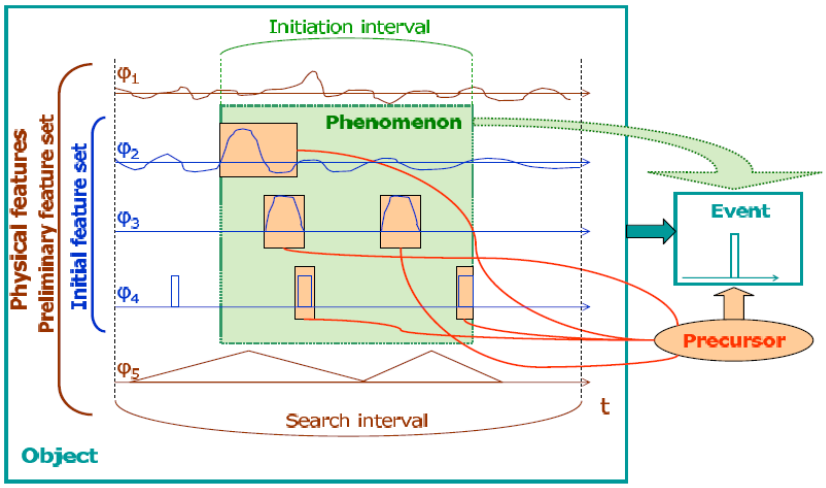


Figure 4.2: Precursor definition [54]

Regardless of the way the problem is looked at, most of the available data comes in the form of a time series. Indeed, the time series data of the states, parameters, and variables of the system can be used to make the prediction. In order to select an appropriate method, characteristics of time series and the type of problem must be understood. Below a brief summary of some basic concepts is provided.

1. *Linear vs. nonlinear*: linear methods approximate the time series through linear correlations. In a linear time series, all nonlinear behavior is due to noise. In real problems, however, linear dynamics are very unlikely, and the linear assumption is not accurate.
2. *Stochastic process*: the assumption is made that the time series is the realization of a random process. The properties of the time series can be evaluated by using statistical models.
3. *Stationarity*: the assumption is made that the stochastic rules do not change in the time series. This means that the distribution of the random variables remains the same. This can be checked with (partial) autocorrelation functions or an augmented Dickey Fuller test [55].
4. *Trend vs. seasonality vs. cycles*: a trend occurs when a certain time series increases or decreases at a specific rate throughout time. Seasonality is a periodic pattern that repeats itself at fixed intervals in time. Cycles are non-fixed patterns.
5. *Univariate vs. multivariate*: in univariate problems, one variable is used and analyzed. In multivariate problems, multiple ones are observed. Various combinations exist, single-input and single-output, multiple-input and single-output, and multiple-input and multiple-output.

6. *Problem dimensionality*: the dimensionality of a problem can increase due to the addition of more variables. However, a problem can also be multidimensional if only one variable is considered. Indeed, one variable can be used to make multiple features, effectively adding extra dimensions.

7. *One-step ahead vs. multistep prediction*: The time series value at $t+1$ is returned in a one-step-ahead prediction, whereas in a multistep forecasting problem, the value at $t+Y$ (where Y>1) is determined. A prediction horizon Y must therefore be defined. Use can be made of iterated predictions where the prediction of $t+i$ is used for $t+i$ (where 1<i<Y), or of direct methods where the forecast for $t+i$ is directly made.

8. *Classification vs. regression*: in the former, the output of the algorithm is a class label/tag, and in the latter, it is a continuous value. In classification, the purpose is to classify as early as possible whether or not the time series will lead to an event. In the regression problem, the purpose is to predict future time series values and detect if values or their associated statistical properties will exceed a certain threshold.

A small overview of some techniques used to extract precursors will be given. Machine learning in the context of KDD (Knowledge Discovery in Databases) in data mining must be considered.

4.1. Supervised methods

A fist apparent attempt to use machine learning techniques for precursor extraction can be seen in Dolenko's work. Her study focused on developing a committee of NN to forecast values of the geomagnetic storm index and to identify the precursors [54, 56, 57]. A search interval is divided into overlapping segments (see Figure 4.3). For each of these segments, a NN is constructed to predict the event using the features. Each of the constructed NN is shifted along the time axis until the minimum time for a prediction to be useful is reached. The precursors belong to the network that can most accurately predict the event.

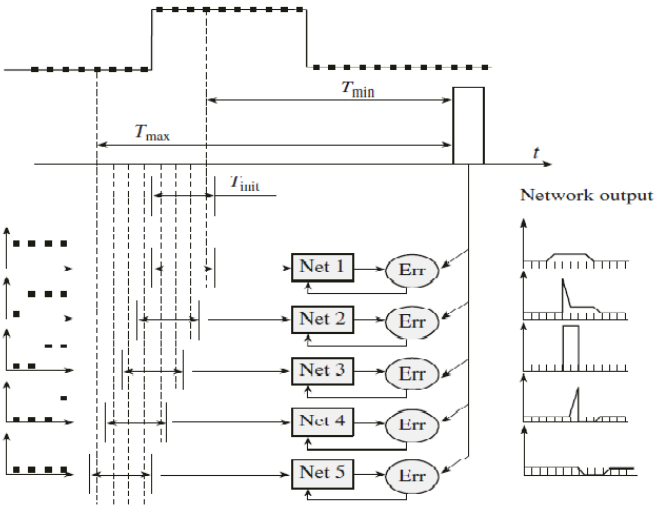


Figure 4.3: Precursor extraction using neural networks [56]

More advanced NN variants have also been used. In [58], a convolution NN is used in combination with a recurrent NN to capture linear and nonlinear dependencies as well as short and long term repeating patterns. Figure 4.4 illustrates the architecture of such a network. Traditionally the convolution NN is used in image analysis. The idea of a 2D image is used here by analyzing multivariate time series to make a robust prediction.

These methods are, however, black boxes meaning that the precise logic that is developed to identify a precursor is not clear. This can be overcome with logic-based techniques where explicit precursors can be returned. This takes the following form [59]:

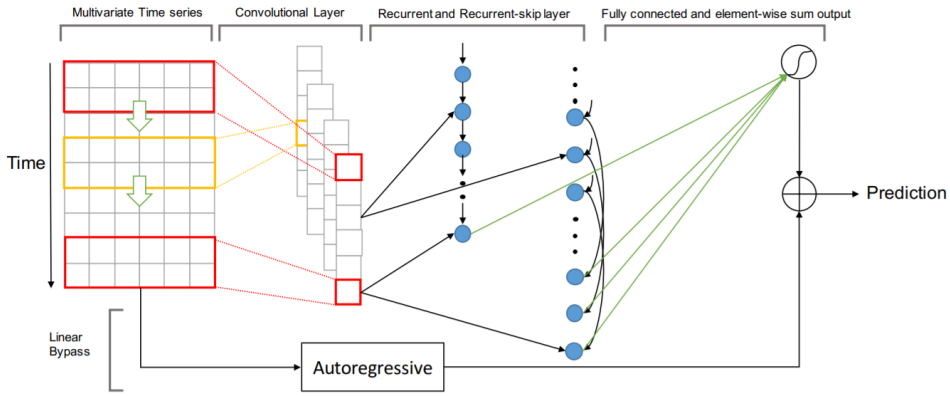


Figure 4.4: A convolution recurrent neural network [58]

if A_1, A_2, ... and A_h, occur within L units of time, then E occurs within time T

where A is a precursor. Deriving these rules, however, is not simple. These methods are traditionally applied in text and speech processing. As opposed to having words or letters following each other, time series have no natural segmentation.

Another logic-based forecasting method is shapelet analysis. The goal is to identify shapes in the time series, which can be used as precursors. Shapelets are supervised segments of series that are highly descriptive of the target variable [60]. Shapelet analysis has proven to be efficient in datasets where local variations (as opposed to global ones) are illustrative of the event [60]. One of the first times shapelet analysis was performed on a MTS was in 2012 through Multivariate Shapelet Detection (MSD) [61]. Shapelets were derived from multiple segments, with each segment coming from one variable. An example is shown in Figure 4.5. A score is assigned to each shapelet, and the irrelevant ones are discarded through a pruning method. The authors demonstrate that early classification of the time series can be achieved by only using 40-60% of the time series length. The method is not entirely descriptive as the extracted shapelets contain segments of the same length and originate from the same time interval. In reality, the precursors can manifest themselves at different locations and could have different time lengths.

Later, a method called Mining Core Feature for Early Classification (MCFEC) to obtain shapelets was proposed [62]. The extracted features are interpretable and effective for early classification. Moreover, MCFEC is faster than MSD on labeled data. The authors, however, acknowledge that using unlabeled data is more challenging and that the pruning process is time-consuming. More recently, in 2018, [63] made use of a multi-layer NN to learn shapelets in MTS. Competitive performance is achieved with the other available algorithms. Finally, a novel method called *ordex* is presented in [64] where relevant and non-redundant patterns are extracted from a high-dimensional search space. They achieve better accuracy than most of the other classification methods.

Overall even though these methods can extract precursors to an event, they have their drawbacks. Labeled data is needed, and for machine learning techniques, in large quantities to train the model. It is also not straightforward to understand what is happening inside these supervised machine learning methods as they are black boxes. With logic-based analysis, the time series problem is brought to a different level of abstraction. Shapes can be extracted with shapelet methods, but this process is extremely time-consuming - especially when the dimensional order of the problem increases. It also depends on the physical shape of the time series, whereas a precursor can manifest itself as an underlying statistical change, or a rare anomaly.

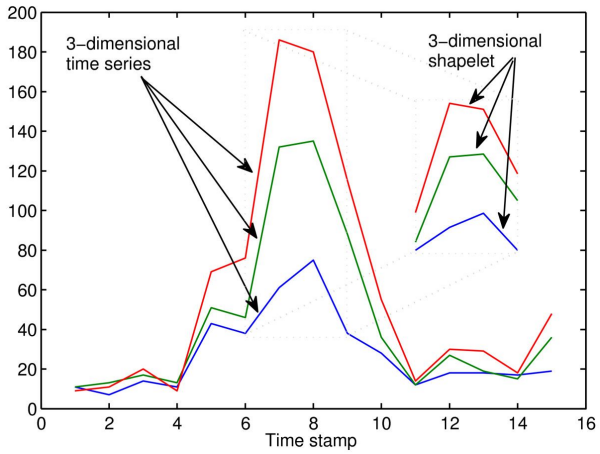


Figure 4.5: Illustration of a 3D shapelet in a three dimensional time series [65]

4.2. Semi-supervised methods

In semi-supervised methods, a portion of the data contains labels, and the rest does not. In the context of anomaly detection, the portion of labeled data is often anomalous-free. Various renowned authors have defined the anomaly detection problem in different ways. A common proposed definition is given below.

"*Anomaly detection is the task of identifying observations whose characteristics are significantly different from the rest of the data*" [66]

This definition seems to refer to anomalies more as outliers, which does not necessarily need to be the case. Other authors have defined anomaly detection as:

"*Finding patterns in data that do not conform to expected behavior*" [67]

In both definitions, similarities can be found. First, it is assumed that the definition of "normal" is well established. Secondly, a reference to a mechanism that detects if an observation conforms to normal behavior is made [68]. Usually, a model is constructed to define normal behavior using a nominal dataset. A detection algorithm is then used to identify anomalies. Even though anomaly detection is traditionally used for fraud detection in credit cards, intrusion detection for cyber-security, or fault detection in safety-critical systems, its application could be extended to LOC prediction.

Distinctions between the types of anomalies first need to be understood.

1. Point anomalies. These are defined as individual data instances that are anomalous with respect to the rest of the data. Usually, these are the types of anomalies that manifest themselves in credit card fraud and network intrusions.
2. Contextual anomalies. These are abnormal behaviors with respect to a predefined situation under a specific context. For example, 0° C is normal during winter but not during summertime.
3. Collective anomalies. These occur when a collection of data is not anomalous by itself, but it is to respect the rest of the data [67].

The type of anomaly drives the type of algorithm to be used. The algorithm itself also depends on the amount of knowledge known beforehand about the anomalies. Few anomalous samples might actually encourage the use of an unsupervised technique to create clusters. Usually, however, classification techniques are used where an observation is classified as nominal or not. In this context, semi-supervised techniques have been thoroughly studied by NASA. The methods developed are often eventually used for the design of FDD systems. The recent work NASA has done in this field will be summarized.

In 2015 NASA released the Adverse Condition and Critical Event Prediction Toolbox (ACCEPT) [1]. This toolbox takes previous ideas of NASA, such as [69–72] and combines them into a central tool. The architecture is presented in Figure 4.6 and is designed to detect anomalies by using a regression and an anomaly detection toolbox [73]. The main idea behind this toolbox lies with the fact that the regression can capture the nonlinear system dynamics.

The goal of the regression toolbox is to estimate a certain parameter that can be used to detect an anomalous condition. In literature, this is mostly the throttle. In order to estimate this parameter, a Multivariate Time Series (MTS) input set of features u_k is used to train a nonlinear model. The training of this model is done in a supervised way. Any model that can take multiple features as an input and can model the system's nonlinearities can be employed. Once a parameter of interest (such as throttle) $\hat{z_k}$ is estimated, it is compared to the non-anomalous target z_k and a residual y_k can be obtained.

Next, the detection toolbox is used to find faults or anomalies. Use is made of both the nominal training data as well as the data containing the anomalies. Anomalies are detected by using alarm systems that employ either a Linear Dynamic System (LDS) or an empirical variant of the redline alarm system. The LDS is used to model the dynamics that the regression was not able to capture. Using an LDS system, the assumption that the plant is modeled through a linear dynamic state-space system driven by Gaussian noise is made. The parameters to be found are presented in the θ vector and are used to build a Kalman filter. These parameters are found by using the process and the measurement noise of the system. Process noise is used as it provides a statistical characterization of the input features u_k, whereas measurement noise provides a statistical characterization of the target parameter z_k. Algorithms such as the expectation-maximization can be used to determine θ [69, 74]. Finally, a Kalman Filter calculates state estimates, and the Kalman prediction is used to propagate the residual in the prediction horizon.

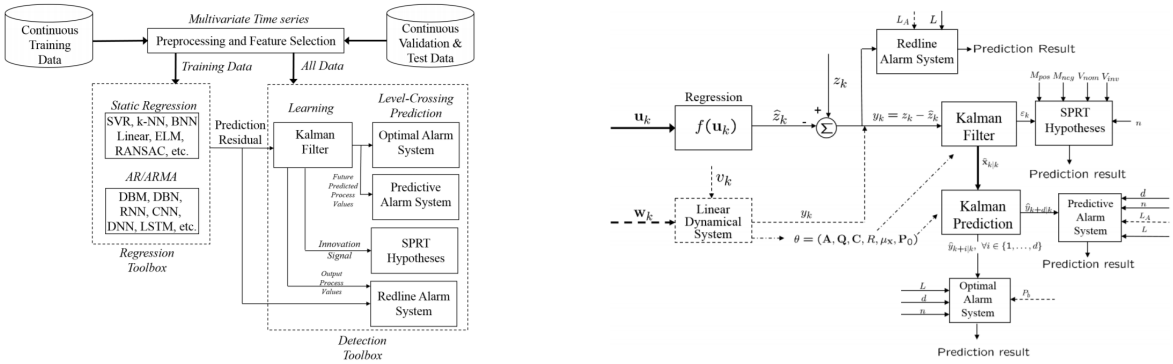


Figure 4.6: Left: functional architecture of ACCEPT, right: signal flow digram of ACCEPT [73]

The detection systems used in the toolbox are briefly described below. A more extensive explanation of these detection systems is found in [73, 75].

1. Standard exceedance: This is the only method that does not require the use of the LDS. Use is made of the residuals of the model. A threshold L_A is defined that warns of a more critical impending event. It is obtained by generating relevant alarm statistics through Monte-Carlo simulations. An envelope $[-L_A, L_A]$ is found, and an alarm is triggered if the residual exceeds it: $\|y_k\| > L_A$.

2. Redline alarm system: Use is made of the residuals of the model. The LDS parameters are used to find the alarm statistics. The alarm is triggered when $\|y_k\| > L_A$.

3. Predictive alarm system: Use is made of the Kalman prediction. Future steps are looked at to determine if the residual will fall in the envelope or not: $\|\hat{y}_{k+Y}\| > L_A$. The alarm statistics can be generated by either Monte Carlo methods or by using numerical integration of the LDS parameters.

4. Optimal alarm system: Use is made of the Kalman prediction. An alarm is provided when the conditional probability of an event exceeds a certain predefined probability: $P(E|y_0, ..., y_k) \leq P_b)$. Monte

[1] https://ti.arc.nasa.gov/opensource/projects/accept/

Carlo methods can again be employed to generate relevant statistics, or alternatively, numerical integration of the LDS parameters can be performed.

5. SPRT tests: Use is made of the innovation signal (a secondary residual whose spectral properties include signal whiteness) obtained from the Kalman filter. An optimal stopping rule is defined that uses test hypotheses to distinguish between probability distributions that characterize anomalous behavior from its counterpart.

Finally, a Reinforcement Learning (RL) technique was also used in the context of semi-supervised precursor extraction. RL is a branch of ML focused on determining the optimal set of actions of a system such that a particular reward function is maximized. Even though this type of algorithm is mostly used in knowledge-based control systems to teach a robot through trial and error how to do something, an interesting application of RL was found in the context of precursor detection. In 2016, NASA's Janakiraman developed an algorithm called ADOPT (Automatic Discovery Of Precursors in Time series data), which employs RL and Markov decision processes to detect precursors to events in commercial aviation data [76]. The framework of ADOPT is presented in Figure 4.7.

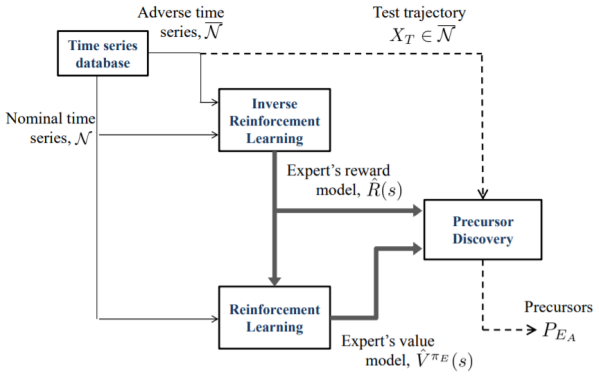


Figure 4.7: ADOPT framework used in [76]

Nominal and adverse time series are used as example trajectories for the model expert. Every time instant in the adverse time series corresponds to a sub-optimal action taken by the expert. Using Bellman's optimality [2], the optimal action that can reduce the likelihood of this event occurring can be determined. By comparing the expert's value function to the sub-optimal action it took that led to an event, the precursors can be found. Finally, by setting a threshold to filter out low strength precursors, the important ones can be identified.

Overall, in these semi-supervised techniques, both nominal and adverse time series are required. This is, in the context of quadcopter failure, not possible to obtain, as each simulation will eventually lead to failure. Additionally, obtaining a reference for what is happening is very challenging without a model. In damaged systems, the nominal model is often not accurate anymore; hence new equations of motion must be developed, which is not always possible.

4.3. Unsupervised methods

In this part, the unsupervised techniques that can be used for forecasting or pattern identification will be explored.

As mentioned earlier, unsupervised techniques do not make use of any labels. The structure within the observations is analyzed, but there is no way of determining if the "right" results are obtained as they are not known. In unsupervised anomaly detection methods, there is no difference between the test and training

[2] This is an optimality principle used in RL.

data. Distance or densities are therefore used to determine what can be qualified as an outlier. In this section, several techniques are presented. For an in-depth classification of the available methods, one is referred to Goldstein's paper [77].

The most popular unsupervised anomaly detection techniques make use of clustering and distance measurement methods. Clustering methods calculate the distance of each point to a cluster center. If numerous points belong to a cluster, the instance can be called normal. The most popular clustering method is k-means, where k clusters are defined. Points that do not belong to any cluster are considered anomalous. Determining if a point belongs to a cluster or not is evaluated using distance measures. A more advanced measure is the Mahalanobis distance, where a covariance matrix is used to return the degree of an instance being an outlier. In [78], it is used for online anomaly detection of autonomous robots.

Overall, the application of unsupervised techniques on time series remains quite abstract. The dimensionality of time series must be reduced, and clustering of time-dependent data points remains a vague concept.

4.4. Signal analysis methods

Finally, signal analysis methods can be considered. Rather than attempting to label a dataset, the signal is analyzed, and specific properties are observed.

One such method is called Critical Slowing Down (CSD). Over the last years, there has been an increasing interest in finding Early Warning Signals (EWS) to critical transitions in a complex system. M. Scheffer has been leading a research group that focuses on identifying such signals in real-world systems [79]. The idea is to determine through time series analysis, significant statistical indicators that can predict this upcoming tipping point.

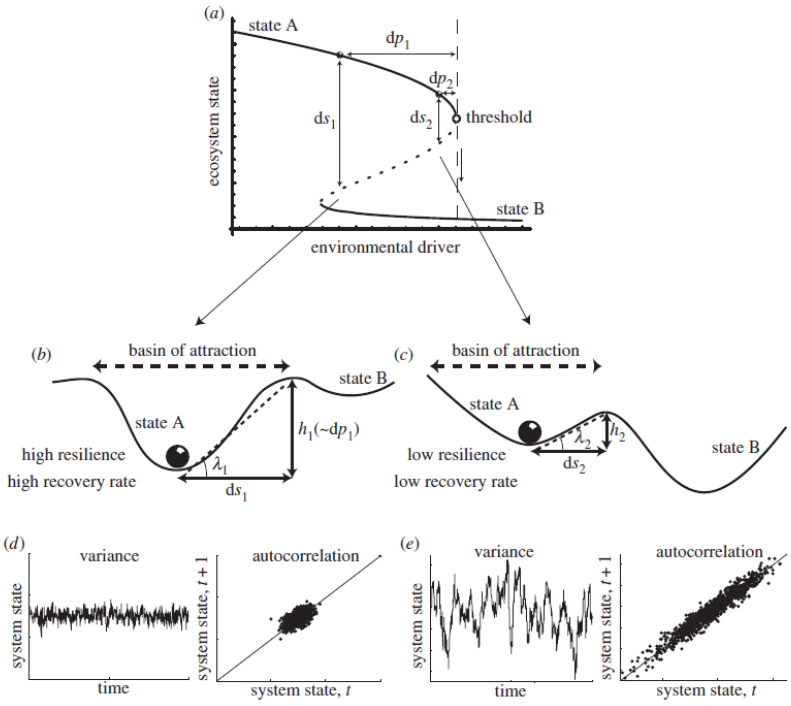


Figure 4.8: The relationship between tipping and early warning signals [80]

The concept behind tipping and how EWS are generated is explained in Figure 4.8. Precursors to tipping can be extracted by looking at the resilience of a system prior to failure. Indeed, as the system approaches the tipping point, the eigenvalues which determine stability reduce, and resilience decreases. In Figure 4.8, resilience depends on two elements: the distance to the critical point, dp, and to the unstable state, ds, when observing the parameter space. As the driver is increased, the state decreases and the system approaches the

threshold. During this process, the basin of attraction becomes flatter; hence its slope decreases. The slope is defined as the ratio between dp/ds.

As the slope tends to zero, the recovery rate of the system also reduces, i.e., the time required for a system to return to a stable equilibrium after a disturbance increases. The system becomes slower hence the name CSD. Signs of CSD can be observed by analyzing specific statistical indicators such as variance and autocorrelation, which should rise as the threshold point is approached.

Not all systems are likely to show EWS. To understand whether this method can be used on a robotic system such as a quadcopter, the flowchart used in [80] must be considered.

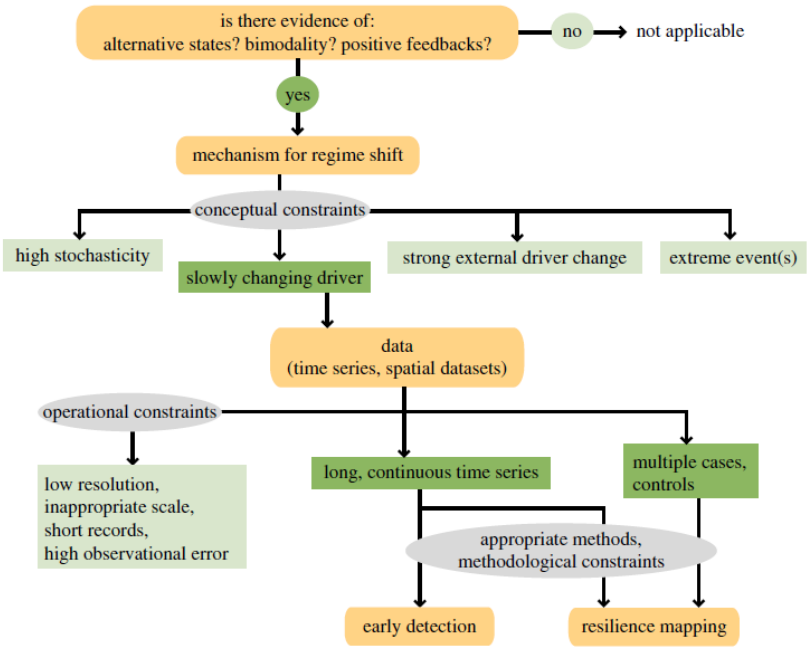


Figure 4.9: Conceptual considerations when using CSD indicators as EWS [80]

The first question can be answered by stating that as the quadcopter approaches critical wind speeds, it tips, and as a consequence, it fails. This failure state could be an alternative one. The mechanism that pushes the quadcopter to failure is the increasing wind speeds. Special attention must be given to the rate of increase of the wind speed as Figure 4.9 suggests that slowly changing drivers are required. Finally, the data considered is a time series. The length of this time series depends on the sampling frequency and the rate of increase of the wind speed. Longer samples are preferred. Additionally, these samples must be of good quality as data plagued by observational errors can also lead to missed warnings. Under these conditions, it could be interesting to use CSD indicators to generate EWS capable of predicting LOC.

Chapter 5

Methodology

In this chapter, the methodology that will be followed is provided. First, the initial research strategy is briefly discussed. Next, the experimental set-up will be elaborated upon. The tests and experiments that will be conducted will be elaborated upon.

5.1. Research strategy

The objective of the research is to extract knowledge from quadcopter flight data to predict loss of control. Use will be made of critical slowing down identifiers. The proposed methods in literature have never been used for such a purpose. Several toolboxes do exist, but they have been used on different types of data structures. As such, in order to determine whether quadcopters do show such behavior before loss of control, a research strategy is proposed. It is presented in Fig. 5.1.

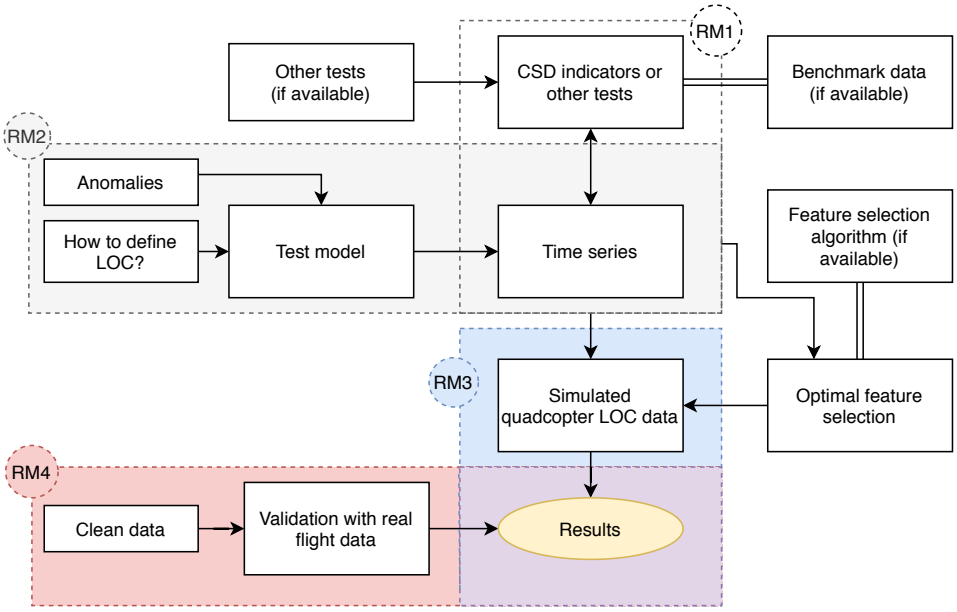


Figure 5.1: Preliminary research methodology

Four modules that can be seen in Fig. 5.1 are described in Tab. 5.1:

1. **RM1**: The first part is the preliminary research. With the literature study finished, a more in-depth research of the CSD indicators can be performed. A better understanding of how the working principles can be achieved and other similar tests can be searched for. These tests will be applied to time series data, which comes from RM2. A double arrow is used in the diagram to mark that this is an iterative

Table 5.1: Description of the different research modules.

ID	Group name	Description
RM1	Preliminary research	The preliminary research will yield a better understanding of the working principles behind CSD. Other tests can also be searched for.
RM2	Test case study	Data from a simple nonlinear model will be generated. Different anomalies can be introduced to simulate LOC. The presence of CSD will be examined, and an idea of the useful features can be obtained. This case study will serve as the verification of the quadcopter results.
RM3	Quadcopter test case	The methods to detect CSD will be applied to simulated quadcopter data. Optimal features will have to be selected.
RM4	Validation	Real time series data of the quadcopter will be used to validate the results obtained.

process, where indicators are used, and time series with different anomalies are generated. Checks can be made with a benchmark dataset to verify correct implementation - if such data exists.

2. **RM2**: The second research module is the development of a test case model. This test case model should be a simple nonlinear system, that is representative of the quadcopter, but possesses fewer states and is easier to understand. Various anomalies can be introduced in order to simulate loss of control. LOC must still be defined in order to ensure that the prediction occurs in a timely manner. CSD indicators will be applied to the time series generated by this test case. Doing so will provide an initial understanding of how they function, and will ultimately verify the results obtained in RM3/4. More information on the nonlinear system will be provided in Sec. 5.2.2.

3. **RM3**: The third research module is focused on identifying indicators of CSD in a simulated quadcopter model. This model already exists and can be readily applied. The quadcopter system is more complex than the pendulum, but the anomaly used to simulate LOC is known to the user. The main challenge is that a complex system has more variables that can be used to predict LOC - hence the right ones must be selected.

4. **RM4**: The final research module is the validation one. Real SRF quadcopter data will be used, and the presence of CSD or not will be examined. The data used will need to be cleaned and pre-processed. Furthermore, comparing the obtained results with the ones from Kersbergen [41] can further validate the validation results.

5.2. Experimental set-up

In this section, the experimental set-up is discussed. The tools required, the types of experiments, and the tests will be discussed. Simulations will be designed to test the algorithm, and a flight test will be performed to collect data.

5.2.1. Apparatus

The proposed project is conducted at the Control & Simulation department of the Delft University of Technology. If data needs to be collected, this can be done in several ways. Flight test can be performed in the Cyberzoo or in the OJF wind tunnel. The OJF wind tunnel enables users to increase the wind speed in order to see how the quadcopter is performing, and the Cyberzoo is an arena where the quadcopter can be flown safely. Data during flight tests can be recorded through cameras, or by connecting to the quadcopter and downloading the time series of the states directly to a ground station.

5.2.2. RM2

A simple nonlinear model that could be used in RM2 is the inverted pendulum on a cart. A force can be applied on either side of the cart in order to balance the pendulum upright. This nonlinear system is simple enough as it has four external states, the position/velocity of the cart and the angle/angular velocity of the pendulum. Depending on the controller used, it also has some internal states that could be analyzed.

Various anomalies can be introduced in this system to simulate LOC situations. A reference signal that the rod has to follow could be introduced and increased throughout time. The friction coefficients between the cart and the rail, but also between the cart and the pendulum can be increased throughout time. Moreover, movement and speed boundaries can be defined, and forces such as wind could be simulated.

Finally, LOC must be detected in the time series data of the inverted pendulum. One way this could be done is by defining the SFE through Monte Carlo analysis. This has, for instance, been done by defining the forward and backward reachable sets of a system for a given time horizon [37, 38]. The moment the SFE is exited, the system is in a LOC condition - hence the precursors must be situated somewhere before.

The SFE can be computed for the inverted pendulum by using a Monte Carlo approach. The initial condition is set at $\theta = 0$ and $\dot{\theta} = 0$ as this is the unstable trim position of the pendulum. Figure 5.2 shows this trim position in white. The reachable set is computed for a time horizon of 1s by generating 102400 random control sequences. At each discrete time step, a random control input $u \in [-25; 25]$ was applied. In Figure 5.2, the shaded area in between both sets is the SFE. This safe set is, however, slightly smaller than the real one. Indeed, in order to obtain the actual boundaries, an infinite amount of samples must be used in the Monte Carlo Simulation.

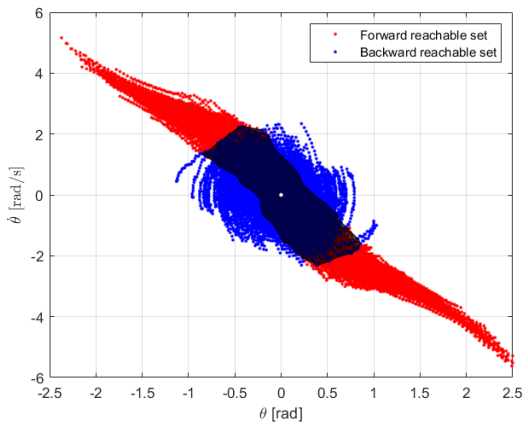


Figure 5.2: In red and blue respectively the forward and backwards reachable sets for a time horizon of 1s. The white point indicates the trim position. The dark shaded region is the SFE.

Another way LOC can be defined is by selecting a moment before failure at which the quadcopter has most of its actuators saturated. Actuator saturation was indeed shown to have a direct link with failure in the work of Kersbergen [41].

5.2.3. Tests

Various tests will be performed in each experiment - especially the test case study and the quadcopter test case. These will be done to verify the results obtained.

1. **Comparative tests**: These tests are intended to compare results obtained with the algorithm to the ones obtained in literature. If results are significantly different, the algorithm might not be applied in the right manner.

2. **Verification tests**: These are simple tests that will be done to check whether the code is working like

it is supposed to. Such verification tests include calculations of indicators through different methods. Similar results indicate proper implementation.

3. **Sensitivity tests**: These tests study the effect of parameters on the outcome of the algorithm. By varying important variables, their influence on the results can be determined.

These tests will be critical during the first experiment case study, where limitations of the CSD indicators are explored. In the quadcopter case study, these tests will be used to ensure their proper implementation. This is indeed needed as the complexity of the system could yield unsatisfactory results that are unrelated to the implementation. Finally, in the last experiment, verification tests will be used to ensure the data is adequately pre-processed. The sensitivity to noise removal can also be examined as possible precursory patterns should not be removed when cleaning the dataset.

5.3. Retrospective analysis

Several retrospective comments about the methodology are given in this section. The moment feature selection is performed, and the decision of not pursuing RM3 are discussed.

5.3.1. Feature selection

Feature selection is important in a multivariate data analysis problem. As such, a feature selection algorithm could be developed and validated in RM2, then be used in RM3. The significant variables can then also be used in RM4.

However, the inverted pendulum on cart problem only had three relevant variables: the states, θ & ω, and the control input, u. No variable selection algorithm is needed as all three variables could be analyzed for precursory signs of LOC. Ultimately, CSD theory encourages the selection of the angular rate as it is more sensitive to show signs of CSD than the angle itself. No variable selection is employed.

The quadcopter has significantly more variables that could be analyzed. As RM3 is not performed (see next section), the selected variables are decided upon in RM4. The techniques employed will be validated rather than applying the same technique in RM2. As such, variable selection was only considered in RM4, rather than RM2 as indicated in Figure 5.1.

5.3.2. RM3

Next, RM3 is discussed. Even though ideally, a simulation of a damaged quadcopter can provide more insight into the type of bifurcation point, RM3 was ultimately not implemented for various reasons.

First, an accurate model of the dynamics of such a vehicle is obtained by fitting a model over flight data. This means that the model is only as accurate as the data collected. Additionally, certain elements can be overlooked and possible bifurcations can be omitted when determining the equations. This was observed when a bifurcation analysis was carried out on a quacopter flying in nominal configuration. The high-fidelity aerodynamic model of a quadrotor in the high-speed flight derived in [40] was analyzed. Only rotational longitudinal dynamic motion was considered:

$$\dot{\theta} = q \tag{5.1}$$

$$\dot{q} = \frac{C_{m,0(V_x,V_z)} + C_{m,1(V_x,V_z)}\, u_1 + C_{m,2(V_x,V_z)}\, u_2}{I_y} \tag{5.2}$$

with

$$
\begin{aligned}
C_{m,0} &= 1.03\times10^{-2}V_x - 6.77\times10^{-4}V_x^2 + 8.64\times10^{-3}V_z + 7.17\times10^{-5}V_x^2V_z + 2.63\times10^{-4}V_xV_z^2 \\
C_{m,1} &= 1.52\times10^{-1} + 1.04\times10^{-3}V_x^2 + 1.66\times10^{-3}V_xV_z - 1.86\times10^{-3}V_x \\
C_{m,2} &= -1.63\times10^{-1} + 8.04\times10^{-3}V_x - 2.11\times10^{-4}V_xV_z - 6.31\times10^{-4}V_x^2
\end{aligned} \tag{5.3}
$$

with V_x, V_z, the velocity projections on the body frame. In the real experiments, the quadcopter does remains at a fixed positions and the wind speed (V) is increased. As such, both velocities are defined as $V_x = Vcos(\theta)$ and $V_z = Vsin(\theta)$. A PD controller was implemented:

$$u_1 = \bar{u} + (K_{p1}(0-\theta) - K_{d1}q) \tag{5.4}$$

$$u_2 = \bar{u} - (K_{p1}(0-\theta) - K_{d1}q) \tag{5.5}$$

and $K_{p1} = 10$, $K_{d1} = 1$ and $\bar{u} = 0.9$ in order to have a normalized value that does not imply completely saturated rotors. Using *Matcont*, a bifurcation analysis was performed. It is presented in Figure 5.3. Several elements that emphasize the point made at the start can be observed. When the wind speed is increased beyond 11 m/s, the quadcopter still seems to be in equilibrium even though it should crash. This is because the equations are only valid for the wind speeds lower than 16 m/s - velocity at which the vehicle crashes. The implementation of a not-so-robust controller can reduce this velocity to 11m/s. From the bifurcations analysis, however, it appears that the system still is in equilibrium beyond this critical value: the pitch angle reduces when the wind speed is increased beyond 11m/s.

Secondly, according to the diagram, the quadcopter does not have a zero pitch angle at zero wind speed. This, again, can be due to the equations not being adequate for such low wind speeds. Nonetheless, both these points highlight why using equations derived from flight data are problematic in a bifurcation analysis.

Finally, a bifurcation point can exist in the real model, but due to the equations being derived from real flight data, the mathematical expressions can remove the presence of such a point. Indeed, the type of bifurcation affects the way eigenvalues move towards the right plane.

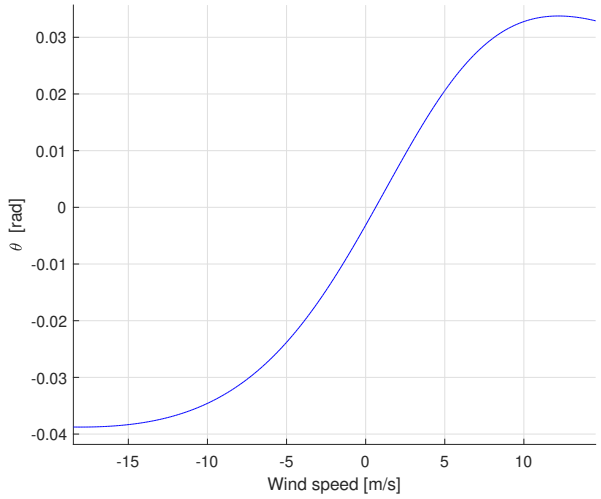


Figure 5.3: Bifurcation analysis of the longitudinal rotational dynamics presented in [40]

Part III

Master of Science Thesis Papers

Critical slowing down indicators as predictive precursors to loss of control in an inverted pendulum on cart system

A. van der Pluijm
Graduate Student, Department of Control and Simulation
Delft University of Technology, Kluyverweg 1 2629 HS Delft, The Netherlands

S. Sun
Ph.D. student, Department of Control and Simulation
Delft University of Technology, Kluyverweg 1 2629 HS Delft, The Netherlands

C.C. de Visser
Assistant Professor, Department of Control and Simulation
Delft University of Technology, Kluyverweg 1 2629 HS Delft, The Netherlands

***Abstract*—The focus of this paper is to provide a first insight into the usage of data-driven methods to predict Loss of Control (LOC) in fast-changing dynamical systems. The use of Critical Slowing Down (CSD) indicators, an approach commonly employed in slow time-varying ecological systems, was further developed to fit the target problem. These CSD indicators have proven to reveal the moments where a system is losing resilience and is tipping to an alternate state. To test this on a robotic system, a nonlinear model of an inverted pendulum was subjected to strong external disturbances, causing it to tip over.**

Early Warning Signals (EWS) were successfully found when the wind force applied on the pendulum was slowly varied, but less so, when this driver was quickly increased. The presence of EWS was validated by controlling the pendulum with a Nonlinear Dynamic Inversion (NDI) controller that prevents it from tipping over. As expected, few EWS were found when the driver was slowly varied, but a considerable amount were observed with the fast-changing driver, highlighting issues when transients are analyzed. The difference in both controllers also strongly suggests that the type of controller used can lead to more or fewer signs of CSD, thus also affecting the presence of EWS.

Finally, the presence of EWS also depends on a variety of underlying parameters that need to be appropriately tuned. EWS are indeed very sensitive to the rolling window sizes used to generate indicator values. Larger window sizes yield more reliable EWS. This, however, also implies that more time is required to determine if a signal can be considered as an EWS. Moreover, these EWS do not always appear at the same location across various model runs. This should not come as a surprise, as the residuals of the time series are analyzed, but does raise doubts regarding potential applications, such as LOC prediction.

***Index Terms*—critical slowing down, early warning signals, loss of control, precursors, critical transition**

I. Introduction

Complex systems often exhibit regime shifts that suddenly cause their qualitative dynamics to change. This change occurs when a system approaches a tipping point, a moment at which it undergoes a *critical transition*. Such transitions occur in response to a small change in the input of a system when it is at a critical state [1]. Tipping effects can be regrouped into three categories:

- Bifurcation tipping, where the output of a system changes abruptly due to a bifurcation;
- Noise-induced tipping, where random fluctuations push the system away from equilibrium;
- Rate-induced tipping, where a system fails to track a stable equilibrium due to a rapid change in parameters causing it to tip.

Examples of tipping can be found in different branches of science, notably in (i) climate dynamics where for example the slow increase in greenhouse gases causes Earth's climate to tip [2]–[5], (ii) in ecology where behavioral changes occur leading to species extinction [6]–[8], in financial systems where market crashes occur [9]–[12] and in pathology, where the start of a disease can cause an outbreak [13], [14].

Avoiding unintentional shifts in complex systems is of great interest for obvious reasons. This can be achieved by predicting critical transitions before they are triggered in order to intervene and prevent the shift from happening. Recent results suggest the existence of Early Warning Signals (EWS) that can be used to determine if such a critical state is being approached [15]. When a system approaches a critical transition, its rate of recovery from perturbations slows down. This can be observed through a rise in statistical indicators, such as autocorrelation and variance, when analyzing the time series of selected system variables. Other indicators that can be observed include skewness, kurtosis, and changes in the power spectrum. These indicators attempt to measure system resilience, and as a result, its distance to a tipping point [2].

The main contribution of this paper lies in applying the traditional loss of resilience indicators to a closed-loop controlled robotic system. In this case, a nonlinear model of an inverted pendulum on a cart was used. Such systems are heavily

dependent on the environment they interact with. Changes in environmental drivers can push robotic systems into a loss of control situation, after which they often fail. By studying the effects of various systems and analysis parameters on the EWS, the limitations of their usage can be better understood. If done properly, the developed methodology can be applied to more complicated robotic systems such as quadcopters - something which has not yet been done.

This paper is structured as follows. First, an in-depth description of Critical Slowing Down (CSD) theory is provided in Section II. Next, in Section III, the methodology is provided. The data collection techniques and the tests that are performed are explained. Section IV gives the results, and a discussion is held in Section V.

II. Critical slowing down

The main principle behind the CSD research rests on the fact that a small disturbance can trigger a substantial change in a real-world system. This occurs when a system reaches a threshold point called a bifurcation or tipping point. When this point is surpassed, an irreversible shift from one equilibrium towards an alternate state occurs [16]. Research has revealed that statistical signals precede these nonlinear transitions. These indicators usually attempt to measure the resilience of a system, i.e., the time it takes for a system to return to an equilibrium state after a disturbance.

Assume that a process $x = x_t : t \leq 0$ can be modeled by a stochastic differential equation that obeys a certain dynamical law $f(\cdot)$, and is subject to perturbations and measurement uncertainties:

$$\mathrm{d}x_t = f(x_t, \vartheta(t))\mathrm{d}t + g(x_t)\mathrm{d}W_t \tag{1}$$

$f(\cdot)$ is also known as the smooth drift function. It describes the deterministic differential equation x_t would obey if $g(\cdot)$ = 0. $\vartheta(t)$ is a time-varying input. Noise is modeled by a one-dimensional Brownian motion $W = W_t : t \leq 0$ multiplied by g(·), an unknown diffusion function which determines how stochasticity interacts with the state variable. In such systems, a potential function ∇V_t is assumed to exist, which reflects the regions the system is most likely to travel towards [17]. This potential is assumed to be smooth and slowly varying. When a bifurcation point is approached, the shape of ∇V_t changes around the equilibrium, and the stable point starts turning into an unstable one. During this process, the real part of the eigenvalues of the system also diminishes smoothly to zero. Excursions of x away from the stable point - that can, for example, occur when the system is subject to perturbations, grow both in length and time. This means that the recovery rate is decreasing until it eventually becomes zero when the critical point is reached. At this moment, the equilibrium vanishes, and the system's eigenvalues move into the right half-plane.

This process is shown in Figure 1. When the system is far away from the bifurcation point, the recovery rate to perturbations is high, and the basin of attraction towards an alternate state is wide. This is shown in the left diagram of Figure 1. As the system approaches the bifurcation, the basin of attraction shrinks and becomes flatter. The system is less

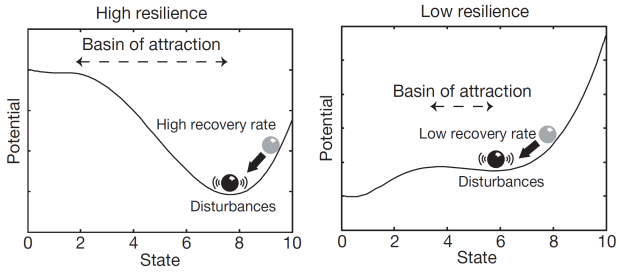


Fig. 1: Resilience of a system as it approaches a catastrophic bifurcation [18]

resilient to perturbations and needs more time to recover from them. Subsequent states $[x_{k-1}, x_k]$ are more alike, resulting in an increase in the system's memory of the perturbation. This leads to a rise in variance and autocorrelation in the time series of the state variables. More information on why this precisely happens is given in Section III-B2, where the indicators used are discussed. By observing local indicator increments, potential Early Warning Signals (EWS) of an upcoming bifurcation can be identified [19]. Additional indicators have since been investigated, and some even go beyond identifying CSD. A brief review of recently published papers that identify EWS in complex systems is provided in Table I. The focus of this paper does remain with the CSD indicators.

It must, however, be noted that not all transitions can be predicted by CSD indicators. In [35], it was concluded that the warning indicators varied strongly in reliability and that agreement between various indicators was low. In [36], the statistically important changes in the time series were detected moments before the bifurcation making them non-usable warning signals. As a last example, in [37], a lack of CSD indicators before a financial crash suggests that not all complex systems undergo critical transitions.

It is, therefore, essential to understand how and why different types of systems can transition to an alternate state. Indeed, regime shifts can occur without the presence of CSD, and vice versa, not all CSD observations involve regime shifts. Moreover, not all regime shifts include bifurcations, and bifurcations can exist in a system with and without there being signs of CSD [38]. This is illustrated in the Venn diagram in Figure 2.

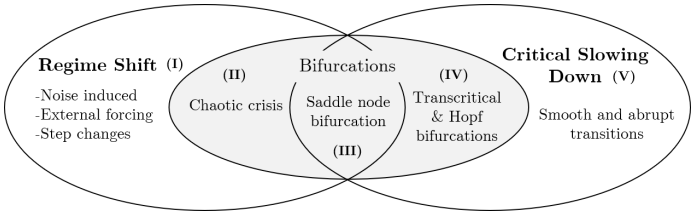


Fig. 2: Venn diagram representing the logical relations between regime shifts, critical slowing down and bifurcations. Adapted from [38].

In the following sections, the contents of this Venn diagram are discussed. In Section II-A, the effects of the type of bifur-

TABLE I: A survey of the most recent EWS literature. The title, corresponding reference numbers, publication date, whether the data used was real or simulated and the indicators used are shown. AR1: lag-1 autocorrelation, SD: standard deviation, PS: power spectrum, DFA: detrended fluctuation analysis, Var: variance, Skw: skewness, Krt: Kurtosis, LE: Lyapunov exponent

Year	Title	Reference	Data	EWS indicators
2019	Predicting noise-induced critical transitions in bistable systems	[20]	Simulated	LE, Shannon entropy
2019	Critical speeding up as an EWS of regime switching	[21]	Simulated	AC1, Var
2019	Detecting bifurcation points in a memristive neuron model	[22]	Simulated	AC1, LE
2019	EWS of recovery in complex systems	[23]	Simulated	AC1, SD, coefficient of variation
2019	EWS in plant disease outbreaks	[13]	Simulated	AC1, Var, Krt, Skw
2019	No evidence for CSD prior to human epileptic seizures	[24]	Real data	AC1, Var
2018	How one might miss EWS of critical transitions in time series data: A systematic study of two major currency pairs	[25]	Real data	AC1, Var, low-frequency PS
2018	Robustness of EWS for catastrophic and non-catastrophic transitions	[16]	Simulated	AC1
2018	EWS precede cyanobacterial blooms in multiplewhole-lake experiments	[6]	Real	AC1, Var
2018	Predicting tipping points of dynamical systems during a period-doubling route to chaos	[26]	Simulated	AC1, Var, Krt, Skw
2018	EWS for Dansgaard-Oeschger events in a high-resolution ice core record	[27]	Simulated	AC1, Var, wavelets
2018	A novel scaling indicator of EWS helps anticipate tropical cyclones	[28]	Simulated	AC1, DFA, decay rate of the PS
2018	CSD as an EWS for financial crises?	[10]	Real	AC1, mutual information, SD
2018	Analysis and Predictability of Tipping Points with Leading-Order Nonlinear Term	[29]	Simulated	Drift diffusion models
2017	CSD as an early warning of transitions in episodes of bipolar disorder: A simulation study based on a computational model of circadian activity rhythms	[30]	Simulated	Var, Skw
2017	Alternative stable states and spatial indicators of CSD along a spatial gradient in a savanna ecosystem	[7]	Real	AC1, Var, PS, Skw
2016	EWS of regime shifts in coupled human –environment systems	[31]	Simulated	Var
2016	EWS for critical transitions in a thermoacoustic system	[32]	Real/Simulated	AC1, Var conditional heteroskedasticity
2016	EWS, nonlinearity, and signs of hysteresis in real ecosystem	[33]	Real	Spatial EWS
2016	Early-warning indicators for rate-induced tipping	[34]	Simulated	AC1, Var

cation point present in a system are covered. In Section II-B, the effects of different drivers are explained.

A. Bifurcations

Bifurcation driven tipping is an important mechanism in which sudden behavioral changes in a system are observed [39]. The stable solution of the deterministic differential equation shown in Equation (1), is called the quasi-static attractor. When $\vartheta(t)$ passes through a bifurcation point, the quasi-static attractor loses stability. Different types of bifurcations have been identified in stochastically forced open systems of the type Equation (1):

i. Dangerous bifurcations, where a sudden and fast dynamic jump to a distant attractor (a different steady-state) is observed.
ii. Safe bifurcations where an attractor state is replaced by another nearby attractor subtly grows in importance.
iii. Explosive bifurcations that lie in between the safe and the dangerous ones. They are less common and are not further discussed. More information about this and the other types of bifurcations can be found in [40], [41]

EWS usually work best when CSD co-occurs with a regime shift (case (III) in Figure 2). This happens when a system passes a dangerous bifurcation [42]. A classic example of such a catastrophic bifurcation is the saddle-node bifurcation, shown in the left graph of Figure 3. An example of an Ordinary Differential Equation (ODE) that contains a saddle-node bifurcation is:

$$\frac{dx}{dt} = r_1 + x^2 \tag{2}$$

where $r_1, x \in \mathbb{R}^n$. For $r_1 < 0$, the ODE has two equilibrium points at $-\sqrt{-r_1}$ and $\sqrt{-r_1}$. For $r_1 = 0$, there is one equilibrium point. This is the bifurcation point. For $r > 0$, there are no real equilibrium points. A system, therefore, contains usually two stable equilibria for a particular control parameter or driver. When increasing the driver, the two equilibria collide, and the system "jumps" to an alternate stable state. This occurs when the eigenvalues of the system intersect zero and move to the right half-plane.

CSD is also present in systems that do not have bi-stable configurations (case IV in Figure 2). These systems usually do not jump to an alternate state, but instead, approach them smoothly, and are thus also labeled as non-catastrophic bifurcations. The regimes are qualitatively different but quantitatively similar [38]. The eigenvalues of the system intersect zero at the tipping point when moving towards the right half-plane. Examples include the transcritical bifurcation (shown in the middle graph of Figure 3), Hopf bifurcations, and pitchfork bifurcations. The former has an ODE of the form:

$$\frac{dx}{dt} = r_1 x - x^2 \tag{3}$$

where $r_1, x \in \mathbb{R}^n$. The bifurcation occurs for $r_1 = 0$. The ODE has two fixed points: $x = r_1$, $x = 0$. When $r_1 < 0$, the fixed point $x = 0$ is stable and $x = r_1$ is unstable. Conversely, when $r_1 > 0$, the fixed point $x = 0$ is unstable and $x = r_1$ is stable.

Finally, there is a class of bifurcations in which no CSD is expected before a regime shift (case (II) in Figure 2). This is the case of crises. In dynamical systems theory, a crisis refers to a sudden change in the dynamics of the system

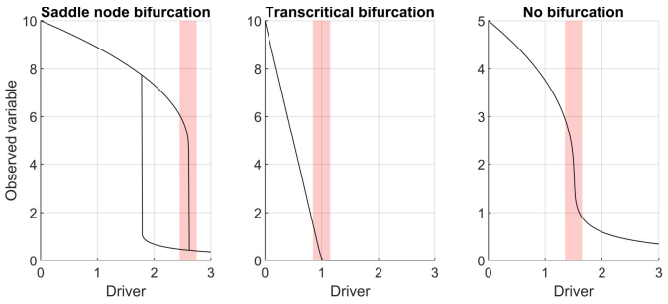


Fig. 3: Various types of bifurcations are likely to occur with CSD. Shaded in pink is the region where the transition occurs. The graph on the left shows a catastrophic saddle-node bifurcation. The middle graph illustrates a non-catastrophic transcritical bifurcation. The right graph shows a strong non-linear response of a system when a driver is increased.

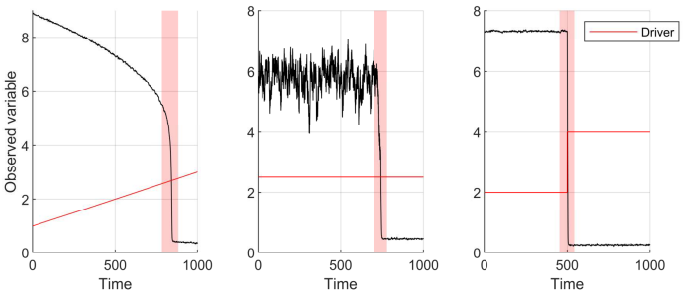


Fig. 4: Different types drivers that can cause a system to tip to another state. The tipping moment is marked by a pink area. From left to right: 1. the driver slowly increases and system tips as a consequence thereof, 2. a noise-induced transition with no change in the driver, 3. a stepwise change in the driver that causes a system to tip.

when a driver or control parameter is slightly changed. Chaotic transient behavior of the system is expected once it reaches such an attractor point [38], [43]. Literature has shown that variance rates can decrease as the threshold region is reached, an indicator that is expected to rise in CSD theory [44]. Crises are not the only case where a regime shift can occur without there being signs of CSD. Indeed, in complex multivariate systems that possess saddle-node bifurcations, CSD might only appear in one of the variables of the system. The importance of identifying the right variable has often been highlighted in literature [8], [15], [38].

B. Drivers

Several mechanisms exist that can push the system into an alternate state. Various types of drivers have been discussed in literature, but the most relevant ones are presented. The most common one in CSD theory is the case where the driver slowly changes, bringing the system closer to a tipping point. By slowly changing, the system tries to find a new equilibrium at each time step, and, as such, exhibits signs of CSD when subject to stochastic disturbances. An example of this case is shown in the left graph of Figure 4.

In the other cases, a regime shift occurs, but signs of CSD are unlikely to be present. These cases correspond to (I) in Figure 2. The first are the so-called noise-induced transitions. The system does not approach or cross a tipping point, but strong external disturbances cause it to shift states (see the middle graph of Figure 4). The second case where it is unlikely to observe CSD, yet a visible shift in the system is observed, is when the external driver changes in a strong stepwise fashion. If the step is large enough to cause the system to move to an alternate state, there should indeed not be signs of CSD. This case is presented in the right of Figure 4.

Finally, other scenarios that can drive the system to tip include stochastic resonance, cyclic transitions, and transients. In the first two, the system can swing between alternate states without crossing a tipping point. These cases also could show signs of slowing down. In the latter, a long transient can cause the system to show a significant nonlinear response without actually having tipped to a different state. Such scenarios are also not expected to show signs of CSD. The system is deviated far away from its equilibrium state and is expected to return to it [19]. These scenarios are, however, not expected to be relevant in applications considered in this paper.

III. Methodology

This paper attempts to make use of CSD theory to uncover possible EWS that could be used as precursors to Loss of Control (LOC). This, in itself, is a challenge. CSD focuses on slow time-varying systems where a slowly increasing driver pushes the system towards a tipping point. A non-complex, low-dimensional, yet highly nonlinear simulation of a robotic system is developed. The influence of the various parameters and the drivers behind EWS - if there are any, can be understood.

A. Simulations

Due to its well understood nonlinear behavior, the inverted pendulum on cart system is used to simulate data containing critical transitions. It is inherently an unstable nonlinear system as the pendulum is kept upright by applying a force on the cart. In this model, a wind force is applied to the pendulum and serves as the control parameter (or the driver). Section III-A1 discusses how the model is set-up, and the different tests performed are elaborated upon in Section III-A2.

1) Model: The free body diagram of the inverted pendulum on cart system is presented in Figure 5 and the equations of motion are given in Equation (4), where m_p is the mass of the pole, m_c is the mass of the cart, and L is the length of the pole. "s" and "c" refer respectively to trigonometric functions sine and cosine.

$$\ddot{\theta} = \frac{(u - F_{wind}\frac{m_c}{m_p})c\theta - m_p L\dot{\theta}^2 s\theta c\theta + (m_c + m_p)gs\theta}{L\left(m_c + m_p s^2\theta\right)} \quad (4)$$

To reduce the complexity of the problem, only two degrees of freedom were considered: $\theta, \dot{\theta}$. Gaussian noise was incorporated into this model to simulate environmental randomness

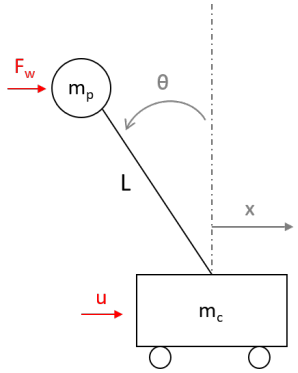


Fig. 5: Free body diagram of the inverted pendulum on cart system with the wind force acting on the pendulum

[16]. Using Equation (1), the stochastic model for the inverted pendulum can be derived as:

$$\frac{d\dot{\theta}}{dt} = f([\dot{\theta}, \theta], [F_{wind}, u]) + \sigma\xi(t) \tag{5}$$

Noise is implemented in a so-called "additive way" as opposed to a "multiplicative one" (such as in [16]). This was done as the assumption that perturbations are independent of the value of the wind force is made. To research the influence of noise color, an Ornstein-Uhlenbeck process was implemented, as suggested in [16]. This process is shown in Equation (6) where $\xi(t)$ is zero-mean Gaussian noise and τ_1 the correlation time.

$$\dot{z} = -\frac{z}{\tau_1} + \frac{1}{\tau_1}\frac{\sigma}{\sqrt{2}}\xi(t) \tag{6}$$

This stationary Gauss-Markov process is usually applied in financial mathematics and physical sciences to model noise [45]. Red-shifted and white noise can both be simulated with this equation by respectively using values of $\tau_1 > 0$ and $\tau_1 \to 0$.

The system overall parameter values are set a shown in Table II. Finally, the simulations are performed in Matlab 2019b using rk4 integration.

TABLE II: Inverted pendulum model and controller constants

Name	Value	Name	Value
g	9.81 m/s^2	K_p	100
L	2m	K_d	40
m_c	1kg	K_{p1}	5
m_p	1kg	K_{d1}	5
$\|noise\|$	0.2	V_{bif}	71.48 m/s

2) Tests: To better understand how CSD manifests itself, three types of failure are simulated:

i. **Test I-A**: The driver is slowly increased. Quasi-steady state behavior is as such simulated at each time step.
ii. **Test I-B**: The driver is increased rapidly. The wind speed is increased such that the pendulum is constantly undergoing a more transient behavior.
iii. **Test I-C**: The driver is slowly increased, but actuators saturate at $\pm 25N$. By setting upper and lower bounds on the control input, the pendulum can tip over without being near a natural tipping point.

In each test, various simulations are performed with different random seeds. Distributions of EWS can be generated and compared using special statistics. One such statistic that is used is the two-sample Kolmogorov-Smirnov (KS) test. This test gives an indication of how different two Cumulative Distribution Functions (CDFs) are.

Two types of controllers are also used in each test: a poorly tuned one (PD controller), and a more powerful one (Nonlinear Dynamic Inversion (NDI) controller). Using the bifurcation analysis software *Matcont*, a saddle-node bifurcation is observed in the closed-loop system (see Figure 6). This implies that the pendulum will always tip at a specific wind force. Signs of CSD should be observed if done so under the right conditions. The PD controlled system is therefore defined as the True Positive (TP) case. The NDI controller, on the other hand, prevents the pendulum from tipping over. No bifurcations are observed (see Figure 6), hence no EWS should be found. The NDI controlled system is defined as the True Negative case (TN).

These concepts are summarized in Table III, where the expected outcomes are presented. Additionally, if no EWS are detected in a simulation with the PD controller, it will be marked as a False Negative (FN) and when an EWS is detected in the NDI system, it will be marked as a False Positive (FP).

TABLE III: EWS outcome as a function of the controller used

		System	
		PD (LOC)	NDI (No LOC)
Classified	EWS	TP	FP
	No EWS	FN	TN

The PD controller is modeled through Equation (7). As only two degrees of freedom are considered, the control loop attempts to keep the pendulum upright by setting θ at a reference of 0 rad (upright).

$$u_{PD} = K_p(0 - \theta) - K_d\dot{\theta} \tag{7}$$

The NDI controller is defined in Equation (8). It is assumed that the nonlinear system can be written as $\dot{x} = f(x) + g(x)u$ and that the control input u follows a control law $k(x)$.

$$u_{NDI} = g^{-1}(\theta, \dot{\theta})\left[K_{p1}(\theta_{ref} - \theta) + K_{d1}(0 - f(\theta)) - f(\dot{\theta})\right] \tag{8}$$

The values of the controller constants are given in Table II. Finally, only the time series data of the angular rate (ω) is analyzed. It can be reasoned that this state variable is of the ones available, the most likely to be sensitive to CSD (the other state variable, the angle, integrates the angular rate hence stochastic effects are less pronounced).

B. Early warning signals

In this section, the method employed to detect EWS is explained. First, the data preparation steps are discussed. Then the various indicators considered are provided. Finally, the definition of an EWS in this paper is given.

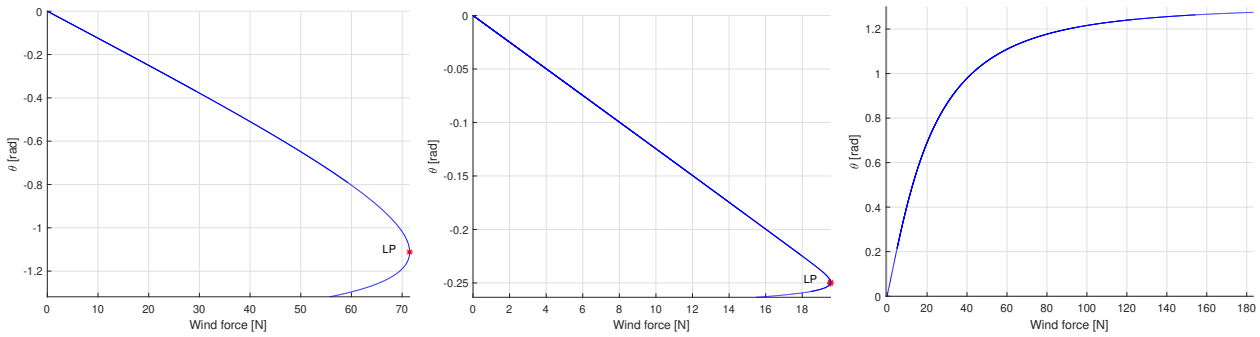


Fig. 6: Bifurcation diagrams of the inverted pendulum on cart system when controlled with different controllers. Left: PD controlled, center: PD controlled with actuators saturating at ± 25N, right: NDI controlled. LP indicates a Limit point, also known as a saddle point.

1) Data preparation: Two elements can usually be observed in a time series:

a. The long term trends of the variable
b. The fluctuations in the residuals.

It is crucial to remove the long term trends in the data to interpret EWS. If not done correctly, the remaining correlations (such as trends or periodicity in the time series) can lead to false indications. Their removal can be done in several ways: through empirical mode decomposition, simple linear detrending (detrended fluctuation analysis), fitting autoregressive models, first-differences, removing running means, Loess smoothing, or by Gaussian smoothing [46]. Trends can be removed both locally through a sliding window, or globally, by looking at the entire time series.

In this thesis, a local Gaussian detrending technique is used as it is most commonly employed in literature [5], [46]–[48]. The kernel can be calculated as:

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{\frac{-x^2}{2\sigma_1^2}} \tag{9}$$

The use of this technique requires defining the bandwidth of the Gaussian kernel, $G(x)$, which is denoted by σ_1. This bandwidth must be selected such that long term trends are removed without filtering out the slow dynamics of the fluctuations. Once the long term trends are established and removed, a residual time series can be obtained. This step is illustrated in the top two graphs of Figure 7. A comparison between the different detrending techniques is not performed in this thesis. However, the effect of detrending the time series with different bandwidths will be looked at.

2) CSD indicators: Various indicators exist in literature. To narrow the scope of the paper, the ones used in [25] are also considered in this paper: the AR1 coefficient, the standard deviation, and the power contained in the low frequencies of the system.

The first indicator is the lag-1 autocorrelation (AR1). It is the most commonly used indicator to measure CSD. An increase in autocorrelation usually refers to an increase in short term memory (correlation at low lags) of the system. An increase in short term memory indicates that the rate of return of a system following a perturbation is reduced. This is best understood by looking at Equation (10), where the assumption is made that a system returns exponentially to an equilibrium position with a recovery rate of $-\kappa$. In this equation, y_t represents the deviation of the state variable of a system from its equilibrium, i.e., $y_t = x_t - \bar{x}$.

$$y_t = e^{-\kappa \Delta t} y_{t-1} + \sigma \xi(t) \tag{10}$$

Consider now a first-order autoregressive process that can be expressed as $y_t = c + \lambda y_{t-1} + \sigma\xi(t)$, where c is a constant. If a zero mean process is considered, then c = 0. Comparing this process with Equation (10), it is visible that $e^{-\kappa\Delta t} = \lambda$, the AR1 coefficient. Theoretically, when a saddle-node bifurcation is approached, κ, the recovery rate of the system converges to zero, meaning that λ approaches a value of one. Transitions where noise plays a role can also have values of λ smaller than one [47].

The second indicator that is used is the Standard Deviation (SD). A slow return rate is usually associated with an increase in SD as the system tends to drift around its stable state. The SD is also the square root of the variance which can itself be determined by subtracting the square of the mean of y_t from the mean of the square of y_t:

$$\text{SD}^2 = \text{var}(y_t) = E\left[y_{t-1}^2\right] - \mu^2 \tag{11}$$

The expectation, μ, of the AR1 process can be found to be zero for c=0. As such the variance of $y_t = \frac{\sigma^2}{1-\lambda^2}$. It was shown that as a system approaches a critical point, $\lambda \to 1$ implying that theoretically, $\text{var}(y_t) \to \infty$.

In a time series, the standard deviation can simply be computed using Equation (12).

$$SD = \sqrt{\frac{\sum_{i=1}^{N} (y_i - \mu)^2}{N-1}} \tag{12}$$

Finally, the low-frequency Power Spectral Density (PSD) is used as the last indicator. The AR1 indicator is oblivious to correlation structure changes at higher lags, which can be captured by the PSD indicator. In a system approaching a transition, a shift of power towards the lower frequencies should theoretically be observed. This shift can be calculated

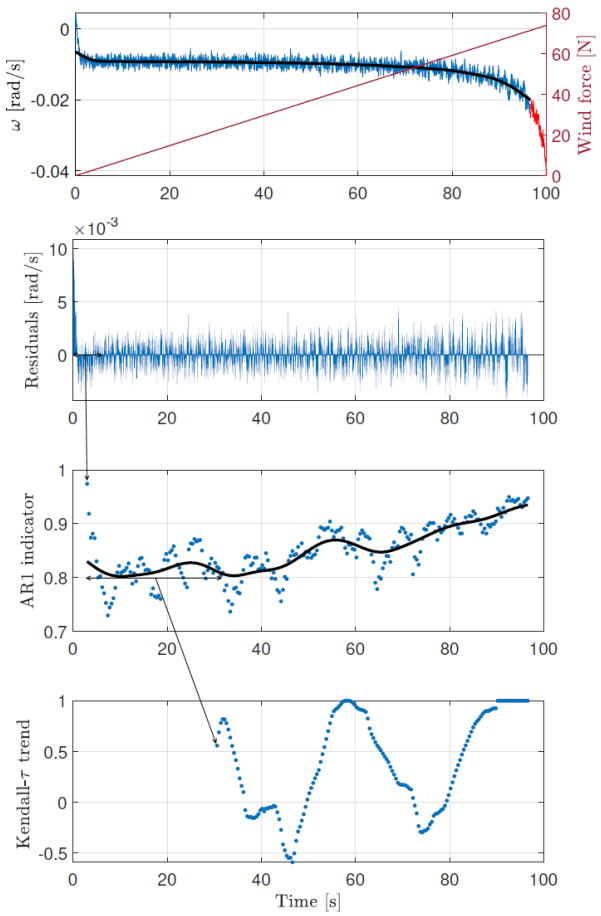


Fig. 7: Rolling window method applied on a residuals time series obtained by detrending the angular rate of one of the simulations of the inverted pendulum. In this example, RWS, the rolling window applied to the residuals, has a size of 500. RWS2, the window applied on the indicator time series has a size of 100.

using the power spectrum obtained from a DFT (Discrete Fourier Transform). The DFT is defined in Equation (13).

$$Y_k \overset{\text{def}}{=} \sum_{n=0}^{N-1} y_n \cdot e^{-2\pi i k n/N} \tag{13}$$

The power is then calculated using Equation (14) and is normalized. The PSD indicator is constituted of the power contained in the first 6% of the PSD sequence [25].

$$PSD = \frac{1}{F_s N}|Y_k|^2 \tag{14}$$

These indicators are locally computed in the residuals time series by sliding a rolling window, RWS, with a rolling step to obtain local indicator values. As a result, a time series of indicator data points is obtained. This step is illustrated in the third graph of Figure 7. As the indicator values can change abruptly due to the step size, the indicator time series is smoothed with a Gaussian kernel of bandwidth 10.

3) Additional indicators: Numerous CSD indicators exist in literature, but three have been selected to scope the project. It is nonetheless still interesting to see if some specific indicators, in combination with the more generic ones, affect the final EWS distribution. The three other indicators considered are skewness, kurtosis, and the cut-off frequency, which contains 50% of the power. The combinations that are tested are presented in Table VI.

Skewness (SKW), the third standardized moment, is an indicator used to detect a phenomenon called flickering. This aspect occurs when a system starts to switch back and forth between alternate states due to significant impacts [49]. This indicator rises depending on whether the current state is larger than the previous one. It can be computed by Equation (15).

$$Skw = \frac{\frac{1}{N}\sum_{i=1}^{n}(y_i - \mu)^3}{\sqrt{\frac{1}{N}\sum_{i=1}^{n}(y_i - \mu)^2}} \tag{15}$$

Another indicator sometimes found in literature is kurtosis (KRT), the fourth standardized moment. Close to a transition, a system is expected to reach critical values. These critical values can be identified through a rise in kurtosis. Kurtosis can be calculated by Equation (16).

$$Krt = \frac{E[y_t - \mu]^4}{\sigma^4} \tag{16}$$

Finally, an alternative method used to measure a shift of power towards the lower frequencies is considered. The cut-off frequency that contains 50% of the power of the spectrum should decrease [9]. This is done by using Equation (13), normalizing it, and computing the cumulative sum of the power at each frequency starting from the smallest one. For each window, the frequency at which the 0.5 mark is exceeded is recorded.

$$\omega_{1/2} = F^{-1}\left(\frac{F(\omega_p)}{\sum_{q=0}^{N-1} F(\omega_q)} = 0.5\right) \tag{17}$$

4) Significance: The significance of these indicators is measured through another rolling window, RWS2, that is slid across the indicator time series with a certain time step. As such, another dataset is obtained from which indicator significance can be interpreted. This is illustrated in the fourth graph of Figure 7. Each datapoint constitutes a potential EWS. Three criteria are considered to assess their significance.

1) Increasing trends;
2) Large endpoints;
3) Concurrence of indicators

First, an EWS corresponds to an increase in the trend of indicator time series values [46]. This increase can be measured using the Kendall rank correlation coefficient, more commonly referred to as Kendall-τ. It is a nonparametric statistical technique used to measure concordance between data points in a list. It is determined by Equation (18). τ is always in the range of [-1, 1]. Large Kendall-τ values suggest a strong trend. The sign indicates if the trend is increasing or decreasing. As highlighted in [16], this method assumes that

data points in the time series are independent, whereas here they are temporally correlated. Nonetheless, as every paper in literature uses this trend estimation method, it is also employed here.

$$\tau = \frac{N_{\text{concordant pairs}} - N_{\text{disconcordant pairs}}}{N(N-1)/2} \tag{18}$$

N is the amount of points in the rolling window. Two points, (x_1, t_1) and (x_2, t_2), are said to be concordant if $t_2 > t_1$ & $x_2 > x_1$. If $x_2 < x_1$, they are said to be disconcordant. If they are equal, they are neither.

Secondly, besides the rising trends, the value of the indicator should also be considered when defining a significant EWS. Indeed, trends can be increasing, yet a low end value of the indicator could suggest that the bifurcation point is far away. As such, an EWS is characterized by a strong enough trend and a large enough endpoint.

To quantify whether both the Kendall-τ and endpoints are strong, their p-value in a historic histogram is computed. For the trends, this histogram is constructed by using the rolling windows on surrogate data that is similar in mean and variance as the residuals. Surrogate trends can then be calculated to populate the historic histogram. This surrogate data can be generated in multiple ways: by using the same Fourier spectrum and amplitude [2], by fitting an ARMA(p,q) model on the residuals [46] or by randomly reshuffling the residuals [25]. The latter option was used due to its simplicity. By reshuffling the residuals 1000 times, to obtain as many samples, and by applying the same rolling windows, trends can be estimated for each indicator using Kendall τ. An example of such a historic histogram is shown for the AR1 coefficient in the left graph of Figure 8.

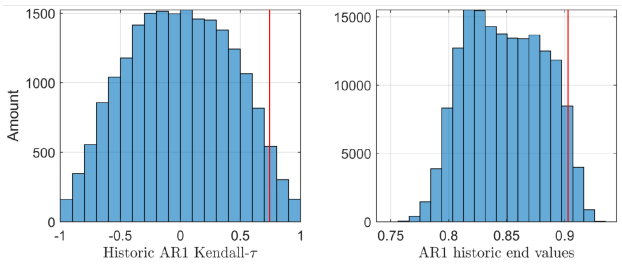


Fig. 8: AR1 indicator significance histograms. Left: historic Kendall-τ values obtained from 1000 simulations, right: historic indicator endpoint values. The red lines are an example of the p-value that defines significance.

For the endpoints, the same method to generate a historic histogram cannot be used as surrogate endpoints are likely to be random and irrelevant. Instead, the original indicator time series is used, and the end value of each rolling window is used. Repeating this procedure across various simulations allows one to construct a histogram of historic indicator end values. An example of the historic histograms for the autocorrelation indicator is shown in the right graph Figure 8.

Finally, an EWS is characterized by concurrent significant indicators. One statistically significant indicator might not be enough to show that an EWS is indeed an EWS and not a false positive. By requiring that all the indicators must be statistically significant, the false positive rate is reduced, and only strong signals are kept.

The values of these various parameters are given in Table IV.

IV. Results

The results of applying the methodology on the inverted pendulum are data are discussed in this section.

A. Test I-A: Slow varying driver

In this first experiment, the wind speed was increased in a continuous monotonic way, allowing the system to return to a quasi-equilibrium state at each time step. The first graph of Figure 7 illustrates this scenario.

When the wind speed is slowly increased, the pendulum approaches a saddle point. With the parameters of the system defined in Table II, this bifurcation occurs at a wind force of 71.48N. The wind force was therefore varied until 74N to ensure the pendulum tips in each run when controlled with the PD controller. The three indicators used to compute EWS are shown in Figure 9. The mean value of the indicator, along with the 95% error margin across 1000 simulations, is shown. In red, the spread of the various indicator values is illustrated; in blue, the same is shown for the trends. As theory stipulates, the AR1 coefficient approaches a value of one. The rise in standard deviation indicates that the system tends to oscillate more around its equilibrium state as the wind force is increased, and the increase in PSD indicates that there is an increase in power contained in the lower frequencies. As the bifurcation point is approached, i.e., at about t=96s, it is also clear that the trends rise. The shaded blue region occasionally reaches a Kendall-τ = 1.

When concurrent trends with histogram p-values of 0.1 are used, EWS are obtained. Doing so across 1000 simulations yields an exponential distribution of the time location of the EWS (see the left graph of Figure 10). As expected, when the tipping point is approached, more EWS are observed.

When the method is applied to the NDI controlled inverted pendulum, the wind force is also increased to 74N. The pendulum does not tip, and no EWS should be observed. However, Figure 10 indicates that several EWS were found in the NDI controlled system. There are significantly less when compared to the PD controlled case, and they seem to occur sporadically throughout the time series.

In total, 7% of runs did not show EWS in the PD controlled case (FN), whereas 16% of the runs in the NDI controlled case did contain EWS (FP).

A two-sample KS-test was performed to compare the PD and the NDI distributions. A KS-stat value of 0.65 was obtained, highlighting the differences between both cases (a KS-stat value close to 1 indicates that distributions differ).

B. Test I-B: Fast changing driver

In this second experiment, the wind speed is also increased until 74N, but the simulation time is reduced from 100s to

TABLE IV: Analysis constants of the model

Parameter	Explanation	Test I-A/I-C	Test I-B
RWS	Number of residual datapoints used to compute the value of one indicator	200	200
step1	The step size of the rolling window for the residual time series	40	40
RWS2	Number of indicator datapoints used to compute one trend value	70	70
step2	The step size of the rolling window for the indicator time series	1	1
T/ dt	The time used to simulate a time series/the time step used for numerical integration	100/0.01	10/1e-3
Bandwidth	The bandwidth used to smooth a target time series and obtain a residuals	200	300
p-value ep/kt	The acceptance thresholds for both the indicator endpoints and the Kendall-τ trends	0.1/0.1	0.1/0.1

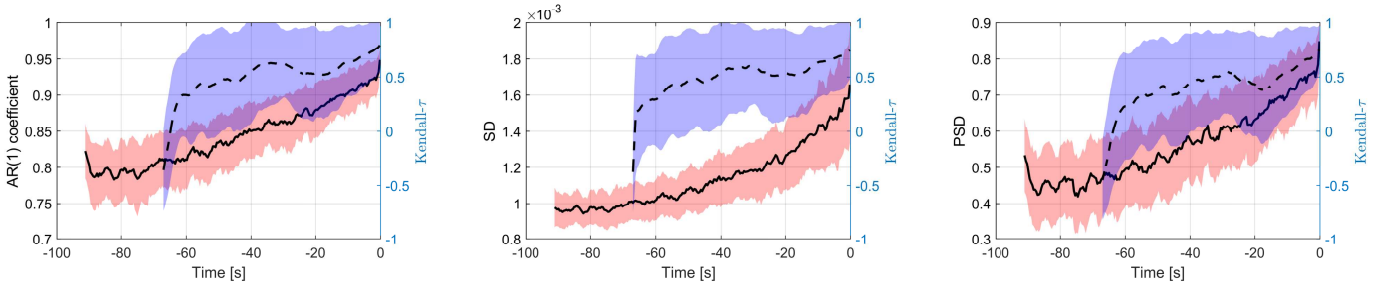


Fig. 9: The three indicators are plotted in a solid black line as a function of the time before bifurcation (left: AR1, middle: SD, right: PSD). The red part represents the spread of values obtained. The dashed line is the average Kendall-τ trend value obtained. In blue, the spread in the various trend values is shown.

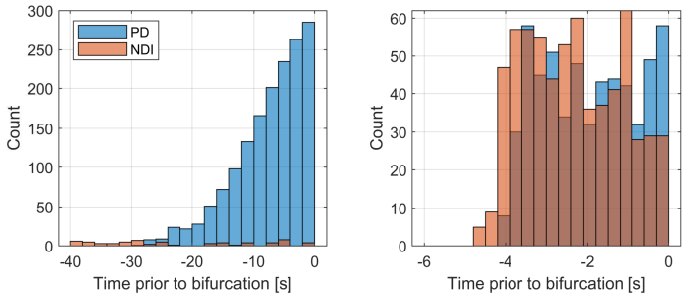


Fig. 10: EWS distribution across 1000 simulations. In the left graph, the EWS for the slowly varying driver are shown (test I-A). In the right graph, the EWS for the fast changing driver are shown (test I-C).

10s. The sampling frequency was increased in order to keep the same number of data points. As the wind speed increases at a faster rate, the system does not return to equilibrium and is constantly subject to quasi-transient behavior. The behavior of the indicators in this model is shown in the appendix, where they are represented for one simulation. Trends are less noticeable, and the indicators are generally noisier when compared with test I-A. Both the SD and the PSD indicators oscillate and have sporadic peaks. This can be due to the large step size.

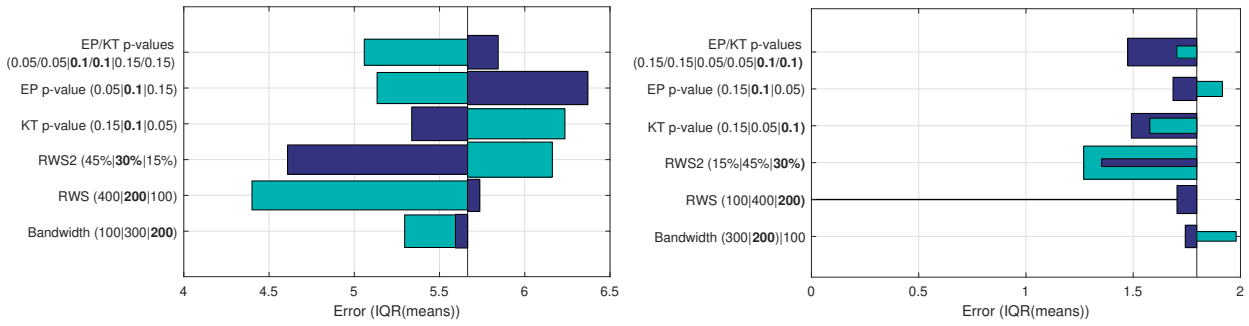
Nonetheless, the locations of concurrent significant indicators are determined, and a distribution of the EWS prior to failure is be obtained. This distribution is illustrated in Figure 10. Instead of observing a particular trend, the EWS are more spread out between five seconds before tipping and the tipping itself. The NDI controlled system also contains significantly more EWS, and the distribution is comparable to the PD controlled case. Again, this can be highlighted by a KS-stat value of 0.13 between both distributions, value which compared to 0.65 in test I-A, is significantly lower. This also means that more false negatives are found: 38% of the simulations with a PD controller did not possess EWS. More false positives are also observed: 61% of the simulations with an NDI controller show EWS.

This can be explained by two reasons. First, more "noisy" information is contained in each window as the sampling frequency is increased. More sporadic EWS are therefore found in the PD controlled case, with some runs not containing any concurrent indicators at all. Second, as the system is not in equilibrium at each time step anymore, it is possible that CSD can no longer be adequately observed in the stochastics. This explains why the noisy sequences in the NDI controller are occasionally identified as EWS.

C. Test I-C: Bounded control system

In this final test, the control input is bounded between $u \in$ [-25, 25]N. In this case, fewer EWS are observed. Indeed, the presence of EWS depends on the distance of the system to its tipping point. Prior to actuator saturation, the closed-loop controlled system behaves normally - just as in the unbounded problem (test I-A). After actuator saturation, however, the system behaves as an open-loop controlled one with different dynamics, hence also a new tipping point. The middle graph of Figure 6 indeed shows a bifurcation at a wind force of 20N.

EWS could be observed in two scenarios. First, if the bifurcation is close enough to the failure, EWS should be found. Second, EWS to this newly induced tipping point can be found if the time between actuator saturation and the moment the pendulum tips is large enough for the methodology to be applied. EWS can thus be observed if the bifurcation point of the closed-loop system is close enough to the moment actuator saturation occurs.



(a) Test I-A (b) Test I-B

Fig. 11: Slow varying driver analysis parameter sensitivity. The x-axis is the independent variable used to measure different spread in EWS distributions when parameters are modified. The baseline "error" is about 6s. In the y-axis, analysis parameters are varied. The bold value corresponds to the ones used in the baseline model. The bars in the chart correspond to the other two values shown on the y-axis. The width of the bar is representative of the amount of TP.

TABLE V: Additional sensitivity information. Across 100 simulations, the average number of EWS in one simulation is given. The number of false negatives and false positives are also provided as well as the KS-stat value comparing the new distribution with the one of the base model.

(a) Test I-A

		PD controller			**NDI controller**	
Parameter		$\bar{\#}$	FN [%]	KS-stat	$\bar{\#}$	FP [%]
Base model	-	**17**	**7**	**-**	**5**	**16**
Bandwidth	100	15	9	0.02	5	17
	400	17	6	0.01	6	14
RWS	100	9	19	0.17	4	15
	400	19	10	0.01	6	13
RWS2	40	16	2	0.06	6	60
	110	23	18	0.05	3	1
KT p-value	0.05	14	16	0.02	5	9
	0.15	20	6	0.01	6	23
EP p-value	0.05	13	13	0.11	5	11
	0.15	21	7	0.08	5	21
EP & KT	0.05	11	27	0.11	4	9
p-value	0.15	24	6	0.08	6	29

(b) Test I-B

		PD controller			**NDI controller**	
Parameter		$\bar{\#}$	FN [%]	KS-stat	$\bar{\#}$	FP [%]
Base model	-	**9**	**38**	**-**	11	**61**
Bandwidth	100	9	44	0.03	10	64
	400	9	44	0.04	10	62
RWS	100	7	99	0.90	6	20
	400	17	24	0.13	14	73
RWS2	40	13	7	0.12	15	92
	110	8	79	0.31	9	26
KT p-value	0.05	7	53	0.02	9	52
	0.15	11	30	0.02	12	70
EP p-value	0.05	8	49	0.06	10	50
	0.15	10	33	0.04	11	70
EP & KT	0.05	7	61	0.07	8	45
p-value	0.15	12	25	0.06	13	77

D. Sensitivity

As the model parameters were set such that reasonable dynamic behavior was observed, and the analysis parameters were set in concordance with values found in literature, it is essential to perform a sensitivity analysis. In order to have comparative results, the independent variable of this analysis is defined as the Inter-Quartile Range (IQR) between the various average EWS time occurrences of the different simulations.

Determining the spread in the EWS mean values of each run with a different random seed allows quantifying the lack of robustness of the EWS. Indeed, a robust method should return similar EWS positions, irrespective of how noise is added in the simulation if all parameters remain unchanged. The spread between the means was measured using the IQR. The IQR in the base model is 6s.

The sensitivity of three classes of parameters can be measured:

1) Analysis parameters,
2) Model parameters,
3) Indicators

1) Analysis parameters sensitivity: The parameters used to calculate EWS are the rolling window sizes, the p-values, and the bandwidth used to detrend the time series. Their effect on the final distribution of the EWS can be measured.

For test I-A, the effect of these various parameters on the independent variable is shown in Figure 11a. Other relevant results are presented in Table Va. Several aspects can be noted. First, from Figure 11a, the effect of the bandwidth on the robustness of results is not significant. The amount of FP and FN also does not vary significantly. Conversely, increasing RWS and RWS2 to capture more information does lead to more robust results. Low KS-stat values are noted, meaning the distributions do not change significantly; however, more FNs are observed for an increase in RWS2. Decreasing RWS2 decreases the FNs but increases the amount of FP significantly.

The acceptance thresholds can also be tuned. As expected, stricter p-values, e.g., 0.05, reduce the average number of EWS and the FPs but increase the amount of FNs. Moreover, the p-value associated with the Kendall-τ has more effect on the FPs and FNs, whereas the endpoint p-value changes the shape of the distribution more. This can be explained by how significance is measured. When a certain run has reduced stochastic effects due to the random seed selected, the significance is still determined in the same historic histograms used for all the runs.

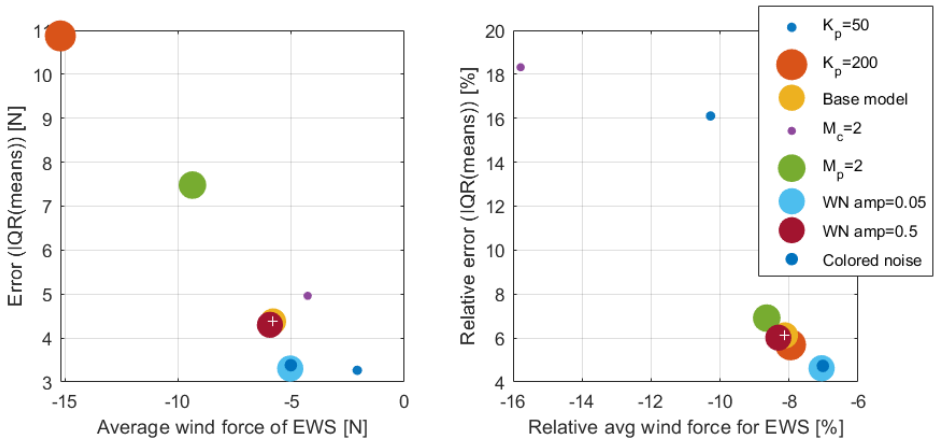


Fig. 12: The effect of the various parameters in the inverted pendulum is shown on the independent variable. The base model is indicated by the white cross. The size of each marker indicates the amount of EWS relative to the base model. As each model has different tipping over wind speed, the wind force prior to the critical one is used in the x-axis of the left graph, and their relative difference is shown on the x-axis of the right graph.

TABLE VI: Combinations of indicators. pvalues for significance are set at 0.1 for both the trend and the endpoints. Results are shown for 100 simulations.

ID	Indicators	FN [%]	FP [%]	# of EWS	Mean [s]	IQR(mean) (s)
C1	All indicators	72	0	66	88.87	9.7
C2	AR1, SD, PSD(6%)	7	16	790	88.75	5.9
C3	AR1, SD, PSD(freq)	7	26	716	88.61	6.7
C4	AR1, SD, KRT, SKW	64	0	104	89.41	8.4
C5	AR1, SD, PSD(6%), SKW	34	3	249	89.46	8.7
C6	AR1, SD	6	31	923	88.76	4.4
C7	AR1, SD, PSD(6%), PSD(freq)	9	15	649	88.55	6.7
C8	AR1, SD, KRT	60	1	138	88.81	7.8
C9	AR1, SD, PSD(freq), SKW	36	7	236	89.61	9.4
C10	AR1, SD, SKW	30	10	311	89.49	7.1

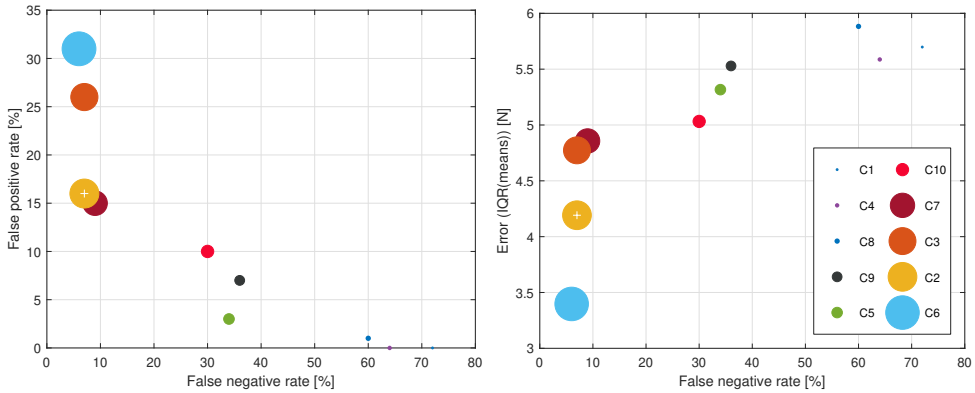


Fig. 13: The robustness of the different combinations of indicators is shown. On the left: the FPs are plotted against the FNs. In the right, the IQR is plotted against the FPs. The base model is indicated by a white cross. The size of the markers is representative of the amount of EWS found by each combination relative to the base model.

The same analysis can be performed for test I-B. The effect of these various parameters on the independent variable is shown in Figure 11b with other important statistics given in Table Va. It can be noticed that some parameters significantly reduce the IQR in between simulations, but that the amount of true positives also greatly reduces. This can be seen when RWS is decreased to 100: the error is zero, but there was only one TP out of the 100 runs.

Overall, it can be seen that the system in test I-B contains multiple FNs and FPs irrespective of the parameter values. A low RWS2 value results in the lowest amount of FNs but also in the highest amount of FPs. Conversely, a low RWS value results in the highest amount of FPs but the lowest amount of FNs. These two parameters are, therefore, fundamental in transient time series analysis. Finally, modifying p-values by 0.05 can result in a variation of up to 20% in the FP or FN amount.

A last aspect to take into account is the interconnection between these different parameters. For the same system, one set of historic histograms is computed. This means, for instance, that when the rolling window size is decreased, acceptance criteria should also be relaxed. If not, larger and smaller indicator fluctuations are harder to qualify as significant.

It can be concluded that the most critical parameters in both cases are the sizes of the rolling windows. The IQR can be lowered by up to 1s in test I-A, and 0.5 seconds in test I-B. The bandwidth has a small influence on the spread of EWS. The threshold criteria have interestingly enough, also a little influence on the spread. Instead, they affect more the false positives/negatives in the system.

2) Model parameter sensitivity: The sensitivity of EWS to model parameters is evaluated in this section. Most of the model parameters were intuitively defined; hence, it could be the case that the system is inherently prone to show results such as the ones obtained. New systems are generated by varying the controller gains, the masses, and the noise added during the simulation.

As each new system (besides the ones with different noise levels) has a new bifurcation point, the spread of the IQR between the average location of EWS as a function of the wind force is used (instead of time, as in the previous section). The new bifurcation positions are shown in Table VII. Recall, that the bifurcation of the nominal system is at a wind force of 71.48N. Figure 12 shows the sensitivity results of the different models. The white cross indicates the base model. The size of each circle corresponds to the amount of EWS found when compared to the base model.

TABLE VII: Bifurcation wind forces for various models and the FNs found

Parameter	Bifurcation [N]	False negatives [%]
Kp = 50	20.28	26
Kp = 200	190.96	6
Mc = 2 kg	27.08	34
Mp = 2 kg	108.3	8

The results show that the type of controller used affects the presence of the EWS. Besides the choice of an NDI controller as the "true negative" case, the type of PD controller used also affects the presence of EWS. Indeed, when the PD controller gains are modified, the natural bifurcation point associated with the critical wind speed that would cause the pendulum to tip, changes. For a similar time interval and similar sampling frequency, the rate of increase of wind speeds is slower or faster. For instance, in the $K_p = 50$ system, the bifurcation point is at 20.28N, which means that in the simulations, the rate of increase is lower when compared to the base model. Fewer EWS are found. This can be seen through the size of the marker in Figure 12. At first, the low error and relative proximity to the tipping point could suggest robustness, yet when observing the right graph of Figure 12, the relative error, and position are large. The lower amount of EWS can be explained by the rate of increase of the driver. Indeed, less information is contained in each rolling window (as the window size is kept the same) as each data point is associated with a wind speed that has not changed. As such fewer signals are detected, and the false negative rate increases from 7% to 26%. Larger rolling windows would be needed to increase the true positive rate. Conversely, in the system with a stronger gain, $Kp = 200$, a drop in the false negative rate is observed. This is explained by the rolling window sizes containing more information about the systems state at different wind speeds. These conclusions are supported by the results found in the previous section, where on a same system, the window sizes were varied. Smaller window sizes also result in more false negatives in both tests I-A,B, as again less information is contained.

The rate of increase of the driver, therefore, also has an influence on the EWS distribution. This was already discussed when both PD gains were modified, but can also be seen when the masses of the system are modified. For a fixed window size, the systems that have slower dynamics (e.g., $K_p = 50$ and $m_c = 2kg$) have a larger spread in EWS distributions across various runs. This is due to the low amount of EWS detected and the high false negative rates. The systems with faster dynamics (e.g., $K_p = 200$ and $m_p = 2kg$) have in absolute numbers more substantial errors, yet when comparing them to the critical wind force, they perform similar to the base model.

Finally, noise within the system should also influence the spread of EWS. Noise does indeed play a central role in CSD theory as, ultimately, the way the system reacts to environmental stochasticity is measured. Colored noise could lead to more autocorrelated samples, whereas the amplitude of a noisy sequence could have effects on the system dynamics. Figure 12 seems to indicate that the amount of EWS found with colored noise is indeed lower. This is because the residuals are overall more correlated from the start meaning that trend increases in indicators are harder to observe. Additionally, the effect of noise amplitude and noise color on the IQR, and the mean location of EWS seems to be minimal. Overall, the conclusions presented in [16] are also visible here. The effect of noise color and amplitude on the EWS distribution of a system containing a saddle-node is small if long term trends are properly removed.

3) Indicator sensitivity: Until now, a combination of the AR(1) indicator, the SD, and PSD was used. It is now interesting to combine these indicators with the ones presented in Section III-B3. The different combinations are shown in Table VI along with their ID.

In Figure 13, the different indicator combinations are plotted. The size of each circle indicates the amount of EWS when compared to the base model (which is itself also indicated by a white cross). In the left graph, the FPs are plotted against the FNs. It can be noticed again that as the number of EWS decreases, more FPs are observed. C2, along with C7, seem to be the optimal combinations. In the right plot, the robustness of the combinations is shown. Once again, for combinations of indicators that yield lower EWS, a more extensive spread in the average position of EWS for each seed is found. As such, C6 is the most robust combination, but also has the largest amount of FPs and the most EWS detected.

Overall, combinations of indicators that yield few false negatives tend to use the more generic indicators, whereas the ones that yield few false positives make use of more specific indicators (skewness & kurtosis). In the left graph of Figure 13, it can be seen that one comes at the cost of the other. The indicator that yields the most consistent distribution across the various runs is the most simple one. This can be explained by the large amount of EWS it detects. The more strict indicator combinations detect fewer EWS (they have a larger false-negative rate), and, from the ones detected, they seem to appear at very sporadic locations as the IQR is quite large.

V. Discussion

The analysis of the inverted pendulum simulation gives insights into the usage of CSD EWS in robotic systems. In test I-A, the EWS are evident. The presence of the saddle-node and the fact that the system remains at a quasi-steady state throughout the simulation suggest that CSD should not only be present but that the frequency of EWS should increase as it approaches the tipping point. The exponential rise of EWS observed prior to the tipping of the pendulum towards its natural equilibrium state verifies that the method does capture CSD in slowly varying robotic systems. To validate the method, an NDI controller that prevents the pendulum from tipping was used. Few sporadic EWS were indeed found in the slow varying NDI controlled system.

The behavior of EWS in the simulated data with the slow varying driver is as expected, but the conclusions that can be drawn from the EWS of the system with a fast varying driver are more ambiguous. EWS are present, but not only in the system that tips over but also in the NDI controlled one. This echoes the doubts found in literature about the applicability of these methods beyond the traditional steady-state analysis. Nonetheless, this does not mean that the EWS found in the PD controlled system of test I-B are wrong. Indeed, the use of an NDI controlled system as a validation case also has its flaws. As the rate of the driver is increased, the performance of any controller is reduced. Stochastic effects become more pronounced, which could lead to concurrent indicators forming EWS.

The robustness of the EWS found is vital to consider when attempting to answer the question as to whether or not they can be used as predictive precursors. Indeed, parameter changes should not affect the outcomes of the results by much. Overall the EWS appear at sporadic locations in the time series - even in the slow varying driver case. The dependence of the method on stochastic system behavior and on the driver that pushes it to the tipping point raises questions about its usage as a robust precursor to LOC. Controllers are being designed to minimize the effect of disturbances, elements that improve system performance, but decrease the potential of CSD indicators. Results seem to indicate that robustness also heavily depends on the values of the parameters used to analyze the model. Larger rolling window sizes are required to reduce the false-negative occurrences and to obtain a more even spread of EWS across various simulations.

Another set of important parameters that should be carefully tuned are the thresholds: the p-values for the trends and the endpoints that define whether a candidate signal can be considered as an EWS. These two thresholds should be chosen such that false negatives and false positives are minimized, with the former being more important. In literature, these parameters are identified by looking at historical values of the indicators, and by identifying the p-value of the current observation. If this p-value is small enough, the signal is considered as important. There are several drawbacks to defining significance in this way. Firstly, doing so requires prior system knowledge, which is not always available. In literature, the rolling window method has always been used on data where a transition already occurred. Obtaining prior knowledge is, therefore, not difficult. Second, obtaining such a historical database can be computationally intensive. The use of, for instance, a fixed threshold value rather than a p-value could reduce computational time at the expense of accuracy.

Additionally, the robustness of EWS is also dependent on the observed variables and the indicators used. Selecting the right variable is of importance as the presence of such signals depends on the ways the fluctuations in the residuals manifest themselves. In the simulation, the analysis was executed on the angular rate as the angle is less sensitive to show signs of CSD. In more complex multivariate systems, a proper selection of the variables to analyze should not be overlooked.

Secondly, the indicators used are also important. The results show that certain combinations of indicators heavily affect the location and existence of EWS in a time series. Fine-tuning each indicator could lead to more true positive cases. However, it could also be the case that the indicators considered are not adequate for the problem at hand. The controller's goal is to minimize disturbances - and kurtosis and skewness attempt to measure more advanced deviations within the residuals. These might occur in an ecological dataset but are less visible in the time series data available. Finally, the second method to measure changes in the power spectrum is a binary one. This makes it harder to interpret trends, and such makes it a less powerful indicator when generating EWS.

Finally, EWS have traditionally been applied in complex systems, such as ecological ones, that have large time spans. In robotic systems, however, the sampling frequency is larger, but

the time span is significantly shorter. When the rolling window methodology is employed, use is made of data points. To obtain a candidate warning signal, datapoints spanning a time of up to 50s could be required, which is considerable in the world of robotics. Time spans can be reduced by decreasing the step in between windows, and by reducing the number of points in a rolling window. The latter option was already seen to cause problems in the simulation as less information is present in these windows (fewer datapoints are used). The former was not investigated as bootstrapping is performed. The amount of time needed to generate an EWS could pose problems for the applicability of such methods in robotic systems, which need much faster algorithms to predict failures. This last point is especially problematic in saturated systems, which will only start to show signs of CSD close to the natural bifurcation point. Two scenarios can be considered:

1) $T_{sat} < T_{bif}$ (eg: damaged quadcopter): LOC due to actuator saturation can only be predicted with CSD EWS if T_{bif} is close enough to T_{sat}. In this case, a saturation predictor (T_{sat} = f(wind speed)) could still be built to determine whether LOC will occur due to saturation or due to the bifurcation.
2) $T_{bif} < T_{sat}$ (eg: aircraft stall): The bifurcation could be predicted with EWS and LOC could be avoided. More knowledge about the tips of tipping points would be needed to make this method fully usable.

In the first option, the fact that actuator saturation occurs can be leveraged to predict LOC. In systems where actuator saturation directly causes the system to fail (such as in the inverted pendulum on cart model), an actuator saturation predictor could be constructed. In cases where actuator saturation occurs without it directly causing the failure, actuator saturation could be used as an indicator - not of CSD, but of a robotic system approaching failure.

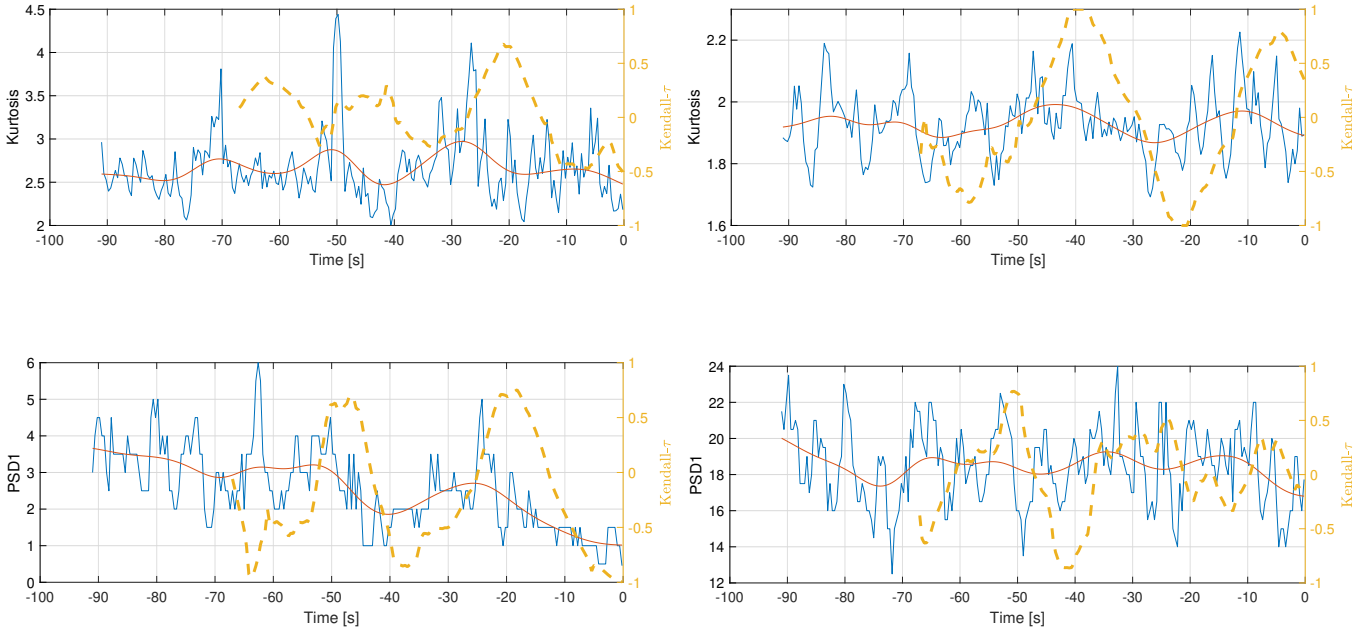
The second option is an interesting one for further research. If tipping does occur under favorable conditions, EWS could be identified and used. However, even in such cases, using them as sole indicators of LOC might not be enough. Instead, another exciting research direction is to make use of bifurcation theory to better understand the path a system takes to reach a LOC condition. The equilibrium paths taken by individual states to reach such a critical point could provide insight on how to best return towards steady conditions. Once these are understood, data-based alarm systems can be implemented, and a control strategy to get out of LOC can be designed. This, for example, is done in [50], where the aircraft deep stall is detected through bifurcation analysis, and a recovery procedure is presented. Equations of motion are, however, used to understand bifurcation behavior - something which is not always available, especially in damaged systems.

VI. Appendix

A. Simulation Indicators

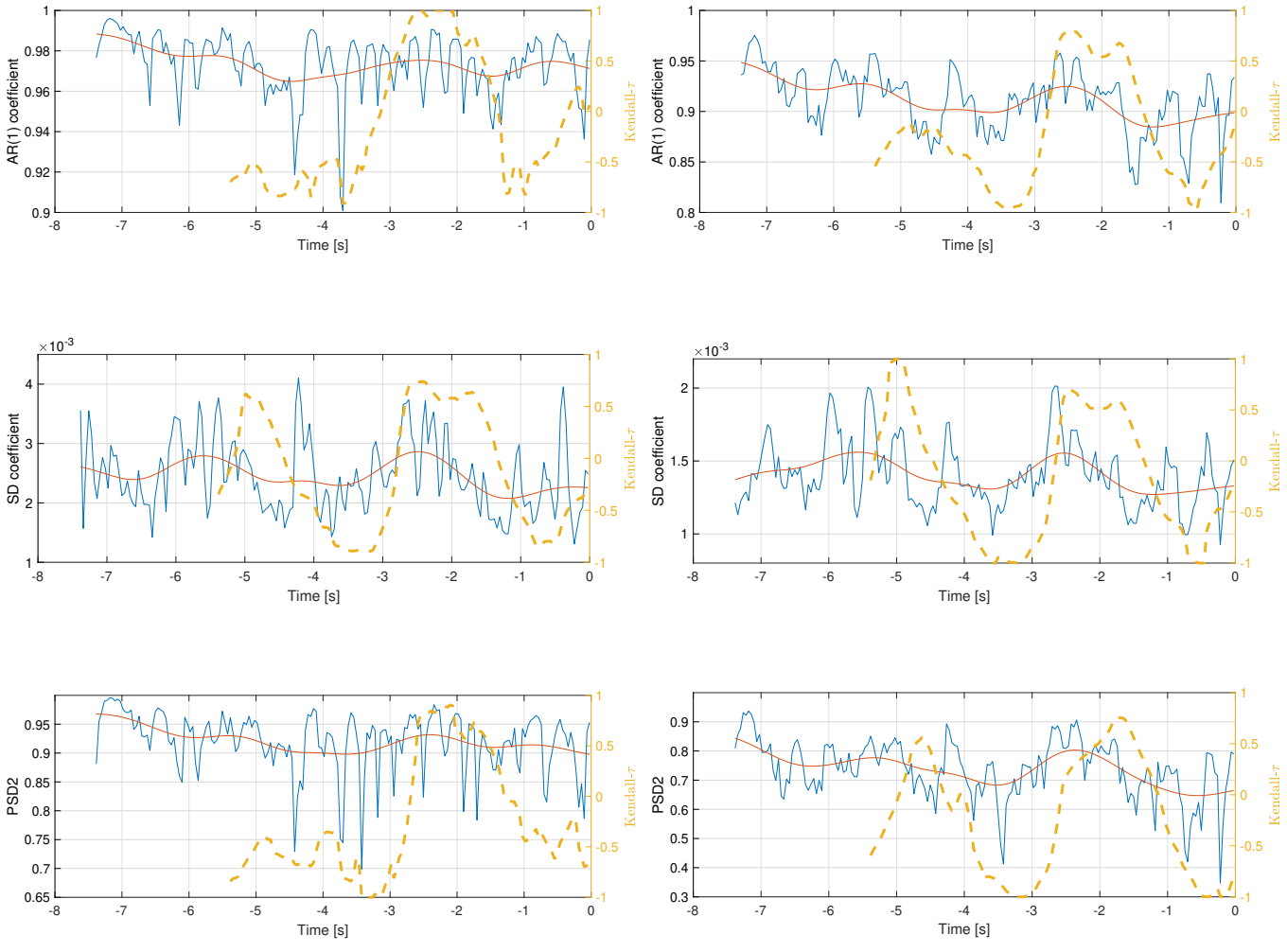
In this appendix various indicators are shown for one random seed.

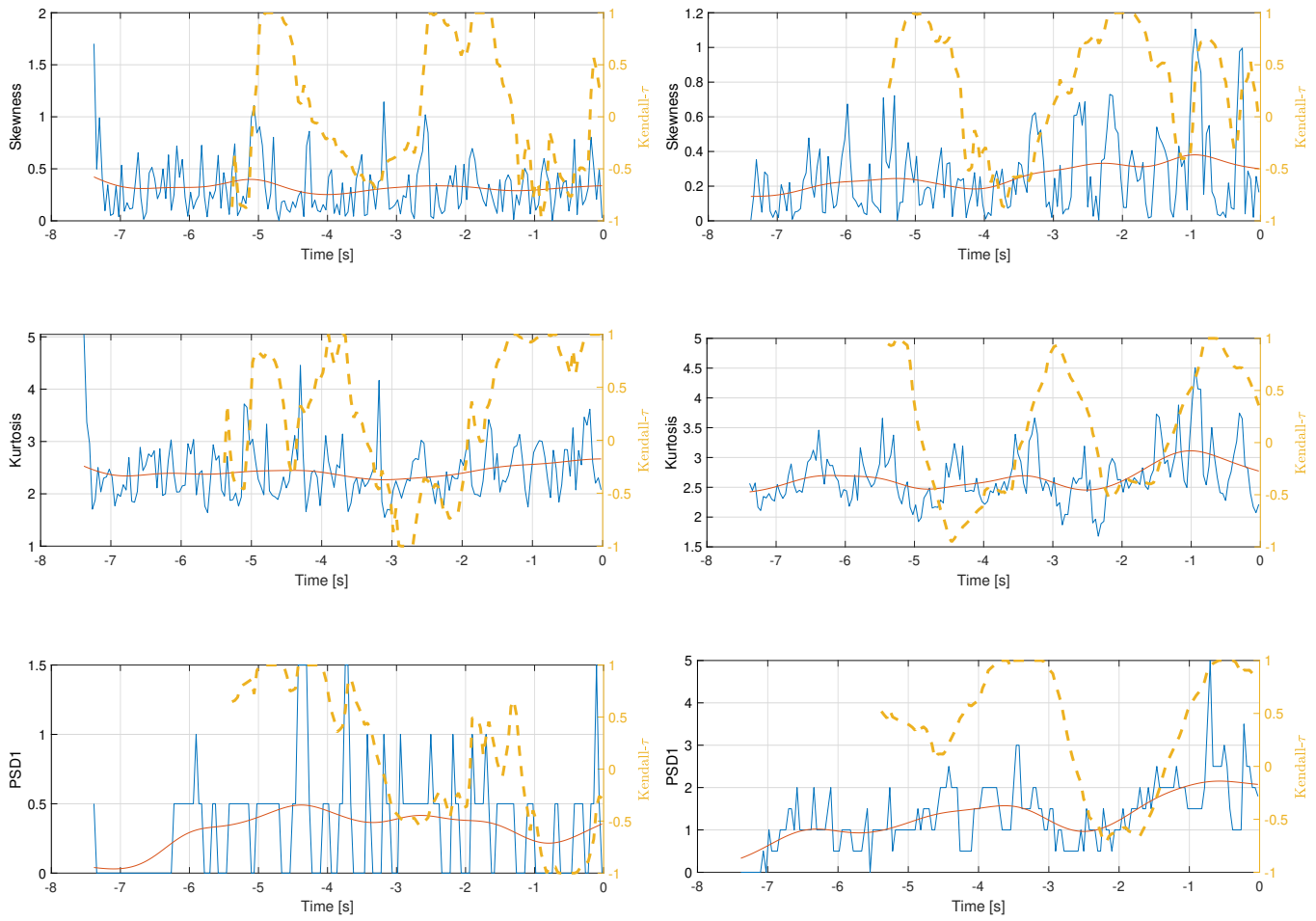
1) Slow varying driver: The indicators shown are constructed for model A. The RWS = 200, and RWS2 = 70. Thresholds were set at 0.1 and 0.1. The bandwidth of the Gaussian kernel used to detrend the time series was set at 200. All the indicators considered in this paper are illustrated for one simulation. The left graphs are representative of the PD controlled pendulum. The right graphs correspond to the NDI controlled pendulum. In each graph, the indicator is plotted, along with its smoothed version. It is with the latter that the Kendall-τ trends are computed. These are indicated in each plot with the dashed yellow line.



2) Fast varying indicators: In this part of the appendix, the indicators are illustrated for system B. The model analysis parameters are a bandwidth of 300, RWS = 200, RWS2=70 and thresholds are set at 0.1 and 0.1 respectively.

PD controlled system **NDI controlled system**





REFERENCES

[1] H. M. Alkhayuon and P. Ashwin, "Rate-induced tipping from periodic attractors: Partial tipping and connecting orbits," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 3, p. 033608, 2018.

[2] V. Dakos, M. Scheffer, E. H. van Nes, V. Brovkin, V. Petoukhov, and H. Held, "Slowing down as an early warning signal for abrupt climate change," *Proceedings of the National Academy of Sciences*, vol. 105, no. 38, pp. 14308–14312, 2008.

[3] J. M. T. Thompson and J. Sieber, "Predicting climate tipping as a noisy bifurcation: a review," *International Journal of Bifurcation and Chaos*, vol. 21, no. 02, pp. 399–423, 2011.

[4] S. Bathiany, M. Scheffer, E. Van Nes, M. Williamson, and T. Lenton, "Abrupt climate change in an oscillating world," *Scientific reports*, vol. 8, no. 1, p. 5040, 2018.

[5] J. Sieber and J. M. T. Thompson, "Nonlinear softening as a predictive precursor to climate tipping," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1962, pp. 1205–1227, 2012.

[6] G. M. Wilkinson, S. R. Carpenter, J. J. Cole, M. L. Pace, R. D. Batt, C. D. Buelo, and J. T. Kurtzweil, "Early warning signals precede cyanobacterial blooms in multiple whole-lake experiments," *Ecological Monographs*, vol. 88, no. 2, pp. 188–203, 2018.

[7] S. Eby, A. Agrawal, S. Majumder, A. P. Dobson, and V. Guttal, "Alternative stable states and spatial indicators of critical slowing down along a spatial gradient in a savanna ecosystem," *Global Ecology and Biogeography*, vol. 26, no. 6, pp. 638–649, 2017.

[8] M. C. Boerlijst, T. Oudman, and A. M. de Roos, "Catastrophic collapse can occur without early warning: examples of silent catastrophes in structured ecological models," *PloS one*, vol. 8, no. 4, p. e62033, 2013.

[9] J. P. L. Tan and S. S. A. Cheong, "Critical slowing down associated with regime shifts in the us housing market," *The European Physical Journal B*, vol. 87, no. 2, p. 38, 2014.

[10] C. Diks, C. Hommes, and J. Wang, "Critical slowing down as an early warning signal for financial crises?," *Empirical Economics*, vol. 57, no. 4, pp. 1201–1228, 2019.

[11] H. Gatfaoui, I. Nagot, and P. De Peretti, "Are critical slowing down indicators useful to detect financial crises?," in *Systemic Risk Tomography*, pp. 73–93, Elsevier, 2017.

[12] C. Tu, P. DOdorico, and S. Suweis, "Critical slowing down associated with critical transition and risk of collapse in cryptocurrency," *arXiv preprint arXiv:1806.08386*, 2018.

[13] S. Orozco-Fuentes, G. Griffiths, A. Baggaley, N. Parker, M. Holmes, R. Ettelaie, and J. Smith, "Early warning signals in plant disease outbreaks," *Ecological modelling*, vol. 393, pp. 12–19, 2019.

[14] T. Quail, A. Shrier, and L. Glass, "Predicting the onset of period-doubling bifurcations in noisy cardiac systems," *Proceedings of the National Academy of Sciences*, vol. 112, no. 30, pp. 9358–9363, 2015.

[15] M. Scheffer, S. R. Carpenter, V. Dakos, and E. H. van Nes, "Generic indicators of ecological resilience: inferring the chance of a critical transition," *Annual Review of Ecology, Evolution, and Systematics*, vol. 46, pp. 145–167, 2015.

[16] P. S. Dutta, Y. Sharma, and K. C. Abbott, "Robustness of early warning signals for catastrophic and non-catastrophic transitions," *Oikos*, vol. 127, no. 9, pp. 1251–1263, 2018.

[17] P. Gloaguen, M.-P. Etienne, and S. Le Corff, "Stochastic differential equation based on a multimodal potential to model movement data in ecology," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 67, no. 3, pp. 599–619, 2018.

[18] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara, "Early-warning signals for critical transitions," *Nature*, vol. 461, no. 7260, p. 53, 2009.

[19] V. Dakos, S. R. Carpenter, E. H. van Nes, and M. Scheffer, "Resilience indicators: prospects and limitations for early warnings of regime shifts," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 370, no. 1659, p. 20130263, 2015.

[20] J. Ma, Y. Xu, Y. Li, R. Tian, and J. Kurths, "Predicting noise-induced critical transitions in bistable systems," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 8, p. 081102, 2019.

[21] M. Titus, Z. Gelbaum, and J. Watson, "Critical speeding up as an early warning signal of regime switching," *arXiv preprint arXiv:1901.08084*, 2019.

[22] Y. Liu, F. Nazarimehr, A. J. M. Khalaf, A. Alsaedi, and T. Hayat, "Detecting bifurcation points in a memristive neuron model," *The European Physical Journal Special Topics*, vol. 228, no. 10, pp. 1943–1950, 2019.

[23] C. F. Clements, M. A. McCarthy, and J. L. Blanchard, "Early warning signals of recovery in complex systems," *Nature communications*, vol. 10, no. 1, p. 1681, 2019.

[24] T. Wilkat, T. Rings, and K. Lehnertz, "No evidence for critical slowing down prior to human epileptic seizures," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 9, p. 091104, 2019.

[25] H. Wen, M. P. Ciamarra, and S. A. Cheong, "How one might miss early warning signals of critical transitions in time series data: A systematic study of two major currency pairs," *PloS one*, vol. 13, no. 3, p. e0191439, 2018.

[26] F. Nazarimehr, S. Jafari, S. M. R. Hashemi Golpayegani, M. Perc, and J. C. Sprott, "Predicting tipping points of dynamical systems during a period-doubling route to chaos," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 7, p. 073102, 2018.

[27] N. Boers, "Early-warning signals for dansgaard-oeschger events in a high-resolution ice core record," *Nature communications*, vol. 9, no. 1, p. 2556, 2018.

[28] J. Prettyman, T. Kuna, and V. Livina, "A novel scaling indicator of early warning signals helps anticipate tropical cyclones," *EPL (Europhysics Letters)*, vol. 121, no. 1, p. 10002, 2018.

[29] F. Romano and C. Kuehn, "Analysis and predictability of tipping points with leading-order nonlinear term," *International Journal of Bifurcation and Chaos*, vol. 28, no. 08, p. 1850103, 2018.

[30] A. Bayani, F. Hadaeghi, S. Jafari, and G. Murray, "Critical slowing down as an early warning of transitions in episodes of bipolar disorder: A simulation study based on a computational model of circadian activity rhythms," *Chronobiology international*, vol. 34, no. 2, pp. 235–245, 2017.

[31] C. T. Bauch, R. Sigdel, J. Pharaon, and M. Anand, "Early warning signals of regime shifts in coupled human–environment systems," *Proceedings of the National Academy of Sciences*, vol. 113, no. 51, pp. 14560–14567, 2016.

[32] E. Gopalakrishnan, Y. Sharma, T. John, P. S. Dutta, and R. Sujith, "Early warning signals for critical transitions in a thermoacoustic system," *Scientific reports*, vol. 6, p. 35310, 2016.

[33] M. A. Litzow and M. E. Hunsicker, "Early warning signals, nonlinearity, and signs of hysteresis in real ecosystems," *Ecosphere*, vol. 7, no. 12, p. e01614, 2016.

[34] P. Ritchie and J. Sieber, "Early-warning indicators for rate-induced tipping," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 26, no. 9, p. 093116, 2016.

[35] A. S. Gsell, U. Scharfenberger, D. Özkundakci, A. Walters, L.-A. Hansson, A. B. Janssen, P. Nõges, P. C. Reid, D. E. Schindler, E. Van Donk, *et al.*, "Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems," *Proceedings of the National Academy of Sciences*, vol. 113, no. 50, pp. E8089–E8095, 2016.

[36] Z. A. Thomas, F. Kwasniok, C. A. Boulton, P. M. Cox, R. Jones, T. Lenton, and C. Turney, "Early warnings and missed alarms for abrupt monsoon transitions," 2015.

[37] V. Guttal, S. Raghavendra, N. Goel, and Q. Hoarau, "Lack of critical slowing down suggests that financial meltdowns are not critical transitions, yet rising variability could signal systemic risk," *PloS one*, vol. 11, no. 1, p. e0144198, 2016.

[38] C. Boettiger, N. Ross, and A. Hastings, "Early warning signals: the charted and uncharted territories," *Theoretical ecology*, vol. 6, no. 3, pp. 255–264, 2013.

[39] P. Ashwin, S. Wieczorek, R. Vitolo, and P. Cox, "Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1962, pp. 1166–1184, 2012.

[40] J. M. T. Thompson, H. Stewart, and Y. Ueda, "Safe, explosive, and dangerous bifurcations in dissipative dynamical systems," *Physical Review E*, vol. 49, no. 2, p. 1019, 1994.

[41] J. M. T. Thompson and J. Sieber, "Climate tipping as a noisy bifurcation: a predictive technique," *IMA Journal of Applied Mathematics*, vol. 76, no. 1, pp. 27–46, 2011.

[42] E. C. Zeeman, "Bifurcation, catastrophe, and turbulence," in *New directions in applied mathematics*, pp. 109–153, Springer, 1982.

[43] C. Grebogi, E. Ott, and J. A. Yorke, "Crises, sudden changes in chaotic attractors, and transient chaos," *Physica D: Nonlinear Phenomena*, vol. 7, no. 1-3, pp. 181–200, 1983.

[44] S. J. Schreiber, "Allee effects, extinctions, and chaotic transients in simple population models," *Theoretical population biology*, vol. 64, no. 2, pp. 201–209, 2003.

[45] C. L. MacLeod, Ž. Ivezić, C. Kochanek, S. Kozłowski, B. Kelly, E. Bullock, A. Kimball, B. Sesar, D. Westman, K. Brooks, *et al.*, "Modeling the time variability of sdss stripe 82 quasars as a damped random walk," *The Astrophysical Journal*, vol. 721, no. 2, p. 1014, 2010.

[46] V. Dakos, S. R. Carpenter, W. A. Brock, A. M. Ellison, V. Guttal, A. R. Ives, S. Kefi, V. Livina, D. A. Seekell, E. H. van Nes, *et al.*, "Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data," *PloS one*, vol. 7, no. 7, p. e41010, 2012.

[47] C. Diks, C. Hommes, and J. Wang, "Critical slowing down as an early warning signal for financial crises?," *Empirical Economics*, pp. 1–28, 2018.

[48] R. Quax, D. Kandhai, and P. M. Sloot, "Information dissipation as an early-warning signal for the lehman brothers collapse in financial time series," *Scientific reports*, vol. 3, p. 1898, 2013.

[49] R. Wang, J. A. Dearing, P. G. Langdon, E. Zhang, X. Yang, V. Dakos, and M. Scheffer, "Flickering gives early warning signals of a critical transition to a eutrophic lake state," *Nature*, vol. 492, no. 7429, p. 419, 2012.

[50] S. Kolb, O. Montagnier, L. Hétru, and T. M. Faure, "Real-time detection of an aircraft deep stall and recovery procedure," *Journal of Guidance, Control, and Dynamics*, vol. 42, no. 5, pp. 1185–1194, 2019.

Critical slowing down and beyond: early warning signals to loss of control in damaged quadcopter flight

A. van der Pluijm
Graduate Student, Department of Control and Simulation
Delft University of Technology, Kluyverweg 1 2629 HS Delft, The Netherlands

S. Sun
Ph.D. student, Department of Control and Simulation
Delft University of Technology, Kluyverweg 1 2629 HS Delft, The Netherlands

C.C. de Visser
Assistant Professor, Department of Control and Simulation
Delft University of Technology, Kluyverweg 1 2629 HS Delft, The Netherlands

***Abstract*—Loss of Control (LOC) is the primary cause of crashes in quadrotors systems. The goal of this paper is to determine if there exist any Early Warning Signals (EWS) in the time series data of a damaged quadcopter that can be used to predict LOC. One possible way that is worth exploring is by using Critical Slowing Down (CSD) indicators. The main idea behind this theory is that a system close to a tipping point needs more time to return to equilibrium after being disturbed. This slowing down can be measured through statistical indicators. To test whether such a system shows signs of CSD, time series of a damaged quadcopter with Single Rotor Failure (SRF) are examined. These time series were obtained from real flight experiments. The analysis was carried out on the time series of the pitch rate and the front rotors. These variables were defined as the most likely to show signs of CSD from all the ones that were deemed important in the LOC classification problem. The latter were identified through a random forest.**

EWS were found in both time series. However, their distribution is not as one would expect. Numerous of EWS are observed early on in the flights, but few remain as the tipping point is approached. When a sensitivity analysis is performed, and signal acceptance conditions are strengthened, the EWS at lower wind speeds remain. Additionally, a significant amount of time is required to quantify whether a signal can be used as a warning. This is a problem as controllers need to detect faults as early as possible to ensure system stability. Nonetheless, these results must still be validated on a proper dataset where the wind speed is increased slowly and monotonically.

Nevertheless, this aspect itself raises questions about the applicability of CSD theory to LOC prediction. The dynamics of a damaged quadcopter are heavily dependent on the wind speed. Meeting the requirements for proper CSD identification is therefore challenging. In an attempt to solve this, an additional indicator was developed that measures the frequency of actuator saturation in the SRF configuration. Indeed, as the tipping point is approached, the rotational rate of the rotors tends to reach its limit. Combined with the CSD indicators, EWS were observed in much more sensible locations. The frequency of EWS increases before failure - notably when the actuator saturation indicator is observed in concordance with standard deviation.

***Index Terms*—Loss of control, quadcopter, CSD, actuator saturation**

I. Introduction

The use of multi-rotor drones has increased significantly over the last years. They are being used in a variety of new fields, including package delivery missions, the film and agriculture industry, and as personal vehicles, to name a few. The method of deployment is also changing. Nowadays, more autonomous platforms are employed, where the drones navigate without there being a human in the loop. A primary concern that comes with such advances is with regard to the safety of these systems.

Loss of Control (LOC) is a failure mode that is currently being researched as it was found to be the primary cause of incidents in both fixed-wing and rotorcrafts [1]. In commercial air transport, LOC directly contributes to 25.8 % of all recorded accidents [2]. In small Unmanned Aircraft Systems (UAS), LOC is responsible for almost 35% of the incidents; more than any of the other causes of failure [3]. With the increase in usage of multi-rotor vehicles, LOC is expected to play a more prominent role in failure incidents making its prevention crucial. For this reason, it is currently being investigated in quadcopters.

Currently, research regarding LOC prevention is directed towards the design of Active Fault Tolerant Controllers (AFTC), whose main objective is to ensure the steady-state performance of the vehicle under both nominal and adverse flight conditions. The two main components of an AFTC are the Fault Detection and Diagnosis (FDD) system and the Fault Tolerant Controller (FTC). The FDD's goal is to detect and isolate the fault in real-time as quickly as possible, whereas the FTC is designed such that stability, desired dynamic and steady-state performance, are achieved automatically. However, most of the research in these fields addresses a specific problem, e.g., rotor or sensor failure. Beyond these specific solutions, there is still no clear way of precisely qualifying LOC of a quadcopter. Recently, the concept of the safe flight envelope has emerged as

a method for quantifying a safe area for dynamics maneuvers. However, such a method relies on knowledge of the physical model of the vehicle. In the case of strongly damaged systems, these equations of motion might not be known.

Data-based methods often do not require knowledge about the physics of the system and could prove to be an excellent alternative to predict LOC in cases where the equations of motion are not known. Numerous methods exist in literature, but one, in particular, has proven to be effective at identifying Early Warning Signals (EWS) in systems that are close to tipping to an alternate state. Use is made of Critical Slowing Down (CSD) theory, where certain patterns can be used to predict an upcoming tipping point. Traditionally, however, this method is used in ecological systems where for example, signs of population extinction are searched for or in climate models, where precursors to climate tipping are identified [4]–[6]. Nowadays, CSD theory is also being used beyond these scenarios and is applied in fast-changing and dynamic environments. Examples of this include the financial sector and the housing market, where crashes are predicted [7]–[12] or in medical data sets where cardiac problems can be predicted [13].

This begs the question as to whether EWS also exist prior to the failure of even more nonlinear systems such as quadcopters. This is where the main focus of this paper lies. The methods used to detect and predict such tipping points in ecological datasets are examined, further developed, and applied to fast-changing, nonlinear dynamic systems such as damaged quadcopters approaching LOC. Results from the CSD indicators on the inverted pendulum on cart system will be used.

The structure of this paper is as follows. First, a brief summary of critical slowing down is provided in Section II. For a more extensive discussion, the reader is referred to the paper on CSD in the inverted pendulum on cart system. Next, the methodology is discussed in Section III. Notable differences with the methodology of the inverted pendulum are Section III-A and Section III-B3. Next, Section IV will present the results, and a discussion will be held in Section V.

II. Critical slowing down

The main principle behind the CSD research rests on the fact that a small disturbance can trigger a substantial change in a real-world system. This occurs when a system reaches a threshold point called a bifurcation or tipping point. When this point is surpassed, an irreversible shift from one equilibrium towards an alternate state occurs [14]. Research has revealed that statistical signals precede these nonlinear transitions. These indicators usually attempt to measure the resilience of a system, i.e., the time it takes for a system to return to an equilibrium state after a disturbance.

Assume that a process $x = x_t : t \leq 0$ can be modeled by a stochastic differential equation that obeys a certain dynamical law $f(\cdot)$ that is subject to perturbations and measurement uncertainties:

$$\mathrm{d}x_t = f(x_t, \vartheta(t))\mathrm{d}t + g(x_t)\mathrm{d}W_t \tag{1}$$

$f(\cdot)$ is also known as the smooth drift function. It describes the deterministic differential equation x_t would obey if $g(\cdot) = 0$. $\vartheta(t)$ is a time-varying input. Noise is modeled by a one-dimensional Brownian motion $W = W_t : t \leq 0$ multiplied by g(·), an unknown diffusion function which determines how stochasticity interacts with the state variable. In such systems, a potential function ∇V_t is assumed to exist, which reflects the regions the system is most likely to travel towards [15]. This potential is assumed to be smooth and slowly varying. When a bifurcation point is approached, the shape of ∇V_t changes around the equilibrium, and the stable point starts turning into an unstable one. During this process, the real part of the eigenvalues of the system also diminishes smoothly to zero. Excursions of x away from the stable point - that can, for example, occur when the system is subject to perturbations, grow both in length and time. This means that the recovery rate is decreasing until it eventually becomes zero when the critical point is reached. At this moment, the equilibrium vanishes, and the system's eigenvalues move into the right half-plane.

However, not all transitions can be predicted by CSD indicators. Indeed, regime shifts can occur without the presence of CSD, and vice versa, not all CSD observations involve regime shifts. Moreover, not all regime shifts include bifurcations, and bifurcations can exist in a system with and without there being signs of CSD [16]. This is illustrated in the Venn diagram in Figure 1.

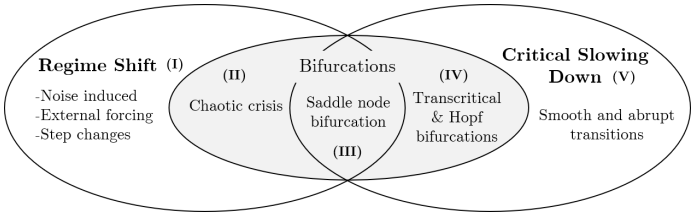


Fig. 1: Venn diagram representing the logical relations between regime shifts, critical slowing down and bifurcations. Adapted from [16].

Early Warning Signals (EWS) usually work best when CSD co-occurs with a regime shift (case (III) in Figure 1). This happens when a system passes a catastrophic bifurcation, a moment at which its equilibrium breaks down, causing it to jump to another state [17]. Non-catastrophic bifurcations (case (IV)), where a system gradually changes towards an alternate state, also can show signs of CSD. However, regime shifts linked to chaotic crises (case (II)) do not necessarily show signs of CSD. Another case where CSD might be missed is in complex multivariate systems that possess saddle-node bifurcations, CSD might only appear in one of the variables of the system. The importance of identifying the right variable has often been highlighted in literature [16], [18], [19].

Additionally, several mechanisms exist that can push the system into an alternate state. The most common one in CSD theory is the case where the driver slowly changes, bringing the system closer to a tipping point. By slowly changing, the system tries to find a new equilibrium and, as such, exhibits signs of CSD when subject to stochastic disturbances. Other ways include noise-induced and rapidly changing drivers.

These cases correspond to (I) in Figure 1. Indeed, in noise-induced transitions, the system does not approach or cross a tipping point, but strong external disturbances cause it to shift states. In stepwise driver changes, the system can exhibit a nonlinear response without tipping or can tip without there being any warning.

III. Methodology

From experiments and video analysis, a damaged quadcopter, suffering from Single Rotor Failure (SRF), undergoes LOC at a forward velocity of around 10m/s [20]. The hypothesis that the system is attracted to a tipping point can be made, as the quadcopter literally does tip over. Instead of jumping to an alternate state, however, it simply enters an uncontrollable motion as stability is lost. In the videos, the quadcopter starts to oscillate more when subject to higher wind speeds. This could be linked to a loss of resilience. Use could, therefore, be made of CSD indicators to quantify this loss of resilience in a bid to generate EWS that can be used to predict failure.

A. Quadcopter data

The SRF quadcopter is a much more complex system than the inverted pendulum. To assess whether CSD occurs in this system, knowledge about the bifurcations is useful. However, the dynamics of such an SRF quadcopter are not fully understood yet, and the models available are usually derived from flight data. This means certain dynamics could have been overlooked. Obtaining a bifurcation plot is therefore complicated, and instead, real-time series flight data will directly be used.

The data used in this paper originates from different sources [20], [21]. The flight experiments were carried out at the TU Delft OJF wind tunnel. Free flights were performed using an off-the-shelf quadrotor (the Parrot Bebop 2) and an open-source autopilot (Paparazzi). The inertial measurement unit runs at 512Hz, and an external motion capture system operates at 360Hz. Test flight where one of the rear rotors was removed will be considered. This means that either the right-back or left-back rotor is removed (respectively rotors 3 and 4 in Figure 2). In the tests performed, the wind speed is increased in a step size fashion. This means that quadcopter behavior is observed for specific wind speeds. Three tests will be performed on this dataset:

i. **Test II-A**: The steady-state behavior is analyzed at each wind speed. One indicator value can be determined per wind speed; hence global trends will be analyzed.
ii. **Test II-B**: The transient behavior is observed by looking at the entire time series. The rolling window methodology is applied here, and EWS distributions are obtained. The results obtained need to be critically assessed as stepwise increments in drivers can themselves be the cause of tipping.
iii. **Test II-C**: The transient behavior is observed by looking at the entire time series. The rolling window methodology is also applied here, but the use of actuator saturation as an indicator of Loss of Control (LOC) will be considered.

Throughout the tests, a balanced dataset will be used: five test flights without the right back rotor (SRF 3), and five others without the left-back one (SRF 4).

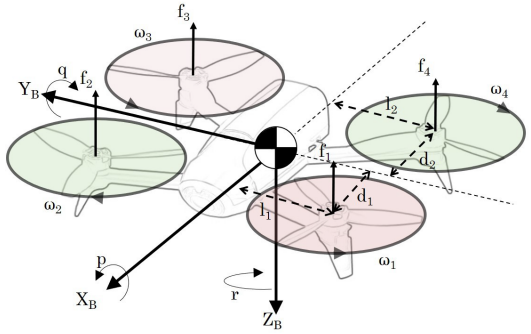


Fig. 2: Parrot Bebop 2 vehicle model in the Body Frame. Rotors 3 and 4 are respectively removed in the SRF 3 and the SRF 4 flights [20].

Finally, the SRF quadcopter has considerably more variables interacting with each other that the inverted pendulum. Selecting the right variables to analyze is crucial, as CSD can be completely missed if the wrong ones are observed. The variables considered are:

1) The six states of the quadcopter ϕ, θ, ψ, p, q, and r;
2) The rotational velocities of the rotors: w1, w2, w4 (If rotor 3 is removed then w3 is not used);
3) The tracking error between the reference command for each rotor and the actual value: $\mathrm{w1}_{err}$, $\mathrm{w2}_{err}$, $\mathrm{w4}_{err}$;
4) A binary feature per rotor indicating whether or not it is saturated;
5) A binary variable indicating how many rotors are saturated.

To narrow the scope, a selection of the most important variables that quantify LOC and that are sensitive to CSD needs to be made. This selection can, in a first step, be achieved by converting the problem as a classification one and using a Random Forest (RF). The RF is known to be a powerful tool when it comes to handling unbalanced data. As for a given flight, the number of time steps associated with LOC are small when compared to the normal ones, a method capable of dealing with such data is required. From all the relevant variables, the two most likely to show signs of CSD will be used in the analysis.

The goal here is to properly classify whether an observation can be quantified as LOC or not. To do so, labels are assigned to various real flight time series. A "1" indicates LOC, whereas a "0" indicates nominal. As the damaged SRF quadcopter eventually fails for a given flight speed, LOC was also defined as a function of the wind speed. The moment failure occurs during a flight, V_{critical}, is identified in the time series (noticeable by a sharp change in data of almost all the variables), and LOC is defined at the wind speed prior to this V_{critical}. Indeed, as the moment LOC occurs is not fully known, its location is varied (e.g., $V_{\text{LOC,1}} = V_{\text{critical}} - 1$ and $V_{\text{LOC,2}} = V_{\text{critical}} - 0.5$). Changes in the variables that quantify the quadcopter's state during LOC are then observed. This can be done using the RF.

To obtain the best results with a RF, weak, and redundant features must be removed. This can be done by removing features that are strongly correlated with each other. The Pearson correlation heatmap of Figure 3 shows the correlation of the independent variables between each other. Variables with a correlation above 0.8 are removed. "Rotor 3" and the "Number of saturated rotors" are strongly correlated; hence one is removed.

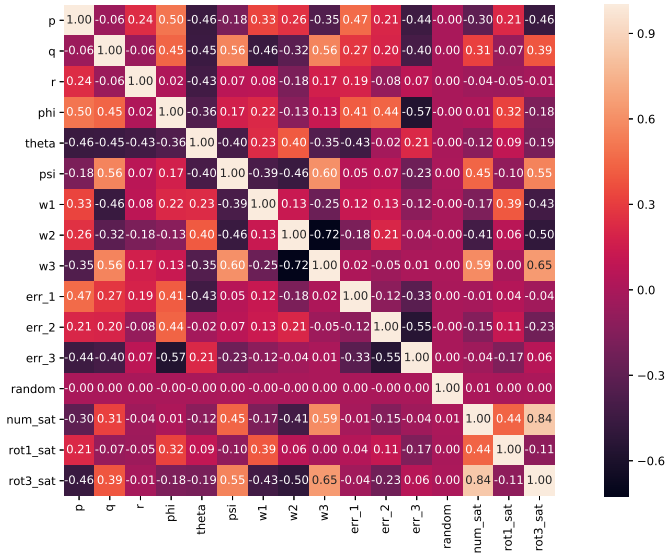


Fig. 3: Correlation matrix of the features used in the RF

The results obtained are validated by applying two other methods: extra trees and gradient boosting classifiers. Gradient boosting is also often used in anomaly detection when there are few samples of the label that is to be predicted [22]. This makes it a good validator of the RF. The extra trees classifier is very similar to the RF except for the way trees are constructed [23]. To avoid overfitting of the data, in each of these trees, the maximum depth was optimized at respectively 10 and 40. The number of trees in the forest was set at 20 for the latter.

B. Early warning signals

In this section, the method employed to detect EWS is elaborated upon. First, the data preparation steps are discussed. Then the various indicators considered are provided. Finally, the definition of an EWS in this paper is given.

1) Data preparation: Long term trends must be removed in the time series to interpret EWS properly. In this thesis, a local Gaussian detrending technique is used as it is the one that is most commonly used in literature [24]–[27]. It was also successfully used in the inverted pendulum tests.

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{\frac{-x^2}{2\sigma_1^2}} \tag{2}$$

The use of this technique requires defining the bandwidth of the Gaussian kernel, $G(x)$, which is denoted by σ_1. This bandwidth must be selected such that long term trends are removed without filtering out the slow dynamics of the fluctuations. Once the long term trends are established, they can be removed from the original time series to create a residual time series.

2) CSD indicators: The three CSD indicators used in this paper are the AR1 coefficient, the standard deviation, and the power contained in the low frequencies of the system. These indicators reflect changes in the properties of a time series generated by the stochastic process presented in Equation (1). This selection was validated in the inverted pendulum system, where multiple indicators were examined.

The first indicator is the lag-1 autocorrelation coefficient (AR1). It is the most commonly used indicator to measure CSD. An increase in autocorrelation usually refers to an increase in short term memory (correlation at low lags) of the system. An increase in short term memory indicates that the rate of return of a system following a perturbation is reduced. Autocorrelation can be approximated through a first-order autoregressive model shown in Equation (3), with $e^{-\kappa\Delta t} = \lambda$, the autoregressive coefficient.

$$y_t = e^{-\kappa\Delta t} y_{t-1} + \sigma\xi_t \tag{3}$$

Theoretically, the recovery rate of the system, $\kappa \to 0$, as the critical point is approached. This implies that $\lambda \to 1$. However, when noise plays a part in the transition, this value can also be smaller than one [24].

The second indicator that is used is the standard deviation (SD). A slow return rate is usually associated with an increase in variance as the system tends to drift around its stable state. Standard deviation is calculated using Equation (4).

$$SD = \sqrt{\frac{\sum_{i=1}^{N} (y_i - \mu)^2}{N-1}} \tag{4}$$

Finally, the low-frequency Power Spectral Density (PSD) is used as the last indicator. The AR1 indicator is oblivious to correlation structure changes at higher lags, which can be captured by the PSD indicator. In a system approaching a transition, a shift of power towards the lower frequencies should theoretically be observed. This can be calculated using the power spectrum obtained from a DFT (Discrete Fourier Transform). The DFT is defined in Equation (5).

$$Y_k \overset{\text{def}}{=} \sum_{n=0}^{N-1} y_n \cdot e^{-2\pi i k n/N} \tag{5}$$

The power is then calculated using Equation (6) and is then normalized. The PSD indicator is constituted of the power contained in the first 6% of the sequence [7].

$$PSD = \frac{1}{F_s N} |Y_k|^2 \tag{6}$$

These indicators are locally computed in the residuals time series by sliding a rolling window with a rolling step to obtain local indicator values.

TABLE I: Analysis constants of the model

Parameter	Explanation	Test II-B	Test II-C
RWS	Number of residual datapoints used to compute the value of one indicator	500	250
step1	The step size of the rolling window for the residual time series	250	10
RWS2	Number of indicator datapoints used to compute one trend value	100	50
step2	The step size of the rolling window for the indicator time series	1	1
Bandwidth	The bandwidth used to smooth a target time series and obtain a residuals	1000	1000
p-value ep/kt(/act sat)	The acceptance thresholds for the indicator end points, the Kendall-τ trends and the actuator saturation acceptance threshold if applicable	0.5/ 0.5	0.8/0.8/0.5

3) Actuator saturation: Numerous CSD indicators exist in literature, but to scope this case-specific topic, only three have been selected. Other CSD indicators were tested on the inverted pendulum on cart system and were seen not to add much additional information when looking at closed-loop systems. In the work performed in [28] - where attempts are made to qualify LOC in a quadcopter, it is concluded that actuator saturation precedes almost every failure case. A recommendation from the inverted pendulum study was to consider case-specific indicators that go beyond the ones suggested in CSD. Actuator saturation will, therefore, be further investigated and developed as an indicator.

By measuring the amount of times a rotor saturates, additional information about the quadcopter transitioning to a LOC situation can be obtained. To do so, saturation must be quantified. The rotors of the Parrot Bebop have a theoretical maximum rotational rate of 12.000 RPM. This value, however, seems never to be reached when observing the time series data of the rotors (e.g., see the time series of w1 in Figure 5). A theory proposed in [20], suggests that actuators of a damaged quadcopter never reach their full limit as they behave as overdamped systems when control over the yaw rate is renounced - which is the case in the experiments considered in this paper. A partial actuator saturation situation must, therefore, be considered. A threshold is defined at which a rotor is considered to be fully saturated. This threshold is fixed at 11.000 RPM, a value which is also suggested in the work of Sijbers [20].

Binary time series are constructed for each rotor. When the rotational velocity exceeds 11.000 RPM, a "1" is assigned. As such, three time series are constructed for the SRF quadcopter (one for each rotor - excluding the one that is removed). The proposed indicator can be calculated with Equation (7), where n is the number of active rotors and $actsat \in \mathbb{Z}_4$.

$$actsat = \sum_{i=1}^{n} z_i \ , \qquad z_i = \begin{cases} 0 & w_i < 11.000 \\ 1 & w_i \geq 11.000 \end{cases} \tag{7}$$

4) Significance: The significance of these indicators is measured through another rolling window that is slid across the indicator time series with a certain time step. As such, another dataset is obtained from which indicator significance can be interpreted from. Each datapoint constitutes a potential EWS. To assess their significance, three criteria are considered.

1) Increasing trends;
2) Large endpoints;
3) Concurrence of indicators

First, an EWS corresponds to an increase in the trend of indicator time series values [27]. This can be measured using the Kendall rank correlation coefficient, more commonly referred to as the Kendall-τ. It is a nonparametric statistical technique to measure concordance between data points in a list. It is given by Equation (8), where N is the number of points in the rolling window. τ is always in the range of [-1, 1]. Large Kendall-τ values suggest a strong trend, whereas the sign indicates if the trend is increasing or decreasing. As highlighted in [14], this method assumes that data points in the time series are independent, whereas, in fact, they are temporally autocorrelated. Nonetheless, as every paper in literature uses this trend estimation method, and as it was successfully implemented in the inverted pendulum analysis, it is also used here.

$$\tau = \frac{N_{\text{concordant pairs}} - N_{\text{disconcordant pairs}}}{N(N-1)/2} \tag{8}$$

Secondly, besides the rising trends, the value of the indicator should also be considered when defining a significant EWS. Indeed, trends can be increasing, yet a low end value of the indicator could suggest that the bifurcation point is far away. As such, an EWS is characterized by a strong enough trend and a large enough endpoint.

To quantify whether the endpoints are "large" enough, their p-value is computed in a historic histogram. This historic histogram is generated by analyzing the time series and observing the end values of each window in order to determine what constitutes "large". This approach is similar to the one presented in the inverted pendulum.

For quantifying the significance of trends, a less time-consuming approach is considered. A simple threshold is defined above, which a trend is considered as "significant". Even though it would be more accurate also to construct a historic histogram (as was done in the inverted pendulum case), the trend values vary between -1 and 1, meaning that a manually defined threshold can be tuned to limit the amount of observed EWS. The values of these various parameters are given in Table I.

Finally, an EWS is characterized by concurrent significant indicators. One statistically significant indicator might not be enough to show that an EWS is indeed an EWS and not a false positive. By requiring that all the indicators must be statistically significant, the false positive rate is reduced, and only strong signals are kept.

An additional comment must be made when actuator saturation is considered as an indicator. As it is a binary one, trends and endpoints are irrelevant. Instead, the mean value

TABLE II: Flight data and global trend results of test II-A. The flight data contains the numbers of wind speed steps in each dataset, the wind speed at which failure occurred at and the SRF configuration. KT = Kendall-τ trends

		idx23	idx24	idx25	idx46	idx48	idx49	idx50	idx51	idx52	idx53	average	
	Number of steps	5	4	4	8	9	7	9	8	6	9		
	Critical wind speed [m/s]	9.4	7.78	7.7	6.3	8.9	7	9	9	6	9.5		
	SRF	4	4	3	4	3	4	4	3	3	3	4	3
q	AR1 KT	1	1	1	0.93	0.56	0.71	0.11	0.57	0.47	0.17	0.75	0.55
	SD KT	1	1	1	1	0.89	0.81	0.44	0.57	0.87	0.33	0.85	0.73
	PSD KT	0.2	0.67	1	0.79	0.22	0.9	0.44	0.57	0.6	0.11	0.6	0.5
w1, w2	AR1 KT	1	1	1	0.79	0.61	0.81	0.39	0.93	0.7	0.78	0.84	0.81
	SD KT	0.8	1	1	0.93	0.94	1	0.89	0.93	1	0.83	0.94	0.94
	PSD KT	1	1	1	0.78	0.5	0.71	0.56	0.79	0.6	0.61	0.8	0.7

of each window is considered. A threshold is defined, which marks when actuator saturation is considered as relevant. The concurrence criterium remains valid for this indicator.

IV. Results

In this section, the results of the tests II-A, B, and C are presented. First, the results of the RF are discussed.

The time series data is divided into a train and test set according to the 80/20 ratio. Bootstrapping was enabled in the RF, and to limit overfitting, the maximum depth and the amount of trees in the forest were both set at 50 and 30. These values were obtained after optimizing the forest on the training sets. The target function that was optimized in this classification problem is the "sensitivity"; which quantifies the amount instances that were labeled correctly as LOC. It is indeed more important to have non-LOC samples mislabeled as LOC, rather than LOC samples mislabeled as non-LOC. Sensitivity can be determined by computing: $\frac{TP}{TP+FN}$.

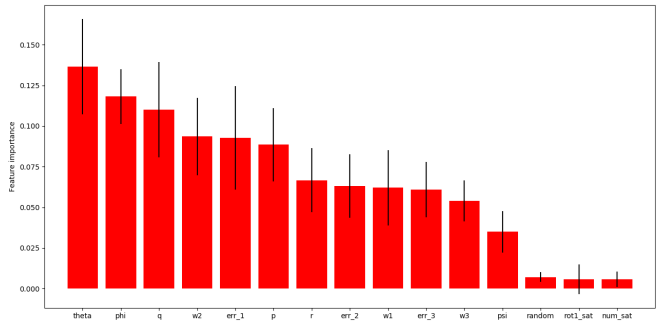


Fig. 4: Feature importance plot of one forest

The important variables can be determined by looking at the feature importance plots of the constructed forest. Figure 4 is an example of the relevant features used by one forest. In the figure, the red bars represent the feature importance and the black ones their inter-tree variability. From the figure, the most relevant variables are ϕ, θ, w2, and q. An additional random variable was added to validate that indeed variables of low significance are not important. It can be seen that the random variable is indeed a non-important feature.

The variables that will be analyzed are the pitch rate (q) and the rotor speed of the rotor in front of the damaged one (w1 or w2 - depending on whether the quadcopter is in SRF 3 or 4). The important angles will not be looked at. Indeed, in the inverted pendulum paper, it was argued that they are less likely to contain EWS. In the appendix, a more in-depth analysis is provided of the essential features, as well as the results of the validation. The relevant information is presented in Appendix A.

Finally, in order to have a visual representation of the time series that will further be analyzed, consider Figure 5. This figure illustrates the time series of q and w1 of flight idx23. The wind speed, on the right y-axis, is increased in steps of two. The top graph shows the pitch rate, whereas the bottom graph shows the angular rate of rotor 1 for a SRF 4 flight. Note that the pitch rate was post-processed with a lowpass of 16Hz to remove noise from rotor vibrations.

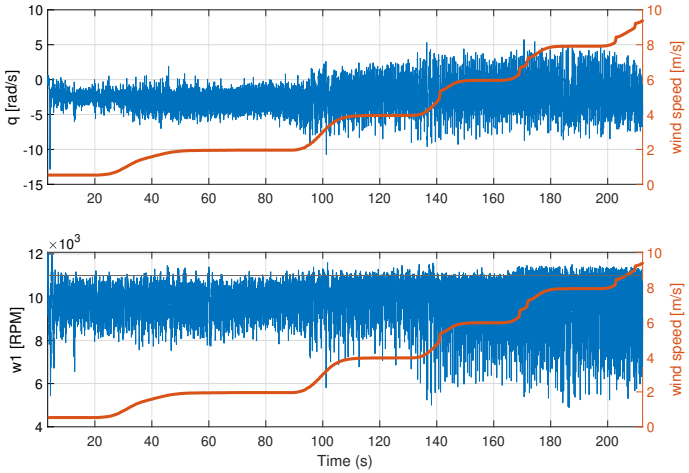


Fig. 5: Time series of a quadcopter flight in an SRF 4 configuration. Top: post-processed pitch rate; bottom: angular rate of rotor 1. Wind speed is plotted in red.

A. Test II-A: Steady-state flight

Steady-state flight behavior can be calculated at every wind speed. Unfortunately, however, the number of wind speed steps at which quadcopter steady-state behavior was observed at is relatively low. Nonetheless, for each wind speed, the indicator values can still be computed. A Kendall-τ trend can then be calculated for all of the indicators, and an indication of whether the general trend is increasing or not can be determined. Locally determining the presence of EWS can not be done as there are not enough indicator data points to perform a rolling window analysis.

TABLE III: Test-II B results. False negatives, average number of EWS, the KS-statistic compared to the base model, the IQR and the mean location of EWS are given.

Parameter		FN		#		KS-stat		IQR [m/s]		mean [m/s]	
		q	w	q	w	q	w	q	w	q	w
base model	-	0	0	63	76	-	-	2.8	3.0	-5.5	-5.5
RWS	250	0	0	116	80	0.21	0.10	2.7	3.1	-6.0	-5.8
RWS2	50	0	0	71	78	0.16	0.12	2.6	3.0	-5.8	-5.5
RWS, RWS2	50, 250	0	0	113	93	0.23	0.2	2.7	2.7	-6.1	-6.1
RWS, RWS2, EP, KT	50, 250, 0.8, 0.7	0	0	25	12	0.40	0.49	2.2	2.3	-6.6	-7.5
RWS, RWS2, EP, KT, step	50, 250, 0.8, 0.7, 10	0	0	283	98	0.39	0.48	2.4	2.0	-6.9	-7.3

Table II gives the global trend results, per indicator, for the two variables considered, for all the flights. Recall that a trend value of one implies rising indicators hence is a sign of CSD. Several observations can be made. First, the weakest indicator for both variables is the power spectral density. The most consistent indicator is the standard deviation followed by the AR1 coefficient. The trends in both the SRF 3 and 4 configurations are significant.

Interestingly, the indicators perform better for SRF 4 than SRF 3. This should not be the case as the quadcopter is expected to behave symmetrically around the z-axis. Nonetheless, this could be attributed to rotor bearing wear or to the fact that the flights in the SRF 3 configuration have, on average more indicator datapoints. They are hence more likely to have a dip in the trends, which in turn can lead to lower Kendall-τ values.

Finally, the rotor speed time series seems to be more sensitive to CSD than the pitch rate. The trends for the various indicators are closer to one. This is interesting, as the pitch rate is a state variable, and in literature, state variables are traditionally observed.

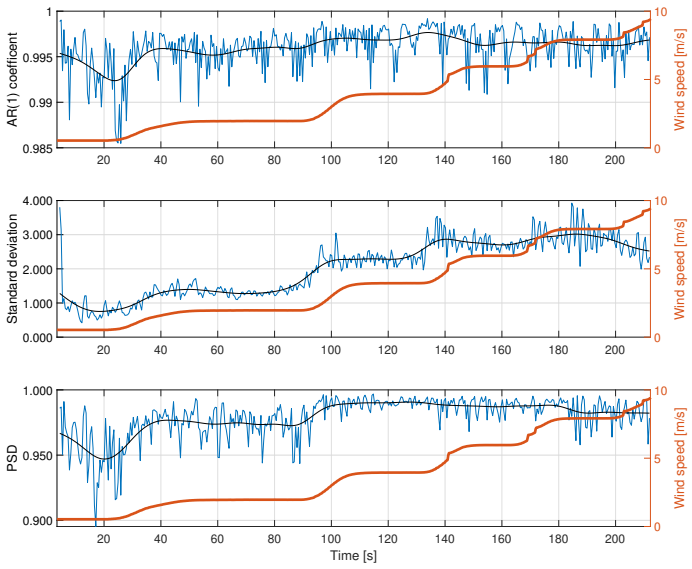


Fig. 6: Computation of the three CSD indicators when looking at the pitch rate time series. The rolling window size is 500 and the step is 250.

B. Test II-B: Transient flight

In this part of the results, the entire time series is analyzed. The values of the analysis parameters are defined in Table I. These values were determined using the results from the inverted pendulum. Large enough window sizes were used in a first attempt in order to ensure a large amount of true positives.

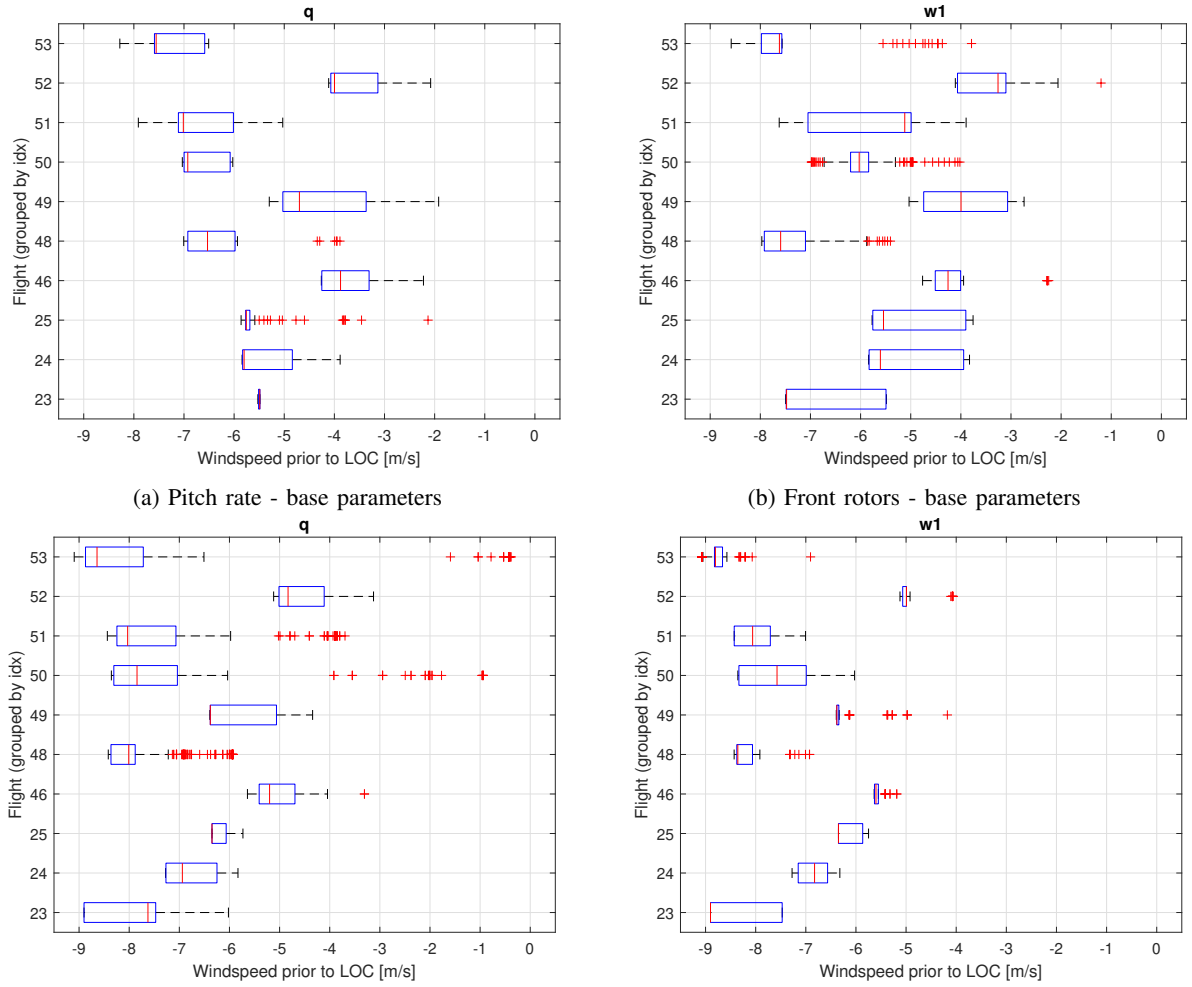
The indicators are shown in Figure 6. The three indicators are plotted when using rolling windows across the time series of the pitch rate. A noticeable aspect is that no direct trend can be observed. The sequence appears to be noisy. In black, a smoothing curve is plotted. Slight increases are observed mostly during transient moments.

Using another rolling window on these indicators results in possible EWS. The location of these signals in each flight is represented in Figure 7a and Figure 7b, when respectively looking at q and the front rotor. EWS concentrations are represented through the boxplots. Outliers, which are indicated by the red cross, simply refer to EWS that were not found near other EWS. The EWS appear to be concentrated towards lower wind speeds. This, in itself, is not the main issue. Instead, the fact that there are no EWS prior to failure is more problematic.

This is also highlighted when observing the second set of results in Figure 11a and Figure 11b. All the EWS across the different flights are aggregated and are plotted with respect to the wind speed they occur at prior to LOC. Again no EWS close to the critical wind speed are found. Additionally, several other interesting elements can be seen in these plots.

First, EWS found when looking at the pitch angle are also shown. The hypothesis that angles contain less EWS and are less sensitive to CSD is validated. Secondly, the EWS found were segregated into different categories. The ones found in SRF 3 are compared with the ones found in the SRF 4 time series (Figure 11a). However, no distinct differences were observed. The same is done for EWS found during a transient moment versus ones identified during steady-state flight (Figure 11b). A transient EWS is defined as one where the last element in a rolling window (RWS2) is situated in a transient moment. Such moments were associated with wind speeds, which were not within 0.1 m/s of the steady-state wind. In this figure, more steady-state EWS are found.

1) Reducing prediction time: A conclusion that was drawn in the inverted pendulum paper was with regard to the applicability of this method in the robotic systems. With the current parameter settings, it takes about 50s to determine if a candidate signal could be an EWS - a number which in ecosystems is small, but in robotics is too large. In this section, the effect of reducing this prediction time will be examined.



(a) Pitch rate - base parameters

(b) Front rotors - base parameters

(c) Pitch rate - reduced parameters

(d) Front rotors - reduced parameters

Fig. 7: EWS found in each fight for test II-B are visually represented through a boxplot. The "box" represents the location of the highest concentration of EWS in the respective flight. The red crosses represent outliers, or in this context, wind speed positions of other EWS. No direct EWS are visible right before LOC occurs (wind speed = 0).

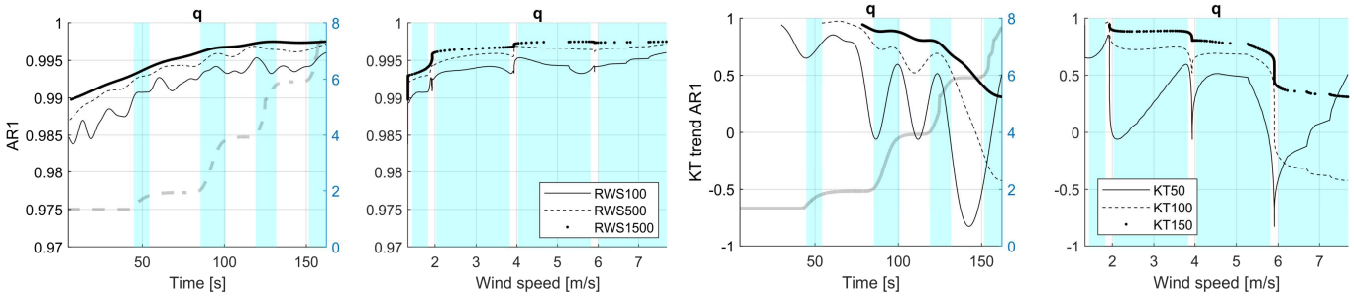
This number depends on the variety of elements, some of which can be adjusted and others which are intrinsic to the system, such as the sampling frequency. The parameters that affect this number are the amount of datapoints in a rolling window and the step sizes. Reducing the step sizes helps to have more local indications - even though on a global scale, the information should not be lost as indicator windows are overlapped. Additionally, the sampling frequency plays a role here. If a low sampling frequency is considered, smaller steps are needed.

Regarding the reduction of datapoints in a rolling window, the results from the inverted pendulum are not optimistic. Reducing the window sizes has effects on the quality of the possible EWS and has been shown to increase the number of false positives. Literature also mentions that window sizes with more datapoints yield better results. Nonetheless, the influence of reducing the required time needed to evaluate a signal will be assessed.

Results are shown in Table III. Combinations are made by lowering the RWS sizes to reduce the amount of required data points. The acceptance thresholds of a candidate signal are strengthened, and the step size can be reduced. The evolution of the overall distribution of EWS is shown in Figures 11c and 11d.

First, the effects of the window sizes are discussed. The numbers presented in the first three rows of Table III do not seem to reveal much difference with the base model. Indeed the KS-stat value remains below 0.25, and the average location of EWS is quasi-identical when compared to the base model. This can be explained by the fact that the rolling windows do not drastically change the locations of the EWS. Some new ones are generated while others disappear. Overall, however, the additional EWS are more concentrated towards the lower end. This can be observed in the left graph of Figure 11c, where RWS and RWS2 are decreased.

The effect of the window sizes can be further understood when analyzing Figure 8. The smaller window sizes were seen to affect the quality of the EWS generated in the inverted pendulum, and here again they skew to the left. In the figure, different window sizes are presented. The shaded areas in the



(a) RWS (b) RWS2

Fig. 8: Different rolling window sizes and their results on the indicator time series, and on the trends when the residuals of q are analyzed. The blue shaded areas represent transients. The trends are generate using an indicator time series of 500.

figure represent transient moments. Two aspects can be noted. First, when observing Figure 8a, the general pattern of the indicator time series is relatively the same across different rolling window sizes. However, when looking at Figure 8b, smaller windows result in more fluctuating trends. Proper tuning of these windows is therefore required - especially as the outcome of RWS2 depends heavily on the result of RWS1. Finally, during the steady-state flight (indicated by the non-shaded areas), the trends always tend to decline. This is because the indicator values should remain constant for a given wind speed; hence no trends should be observed. Again, the poor data structure does not help in extracting useful information with smaller window sizes.

Next, the acceptance p-values can be made stricter in a bid to keep the strongest EWS. Table III shows that the amount of EWS does indeed decrease, yet the average location of these EWS also decreases. This is also visible in the middle graph of Figure 11c, where the p-values of the endpoints and the threshold value are increased. A larger concentration of EWS is obtained. If stricter conditions are posed on the EWS with larger rolling windows, few EWS remain (see the right graph of Figure 11c). The interplay between the different parameters is highlighted here.

Finally, step sizes can be decreased. The step between windows is lowered from 40 to 10 datapoints (i.e., 0.07s to 0.02s). Overall, as expected, more EWS are detected. This is visible when taking a closer look at the middle graph of Figure 11d, and noting the order of magnitude on the y-axis. However, the distribution has not changed much besides it now having tails at wind speeds closer to the tipping point.

Acceptance criteria of the remaining signals can now be toughened to remove weak EWS. Similar behavior as before is observed: the distribution of signals skews to the left (see right graph in Figure 11d). Furthermore, the values in Table III suggest that the distribution is quite different from the base model, but comparable to the one with a larger step. The mean location of EWS is indeed the same in the last two rows: about 7.4 m/s when looking at the rotor time series and about 6.7 m/s when looking at q.

Finally, under the reduced condition, i.e., the last row of Table III, it takes slightly less than 1.5 seconds to compute a signal. The location of the EWS can, as in the base model, be plotted for each flight. This is shown in Figures 7c and 7d for both the pitch rate and the front rotor angular velocity. The shift of the EWS to the lower wind speed when comparing the plots to Figures 7a and 7b is quite noticeable. This is not what is wanted, as more EWS as failure is approached are required.

C. Test II-C: Beyond CSD

In this section, the results of introducing actuator saturation as an indicator of LOC will be presented. After examining the suitability of the indicator, the EWS found when using it on the time series will be shown.

First, the hypothesis that was made in the methodology is assessed: does the amount of times an actuator saturates increase as LOC is approached? The time series of w1 in Figure 5 does seem to indicate this. The frequency of peaks that are above the 11.000 RPM mark indeed increases over time. The same rolling window applied to the residuals of variables is applied to the binary time series that defines how many rotors are saturated. Each indicator datapoint corresponds to

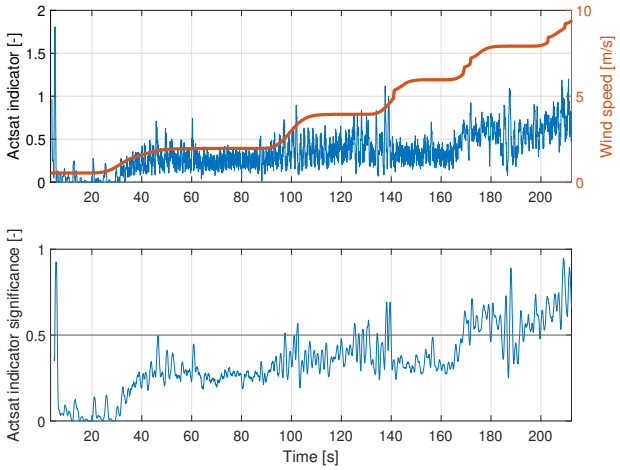
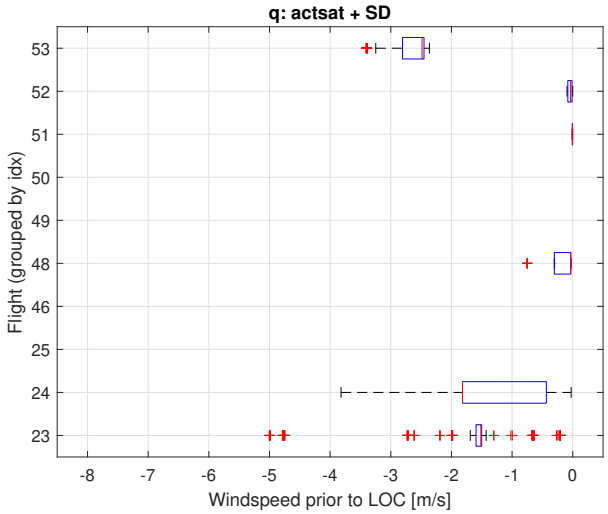


Fig. 9: Actuator saturation indicator. Top: indicator series when RWS is applied, bottom: Significance time series when RWS2 is used. The 0.5 threshold of significance is marked.

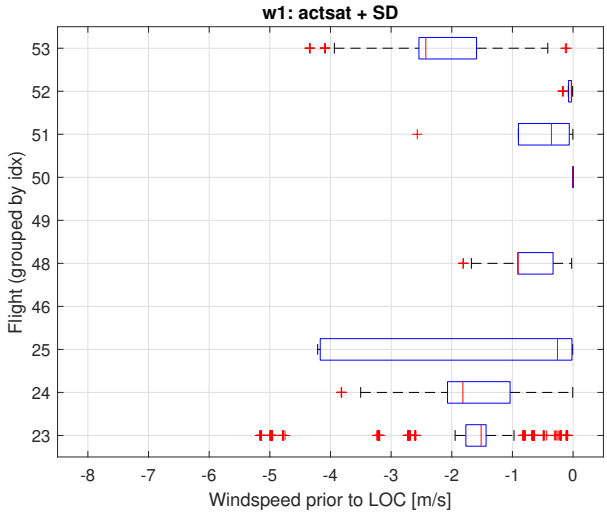
the mean of the contents of the rolling window. As such, a value of the actuator saturation indicator of “1” indicates that at least one actuator was saturated at each time step of the rolling window. Values above 1 indicate that at all time steps,

TABLE IV: Results of test II-C. "-"indicates EWS were found, "x" indicates no EWS were found. The average number of EWS found across all the flights are given along with the IQR and the average location at which they were detected. "sat %" refers to the percentage saturated time steps in the entire time series

Flight idx	23		24		25		46		48		49		50		51		52		53		#		IQR [m/s]		Wind speed [m/s]	
sat [%]	34.3		30.9		16.3		8.2		15.7		0.2		0.4		5.4		3.4		25.1							
	q	w	q	w	q	w	q	w	q	w	q	w	q	w	q	w	q	w	q	w	q	w	q	w	q	w
AR1	-	-	-	-	-	x	x	x	-	-	x	x	x	x	-	-	x	x	-	-	62	93	3.5	1.1	-2.8	-1
SD	-	-	-	-	x	-	x	x	-	-	x	-	x	x	-	-	-	-	-	-	690	794	1.8	1.5	-1.7	-1.4
PSD	-	-	-	-	-	-	x	x	-	x	x	x	x	x	-	x	x	x	x	x	633	111	4.0	2.7	-3.9	-2.6
AR1+SD	-	-	-	-	x	x	x	x	-	x	x	x	x	x	-	-	x	x	-	-	27	40	0.95	1.1	-1.7	-1
PSD+SD	-	-	-	-	x	x	x	x	x	x	x	x	x	x	-	x	x	x	x	x	147	34	1.2	1.3	-1.7	-1.4
All	-	-	-	-	x	x	x	x	x	x	x	x	x	x	-	x	x	x	x	x	23	11	0.3	1.3	-1.9	-1.5



(a) Pitch rate SD and actuator saturation



(b) Front rotor SD and actuator saturation

Fig. 10: Boxplots of the EWS found in the flights when actuator saturation is used in concordance with the SD indicator (test II-C)

the rotors were saturated and that at least at one time step, two rotors saturated. The resulting indicator is shown for a rolling window size of 250 in the top graph of Figure 9.

To determine whether an element is significant or not, the second rolling window is applied. Rather than measuring the trends, the exact same as was done in the first window is done here. Doing so smooths out local peaks due to transients, and locations, where actuator saturation often occurs without there being a local peak, are amplified. This can be seen in Figure 9, where for example, at t=140s and t=190s, actuators saturate often but only the effect at t=190s is amplified.

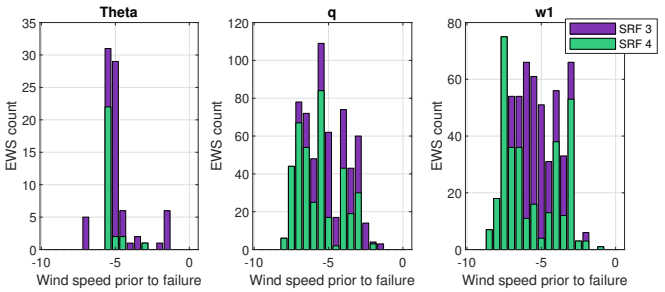
Looking at Figure 9, it can be seen that the general trend is increasing. This means that as the wind speed is increased, actuator saturation does indeed occur more often. Additionally, a quadcopter is more likely to fail during a transient. It is noticeable that peaks are obtained with this indicator at such moments.

Another reason this indicator is interesting is that for wind speeds where actuators often saturate, the probability of them saturating even more increases; unless the wind speed decreases. Making use of the previous times' saturation was observed could, therefore, prove to be useful.

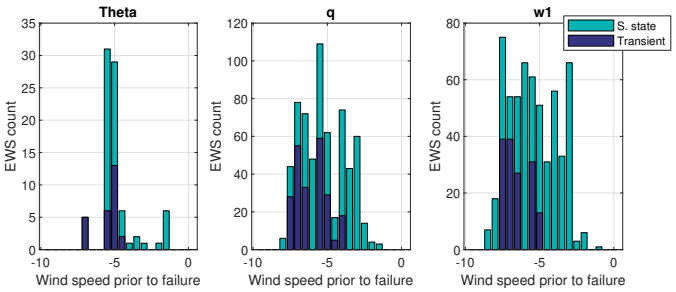
The actuator saturation indicator is tested on the ten flights and is combined with the CSD indicators. The combinations that are tested along with results are shown in Table IV. The values used for the parameters in the analysis are given in Table I. As CSD is not tested for anymore, the values of the reduced parameters that returned poor results in Section IV-B1 are employed.

Different combinations of indicators are tested to compute EWS. They are presented in Table IV along with the results for each flight, in both the q and w time series. Several aspects can already be observed. First, the average location of these EWS is at wind speeds that are closer to tipping. Second, combining the statistical indicators with the rotor speed seems to give the best results. The EWS are closer to LOC than the ones obtained when analyzing q, and their spread is smaller. Finally, certain combinations yield better results than others. For instance, combining the actuator saturation indicator with the SD or PSD results in significantly more EWS than any other combination.

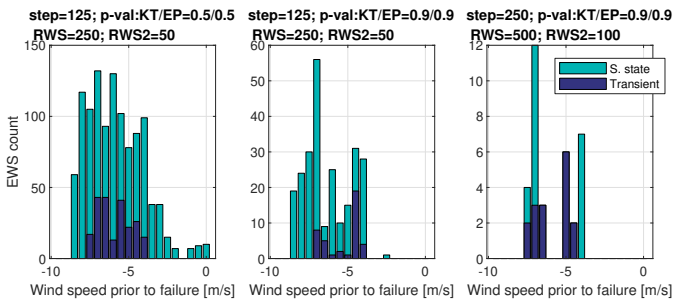
Finally, histograms of the distribution of these EWS are provided in Figure 12. The different combinations are shown when considering both q and w. The combination with the SD indicator seems to be the best. EWS are indeed in large quantities close to the tipping point. For this combination, there are also the least amount of false negatives (marked by "x" in Table IV). The position of the EWS found when



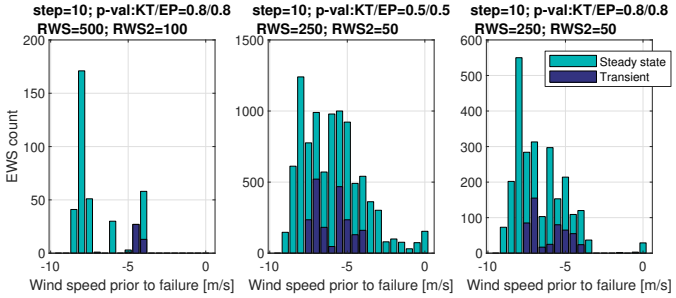
(a) Distinction made between SRF 3 and SRF 4



(b) Distinction made between steady state and transients

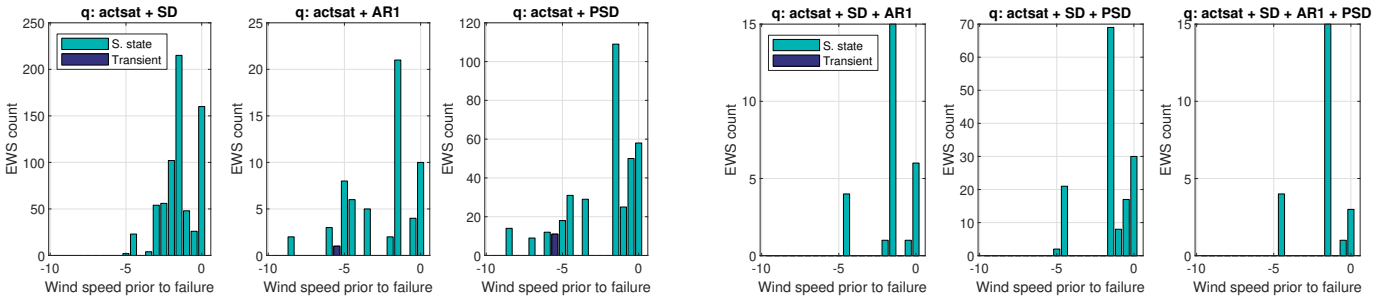


(c) Effect of reducing window sizes and p-values on the EWS distribution in the front rotor time series

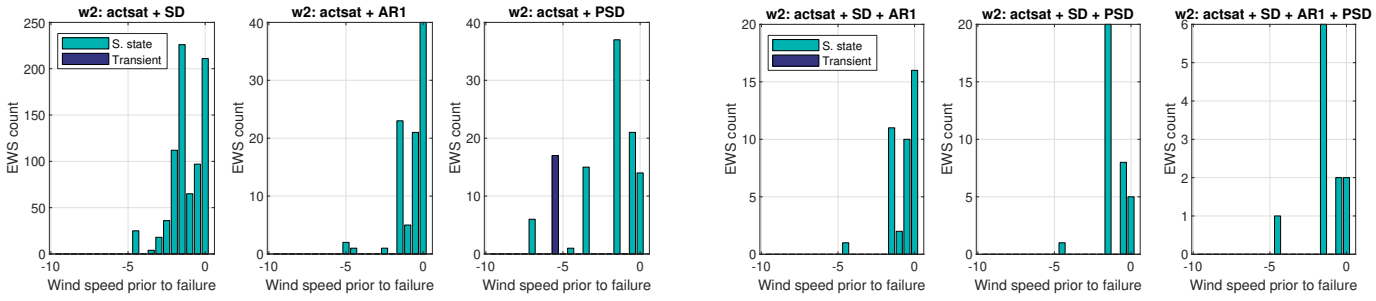


(d) Effect of reducing the step sizes on the EWS distribution in the front rotor time series

Fig. 11: Distribution of the EWS in test II-B. The EWS found across all the flight cases are aggregated and shown as a function of the wind speed. When conditions are made stricter, remaining EWS are concentrated a lower wind speeds.



(a) Using statistical indicators on the pitch rate time series



(b) Using statistical indicators on the front rotor time series

Fig. 12: Distribution of the EWS in test II-C. The EWS found across all the flight cases are aggregated and shown as a function of the wind speed. Different indicator combinations are shown for both the pitch rate and front rotor time series. When multiple indicators are overlapped, few EWS remain.

combining the actuator saturation indicator with the SD is therefore shown in Section IV-C. Again, boxplots are used to represent EWS locations. Unlike test II-B, the EWS are mostly skewed to the right.

V. Discussion

A. CSD in a SRF quadcopter

Signs of CSD in a damaged quadcopter was examined both globally (test II-A) and locally (test II-B). In test II-A, the rising indicator trends in steady-state flight seem to indicate that there is some form of CSD occurring. All of the indicators have an average Kendall-τ trend larger than 0.5 across all ten flights when analyzing both the pitch rate and the front rotor speed. Certain CSD indicators do perform better. For instance, in the majority of the flights, the standard deviation performs better than the low-frequency power spectrum. Additionally, the flights where lower trend values were observed, tend to have more wind speed steps. If during one of those steps, the quadcopter did not remain at equilibrium long enough, then the indicator is also affected. Indeed, fewer data points would be used to compute its value.

A rise in the global trend of the indicator indicates that CSD could be occurring. However, to accurately assess this, more research needs to be done on better quality of data. Slower increments in the wind speed, especially towards the critical wind force, could lead to an improved understanding of the dynamics between the stochastic effects and the quadcopter. Moreover, only two variables, the pitch rate and the rotor speed of the front rotor across the one that failed, were analyzed. These variables were identified through a RF which quantifies LOC by using specified features. A better way to find relevant variables that are susceptible to showing signs of CSD should be considered. Indeed, by eliminating variables that do not play an essential role in quantifying LOC, an essential one that could contain signs of CSD might have been eliminated.

Locally, several signs of CSD were found, but questions marks must be raised as their location in the time series is intriguing. Indeed, Across the ten flights, the average location of EWS was about 5.5 m/s prior to LOC when both pitch rate and front rotors were analyzed. When the window sizes were reduced, in a bid to make the method more sensitive to shorter detection times, the distribution of EWS skews to even lower wind speeds. The appearance of signals so early on, followed by a disappearance of EWS when LOC is approached is not conform to theory. The inverted pendulum research indeed revealed that an exponential distribution should be expected with a slow driver or a more uniform one with transients, yet none seem to occur here. This can be explained by several reasons.

First, the best way to detect signs of CSD is by slowly increasing the driver. The data analyzed, however, contains steps in the wind speeds. Besides being problematic for analysis, CSD literature has signaled that tipping can be caused by a stepwise increase in a driver effectively removing CSD signs. This means that the flights fail due to heavy transients.

Second, the stability of the quadcopter seems to shift around 5.5m/s. This is indeed the moment when the actuator saturation frequency starts to increase. As such, the stability of the system is continuously changing from closed to open-loop for short time bursts. This could affect the quality of the results. When constructing historic histograms, endpoint values in this part can be lower, and trends can be harder to observe, resulting in fewer EWS detected.

Finally, the time required to compute an EWS was lowered from 50s to 1s. This was done by decreasing the values assigned to the analysis parameters. However, the results are mixed. The EWS that remain are concentrated at lower wind speeds. The fact that no EWS occur at larger wind speeds suggests that only using CSD indicators to generate EWS might not be possible. However, to accurately assess whether this is due to actuators saturating, the poor data structure, or due to the system actually not undergoing CSD, more research must be done.

B. Actuator saturation as LOC indicator

From the inverted pendulum paper, a recommendation that was made included developing case-specific indicators such as actuator saturation. Combined with certain CSD indicators and smaller rolling windows, signals in the data can be found that warn of imminent failure. Both the pitch rate and the front rotor were analyzed, and unlike test II-B, the EWS are skewed to the right. Indeed, moments before LOC occurs, most of the flights have EWS that could be used. Even though this seems positive, certain assumptions were made, and significantly more research must be undertaken. ‘ Firstly, not all the flights contained the same amounts of EWS, with some flight showing none (idx46, 49, 50). This can be linked to the frequency of actuator saturation occurring in the entire time series. Indeed, each of these flights was saturated for less than 10%, with the latter two being saturated under 1%. This means that these three flights rarely had RPMs exceed the 11.000 RPM limit that defines saturation. Instead of setting this limit as a general threshold, making it case-specific could solve this problem. However, it could also be possible that not all flights have actuators that reach their maximum. For instance, flight idx 46 fails at 6m/s, wind speed at which other flights have not yet necessarily seen signs of saturation.

Secondly, even though this method goes beyond CSD theory, analyzing the correct variable remains essential. Indeed, statistical shift are still measured, meaning variables prone to show such changes should be considered.

Overall, using actuator saturation along with the statistical indicators of the time series of the front rotors gives more positive results than when the pitch rate was used. Indeed, two more false negatives are detected when the pitch rate is analyzed. The mean position of EWS is closer to the bifurcation point, and the spread between these signals is smaller, indicating that EWS robustness is also improved if the proper variable is analyzed.

C. Research direction

This paper, along with the inverted pendulum analysis, can be used as the foundation for data-based methods to uncover potential EWS to LOC. Based on this, it is recommended

to further research how actuator saturation can be used as an indicator. Another research direction worth investigating is how transients can be properly analyzed. CSD in itself should also be investigated further before fully discarding it.

Indeed, CSD theory is extremely sensitive to the driver in the system. A slow and monotonic increase is required to observe signs that can lead to EWS properly. A damaged quadcopter is often subject to sudden variations in wind speed. Transients and decreasing drivers are likely to occur; hence the suitability of CSD theory is put into question. Secondly, CSD is traditionally applied to slow time-varying systems. These systems are at each time step at a quasi-steady state, which implies that slow recovery rates and shifts in statistical indicators are easier to identify. Finally, EWS work best when a regime shift co-occurs with CSD, i.e., when the system contains a saddle-node bifurcation. As the model behind the damaged quadcopter is not fully known, the existence of a saddle-node bifurcation cannot be confirmed. The quadcopter also does not behave as a system that contains a saddle-node. Rather than jumping to an alternate equilibrium, the system enters a LOC condition, after which it crashes. One could argue that the alternate equilibrium is a failure state, but a more in-depth analysis should be done.

The results presented in this paper are also based on a suboptimal dataset. The experiments should be redone to obtain proper data. The wind speed should be slowly increased, allowing the quadcopter to return to equilibrium at each step. Second, additional indicators that are more suitable for nonlinearities can be looked at. The Lyapunov exponent, for example, has seen an increase in usage in literature [29], [30]. Other indicators such as the permutation entropy [31], [32] are also interesting to consider as they are known to deal well with time series plagued by dynamical or observational noise. Moreover, a more detailed analysis of identifying the variables most likely to show signs of CSD should not be overlooked. The differences in EWS between two of the most important variables in the SRF quadcopter were apparent. Instead, identifying the direction of lowest resilience in a multivariate system can lead to finding the most CSD-sensitive variables.

The actuator saturation indicator has shown that smaller windows can be used, and if the proper variable is observed, reliable EWS can be found. Nonetheless, this indicator must be further developed and further validated if it is to be used for predictive purposes. A sensitivity to the window sizes should be performed to understand better the optimal number of datapoints required in a window. The acceptance thresholds should also be optimized. Finally, different variants of this indicator can be considered. Currently, the sum of three binary time series is used. However, it would be interesting to observe if each time series is used independently.

Beyond the contents presented, other areas that are interesting to further research include rate-induced tipping. "A dynamical system is said to undergo rate-induced tipping when it fails to track its quasi-equilibrium state due to an above-critical-rate change of system parameters" [33]. The tipping does not occur due to noise or bifurcations. Rate-induced tipping is a recent branch of research that focuses on detecting EWS in nonautonomous (time-dependent) and random dynamical systems that do not have a slow changing driver. The same statistical measures are often employed, but with a delay. This means that the signals occur after failure [33].

Ultimately, once robust EWS are found in the time series data, an alarm signal must be designed. This can be done in various ways. For instance, a threshold can be defined, after which an alarm is given. This threshold could consider the cumulative distribution of EWS, the frequency at which they occur, or the co-occurrence of EWS across several key variables. This alarm signal must be capable of predicting failure soon enough for the controller to act accordingly.

APPENDIX

In this appendix, the contents of Appendix A are discussed. This table contains information about the validation of the RF results. The results for the three classifiers are provided in each row. The wind speed that defines LOC is varied in the "wind" column, and the "sensitivity" criterium is quantified. The five most relevant variables are then observed. This is done for several flights, where the SRF3 and SRF4 configurations are examined.

REFERENCES

[1] C. M. Belcastro, J. V. Foster, G. H. Shah, I. M. Gregory, D. E. Cox, D. A. Crider, L. Groff, R. L. Newman, and D. H. Klyde, "Aircraft loss of control problem analysis and research toward a holistic solution," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 4, pp. 733–775, 2017.

[2] G. Wild, J. Murray, and G. Baxter, "Exploring civil drone accidents and incidents to help prevent potential air disasters," *Aerospace*, vol. 3, no. 3, p. 22, 2016.

[3] C. M. Belcastro, R. L. Newman, J. Evans, D. H. Klyde, L. C. Barr, and E. Ancel, "Hazards identification and analysis for unmanned aircraft system operations," in *17th AIAA Aviation Technology, Integration, and Operations Conference*, p. 3269, 2017.

[4] V. Dakos, M. Scheffer, E. H. van Nes, V. Brovkin, V. Petoukhov, and H. Held, "Slowing down as an early warning signal for abrupt climate change," *Proceedings of the National Academy of Sciences*, vol. 105, no. 38, pp. 14308–14312, 2008.

[5] T. M. Lenton, "Early warning of climate tipping points," *Nature Climate Change*, vol. 1, no. 4, p. 201, 2011.

[6] J. M. T. Thompson and J. Sieber, "Predicting climate tipping as a noisy bifurcation: a review," *International Journal of Bifurcation and Chaos*, vol. 21, no. 02, pp. 399–423, 2011.

[7] H. Wen, M. P. Ciamarra, and S. A. Cheong, "How one might miss early warning signals of critical transitions in time series data: A systematic study of two major currency pairs," *PloS one*, vol. 13, no. 3, p. e0191439, 2018.

[8] C. Diks, C. Hommes, and J. Wang, "Critical slowing down as an early warning signal for financial crises?," *Empirical Economics*, vol. 57, no. 4, pp. 1201–1228, 2019.

[9] H. Gatfaoui, I. Nagot, and P. De Peretti, "Are critical slowing down indicators useful to detect financial crises?," in *Systemic Risk Tomography*, pp. 73–93, Elsevier, 2017.

[10] V. Guttal, S. Raghavendra, N. Goel, and Q. Hoarau, "Lack of critical slowing down suggests that financial meltdowns are not critical transitions, yet rising variability could signal systemic risk," *PloS one*, vol. 11, no. 1, p. e0144198, 2016.

[11] C. Tu, P. DOdorico, and S. Suweis, "Critical slowing down associated with critical transition and risk of collapse in cryptocurrency," *arXiv preprint arXiv:1806.08386*, 2018.

[12] J. P. L. Tan and S. S. A. Cheong, "Critical slowing down associated with regime shifts in the us housing market," *The European Physical Journal B*, vol. 87, no. 2, p. 38, 2014.

TABLE V: Random forest validation results

	case 1: RB rotor removed			case 2: LB rotor removed			case 3: LB rotor removed - low speed failure.		
	wind	sensitivity [%]	variables	wind	sensitivity [%]	variables	wind	sensitivity [%]	variables
Random forest	8.5	93.4	θ, ϕ, q, w2, w4	8.5	95.3	ϕ, θ, r, w2, $w1_{err}$	7	93.3	ϕ, θ, q, w2, p
	9	92.6	θ, ϕ, q, w2, w4	9	90.5	ϕ, θ, w2, r, q	7.5	90.4	θ, q, ϕ, $w1_{err}$, w2
Extra trees	8.5	95.0	θ, ϕ, w4, q, w1	8.5	96.7	ϕ, θ, w2,r, q	7	91.5	ϕ, θ, q, w2, p
	9	95.2	θ, ϕ, w4, w1, q	9	93.4	ϕ, w2, θ, q, r	7.5	93.7	θ, q, ϕ, w2, $w1_{err}$
Boosted tree	8.5	96.0	θ, ϕ, w2, w4, q	8.5	95.6	ϕ, θ, r, q, w2	7	94.5	ϕ, θ, q, p, r
	9	96.7	θ, ϕ, w2, q, w4	9	86.8	ϕ, r, q, w1, θ	7.5	91.2	θ,ϕ, q, $w1_{err}$, $w3_{err}$

[13] T. Quail, A. Shrier, and L. Glass, "Predicting the onset of period-doubling bifurcations in noisy cardiac systems," *Proceedings of the National Academy of Sciences*, vol. 112, no. 30, pp. 9358–9363, 2015.

[14] P. S. Dutta, Y. Sharma, and K. C. Abbott, "Robustness of early warning signals for catastrophic and non-catastrophic transitions," *Oikos*, vol. 127, no. 9, pp. 1251–1263, 2018.

[15] P. Gloaguen, M.-P. Etienne, and S. Le Corff, "Stochastic differential equation based on a multimodal potential to model movement data in ecology," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 67, no. 3, pp. 599–619, 2018.

[16] C. Boettiger, N. Ross, and A. Hastings, "Early warning signals: the charted and uncharted territories," *Theoretical ecology*, vol. 6, no. 3, pp. 255–264, 2013.

[17] E. C. Zeeman, "Bifurcation, catastrophe, and turbulence," in *New directions in applied mathematics*, pp. 109–153, Springer, 1982.

[18] M. Scheffer, S. R. Carpenter, V. Dakos, and E. H. van Nes, "Generic indicators of ecological resilience: inferring the chance of a critical transition," *Annual Review of Ecology, Evolution, and Systematics*, vol. 46, pp. 145–167, 2015.

[19] M. C. Boerlijst, T. Oudman, and A. M. de Roos, "Catastrophic collapse can occur without early warning: examples of silent catastrophes in structured ecological models," *PloS one*, vol. 8, no. 4, p. e62033, 2013.

[20] S. Sun, L. Sijbers, X. Wang, and C. de Visser, "High-speed flight of quadrotor despite loss of single rotor," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3201–3207, 2018.

[21] S. Sun, C. C. de Visser, and Q. Chu, "Quadrotor gray-box model identification from high-speed flight data," *Journal of Aircraft*, vol. 56, no. 2, pp. 645–661, 2018.

[22] A. Ravanshad, "Gradient boosting vs random forest," April 2018. Accessed on 22 November 2019.

[23] F. Ceballos, "An intuitive explanation of random forest and extra trees classifiers," July 2019. Accessed on 22 November 2019.

[24] C. Diks, C. Hommes, and J. Wang, "Critical slowing down as an early warning signal for financial crises?," *Empirical Economics*, pp. 1–28, 2018.

[25] J. Sieber and J. M. T. Thompson, "Nonlinear softening as a predictive precursor to climate tipping," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1962, pp. 1205–1227, 2012.

[26] R. Quax, D. Kandhai, and P. M. Sloot, "Information dissipation as an early-warning signal for the lehman brothers collapse in financial time series," *Scientific reports*, vol. 3, p. 1898, 2013.

[27] V. Dakos, S. R. Carpenter, W. A. Brock, A. M. Ellison, V. Guttal, A. R. Ives, S. Kefi, V. Livina, D. A. Seekell, E. H. van Nes, *et al.*, "Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data," *PloS one*, vol. 7, no. 7, p. e41010, 2012.

[28] K. Kersbergen, "Quantifying loss-of-control of quadrotors," 2018.

[29] J. Ma, Y. Xu, Y. Li, R. Tian, and J. Kurths, "Predicting noise-induced critical transitions in bistable systems," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 8, p. 081102, 2019.

[30] Y. Liu, F. Nazarimehr, A. J. M. Khalaf, A. Alsaedi, and T. Hayat, "Detecting bifurcation points in a memristive neuron model," *The European Physical Journal Special Topics*, vol. 228, no. 10, pp. 1943–1950, 2019.

[31] G. Schiepek and G. Strunk, "The identification of critical fluctuations and phase transitions in short term and coarse-grained time series—a method for the real-time monitoring of human change processes," *Biological cybernetics*, vol. 102, no. 3, pp. 197–207, 2010.

[32] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical review letters*, vol. 88, no. 17, p. 174102, 2002.

[33] P. Ritchie and J. Sieber, "Early-warning indicators for rate-induced tipping," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 26, no. 9, p. 093116, 2016.

Part IV

Discussion

Chapter 6

Discussion

In this chapter, the conclusions will be provided as answers to the research questions. Next, recommendations are listed.

6.1. Conclusions

The objective of this work was to determine if loss of control of a damaged quadcopter can be predicted using data-based methods. After a thorough literature study was conducted to understand which data-based methods could best be used for such a problem, CSD applications emerged as the most interesting option to pursue. This led to the development of the following main research question:

"Do the time series of a quadcopter suffering of single rotor failure contain early warning signals that can be used to predict loss of control?"

A quick answer to this question is given: EWS found by solely applying CSD theory on the quadcopter system might on their not be capable of predicting LOC. The necessity of quasi-stationarity for proper EWS identification sets unrealistic constraints on how the wind speed should behave. However, when looking beyond traditional CSD indicators and using the actuator saturation frequency, more reliable EWS are generated. Understanding how to best use these signals as robust LOC predictors is the next research step.

This conclusion is generated after researching three sub-questions that were defined at the start of the thesis. The next sections will discuss each of these questions.

6.1.1. CSD theory for EWS generation

CSD indicators are traditionally applied to slow time-varying systems, such as ecosystems, in a bid to find signals that can be used as warnings to an upcoming disaster. A better understanding of how these indicators generate an EWS is required if the ultimate goal is to apply them to quadcopter time series.

To test this, a literature review on the critical slowing down topic was undertaken, and the most relevant indicators were identified. These include the autocorrelation coefficient, the standard deviation, the low power frequency spectrum, skewness, kurtosis, and the cut-off frequency at which 50% of the power of the signal is contained. These six indicators were identified in a first stage as the most used ones for early warning signals analysis in both financial and ecological datasets. Different combinations were tested and applied to the simplified inverted pendulum problem. A first conclusion that was drawn from this experiment is that not all indicators are equally sensitive to show signs of CSD. Whereas some indicators are more sensitive to outliers, others are more sensitive to the underlying shifts in the power spectrum. This is problematic for closed-loop controlled systems. Indeed a robust controller attempts to minimize the effect of disturbances on the system. Therefore indicators such as standard deviation, skewness, and kurtosis, that quantify the magnitude of a disturbance on the system must be used with caution. This was indeed shown in the inverted

pendulum analysis, where usage of the latter two resulted in more false negatives. A system controlled by a more robust controller is also less likely to contain EWS if analyzed by the wrong indicators. The AR1, SD, and PSD were, as such, found to be the most optimal combination of indicators to generate EWS in the inverted pendulum tests. Other indicator combinations either resulted in a large number of false negatives barely picking up any EWS.

Besides the combination of indicators, it is essential to analyze the right variable. In the inverted pendulum tests, this problem is trivial as there are a total of four possible variables to analyze, but becomes a necessity in the multivariate quadcopter. A random forest was constructed to identify variables closely linked to LOC. The pitch and roll angles, the pitch rate, and the rotor speed of the front-facing rotor directly in front of the damaged one were found to be essential in the LOC classification problem. This, however, does not mean that their time series are likely to show signs of CSD. A variable that contains responses to the disturbances in the time series should be used. Angular rates or acceleration are, therefore, more likely to be sensitive to CSD. It was indeed found that the amount of EWS found in the time series of the pitch angle is a lot lower than the amount found when looking at the pitch rate or the rotor speed.

The first sub-question is as such answered. Unlike the ecosystems or financial structure on which CSD theory has been applied to, selecting the right set of indicators and choosing the proper variable to analyze, are of utmost importance in the quadcopter. The next part which can be discussed is the purpose of these EWS. The initial idea was to use EWS to predict LOC. This means that the moment a system enters a LOC condition must be defined.

6.1.2. Defining LOC in a time series

LOC is still a relatively new concept. It is, until now, characterized through the SFE - zone, which marks the safe space for dynamic maneuvers. Section 6.1.2 indicates where LOC due to the SFE occurs on the time series of the angular rate of the inverted pendulum. The envelope calculated in Figure 5.2 was used to define the safe states. The SFE is highly dependent on the time horizon. For larger time horizons, a larger safe space is obtained, implying that more extreme parts of the safe space can be reached. Currently, this parameter is manually defined. In this analysis, however, no direct link between this time horizon and the actual moment the pendulum starts to tip was found. It was concluded that using EWS to predict LOC when the SFE is exceeded is currently not possible.

The second idea was to define LOC as the moment actuator saturation occurs. In systems such as the inverted pendulum where saturation directly leads to tipping, it is possible to define LOC as such. This moment is also shown in Section 6.1.2. However, in systems such as quadcopters, actuator saturation is not always directly linked to tipping. It can occur as a response to an external disturbance without it actually crashing. Therefore, even though saturation almost always occurs prior to LOC, defining LOC as the moment the control input saturates was also not deemed as accurate.

Thirdly, LOC can be defined as the moment a system approaches a bifurcation point. According to CSD theory, this is the moment a system starts to tip towards an alternate state. This point, also known as the tipping point, is directly linked to EWS. It describes the exact moment the eigenvalues of a system move to the right half plane. However, using this bifurcation point in a LOC context is challenging. The location and nature of this point are highly dependent on the underlying system dynamics. Indeed, certain types of bifurcation points imply the presence of EWS while others do not. Unfortunately, determining the type of bifurcation when the equations of motion are not known is challenging. This has proven to be problematic in a damaged quadcopter context where the dynamics of the system are not fully understood. Additionally, in a system that has control inputs that can saturate, tipping can occur prior to this critical point. As such, in the inverted pendulum tests, it was found that when the system tips due to a poorly tuned controller, a tipping point is approached, and EWS are found. When the pendulum tips due to actuators that saturate, the presence of EWS depends on the proximity of the bifurcation point to actuator saturation. If this distance is too large, EWS signals cannot be used. This concept is shown in Section 6.1.2 where the green lines show a sequence of possible EWS. The catastrophic saddle-node bifurcation point is also shown. The EWS are found prior to this point but not prior to the critical actuator saturation moment. It was concluded that in systems that tip directly due to actuator saturation, failure is best predicted by predicting actuator values. In systems

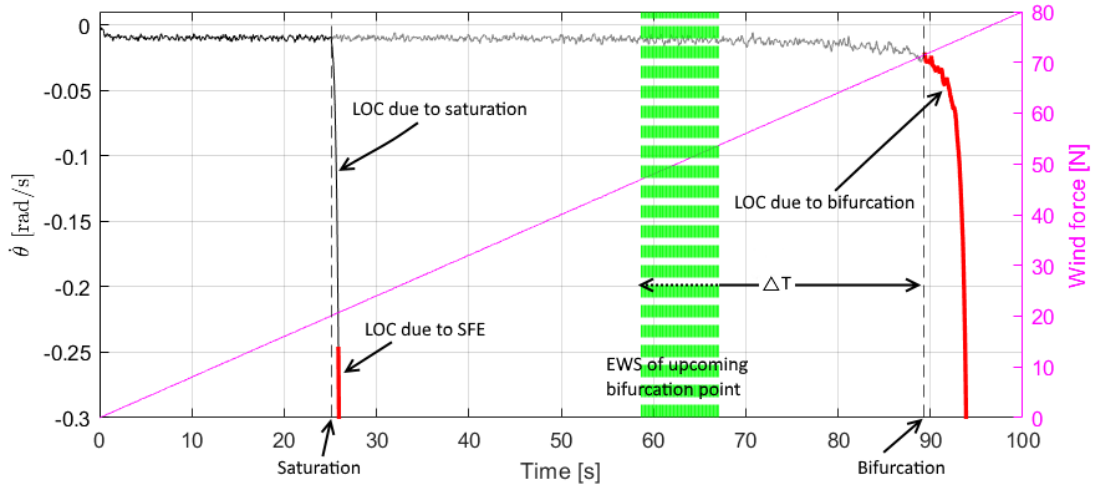


Figure 6.1: Time series of the angular rate of the inverted pendulum on cart problem. LOC can be defined as the moment the system exits the SFE, the moment actuator saturation occurs, or when the bifurcation point is approached. In green, a set of EWS is shown. An EWS is defined by rising trends in the AR1, SD and PSD indicators.

that do not directly fail due to actuator saturation, CSD can lead to EWS that could be used to predict tipping.

Finally, in systems where the bifurcation cannot be determined, such as the damaged quadcopter, LOC was ultimately specified as a variable region. The wind speed the system crashes at, ($V_{critical}$), was identified in the time series and LOC was defined as the region prior to the crash: LOC = $[V_{critical} - 1, V_{critical}]$. The lower bound was varied to examine the sensitivity of the results. Depending on the $V_{critical}$ of the flight, the lower bound was set closer or not to the failure wind speed. Overall, however, the prediction horizon was set as the time associated with $V_{critical}$.

These conclusions provide an answer to the second sub-question. In the context of EWS, LOC is best modeled as the moment a system passes a bifurcation point. The SFE seems to be a different type of problem, and actuator saturation can either be a direct cause of failure or a useful precursor that can be used in combination with the EWS obtained through CSD theory.

6.1.3. Predicting LOC with EWS

The last conclusions of this thesis work relate to the capacity of these EWS as LOC predictors. EWS are usually used retrospectively in slow time-varying systems to determine whether or not a failure could have been predicted. The limitations when applied to the faster time-varying system were therefore also explored.

A key element when analyzing a time series for EWS is the behavior of the external driver that pushes the system to tip. In the inverted pendulum, the wind speed was increased both slowly as rapidly. Out of the 1000 simulations performed on the slow time-varying system, 7% were false negatives, while 16% were false positives. When the driver was increased rapidly, these numbers deteriorate quickly, 38% of the simulations were false negatives, and 61% were false positives. Even though the suitability of an NDI controlled pendulum (that never tips) as a true negative scenario can be put into question, this highlights the requirements of the external driver. It should increase slowly and monotonically, allowing the system to be in a quasi-steady state condition at all times. A sensitivity analysis furthermore revealed that the most important parameter is the size of the rolling window used, especially in the fast time-varying system. The amount of EWS found drastically changes with the number of data points considered. Larger windows yield less false positives, yet for robotic applications, smaller windows are preferred. Indeed, it can be reasoned that larger windows require more data points hence more time to quantify whether a portion of a time series can be considered as an EWS. Furthermore, it was found that a larger sampling frequency does not directly contribute to improving the results obtained in the fast-changing dynamical system.

The quadcopter yields similar results. A concentration of EWS at lower wind speeds is observed across multiple flights when the entire time series is considered. When the criteria of acceptance of a candidate signal are strengthened, the remaining EWS concentrate at lower wind speeds (e.g., 2m/s). This raises questions about the significance of these EWS as quadcopters crash at larger wind speeds (e.g., 7m/s). Unlike the inverted pendulum, no EWS are found when this critical wind speed is approached. This could be due to several reasons. First, the data structure is inadequate. Instead of monotonically increasing the wind speed, it was increased during the experiments in a stepwise way. Secondly, the actuators of the damaged quadcopter saturate more often at higher wind speeds. Actuator saturation causes the dynamic behavior of the system to change from closed to open loop, hence the types of bifurcations. This could affect the presence of EWS as they depend heavily on the nature and position of the bifurcation point.

Nonetheless, several positive results were obtained. When analyzing the steady-state behavior of the quadcopter at each wind step, a rise in the indicators across most flights was found. This indicates that a form of CSD is occurring. The time series of the front rotor was more prone to show this behavior across the various indicators than the pitch rate. Secondly, the increasing frequency of actuators saturating when failure is approached was leveraged. When used in combination with CSD indicators, encouraging results were obtained. Out of the ten flights considered, only three did not contain EWS. Most likely, this is due to how saturation was defined. The remainder of the flights showed increasing EWS as tipping was approached. Here too, the amount of EWS found was larger when the front rotor speed was analyzed instead of the pitch rate time series. Most of these EWS were concentrated at wind speeds close to $V_{critical}$.

Finally, the robustness of the EWS found in the various systems is low. In the slow time-varying inverted pendulum, a different random seed was used across the simulations to model various stochastic conditions. The time location of EWS across the different runs varied from 40s prior to tipping (out of a 100s simulation) to just before tipping (T_{bif}=96s). If a method is to be used as a LOC mechanism, the time location of EWS across the different runs should be the same. Various parameters were modified in a bid to improve robustness across various runs, but no single one was found to cause a significant change.

Concluding on sub-question three, EWS due to CSD are found in slow time-varying systems but less so in fast-changing ones. Using them as reliable predictors of LOC appears more problematic as for various simulations, they appear at more sporadic locations. Poor results were also found when analyzing the transient quadcopter data. However, a steady-state analysis shows that there might be more to be uncovered with a proper data set. Additionally, when expanding the concepts of CSD to include actuator saturation as an indicator of LOC, more encouraging results are found.

6.2. Recommendations

Several recommendations and future steps are provided in this section. First, recommendations on the work in this thesis are provided.

1. The six most used indicators were considered at the beginning of the work. However, the Lyapunov exponent, which has recently been used in a variety of more complex systems, has not been examined. Whereas the other indicators perform best when a saddle-node bifurcation is present, the Lyapunov exponent can be a good predictor in different critical transitions. Additionally, entropy measures such as permutation entropy, dynamics complexity, or recurrence quantification could also provide more insight - especially when critical fluctuations cause the tipping of the system.

2. The current results for the quadcopter analysis were obtained using a poor dataset. The wind speed was increased with steps. Instead, an experiment should be performed where the wind speed is increased in a slow and monotonic way until quadcopter failure. This will validate the current results further.

3. Variables for analysis in the quadcopter tests were selected with a random forest. This does not imply that these variables are the most sensitive to CSD. Selecting the variables in the direction of the lowest resilience could yield improved results.

4. In this work, a significant EWS is defined as one with large trends and large endpoints. In both cases,

"significant" is determined by comparing the value to a threshold mark. For the endpoints, this threshold is defined using prior knowledge about the possible outcomes of an indicator in order to define what significant could be. In order to use this method online, significance should be defined without prior knowledge.

5. The actuator saturation indicator can be further expanded and researched. One approach to use this phenomenon was proposed in this work, but there could be other ways to extract useful information.

6. In this work, the sensitivity of the analysis parameters was examined. However, these parameters could also be optimized in a bid to keep the important EWS.

The most important conclusion, however, was to expand the search for EWS beyond traditional CSD theory. Possible research directions to undertake in future steps are provided below.

1. Signs of CSD are strongly dependent on the type of bifurcations present in the system. One possible research direction is to further research and explore the types of bifurcations present in a quadcopter.

2. To best observe EWS, a slow passage through a bifurcation point is also required. A method capable of dealing with transients is required if data-based methods are to be used for LOC prediction. From a mathematical perspective, this could be linked to rate induced tipping. In such tipping, failure is caused as the system is not capable of tracking a continuously changing quasi-steady state. This could be one direction worth exploring, even an initial literature study on the topic indicated that the same CSD indicators are used.

3. The quadcopter is inherently a stable system. The quadcopter will most likely enter a LOC condition when one or more actuators saturate. This causes the dynamics to change from closed to open loop. This was shown to be problematic when searching for EWS. Instead, this method could be applied to dynamical systems that do not suffer from actuator saturation, such as aircraft stall.

4. An alarm system should still be developed. In order to do so, a better understanding between the EWS and the subsequent failure, if any, must be developed. A relationship between the EWS frequency and failure should be derived.

Bibliography

[1] Quadcopter Arena. The history of drones and quadcopters, Accessed 01-2019.

[2] FAA. Faa releases aerospace forecast, 2018.

[3] Ewoud JJ Smeur, Qiping Chu, and Guido CHE de Croon. Adaptive incremental nonlinear dynamic inversion for attitude control of micro air vehicles. *Journal of Guidance, Control, and Dynamics*, 38(12):450–461, 2015.

[4] Parrot. Parrot bebop 2, Accessed 01-2019.

[5] Youmin Zhang and Jin Jiang. Bibliographical review on reconfigurable fault-tolerant control systems. *Annual reviews in control*, 32(2):229–252, 2008.

[6] Xuewu Dai and Zhiwei Gao. From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis. *IEEE Transactions on Industrial Informatics*, 9(4):2226–2238, 2013.

[7] A Freddi, S Longhi, and A Monteriu. A model-based fault diagnosis system for a mini-quadrotor. In *7th workshop on Advanced Control and Diagnosis*, pages 19–20, 2009.

[8] Alessandro Freddi, Sauro Longhi, and Andrea Monteriu. Actuator fault detection system for a mini-quadrotor. In *2010 IEEE International Symposium on Industrial Electronics*, pages 2055–2060. IEEE, 2010.

[9] Ngoc Nguyen and Sung Hong. Sliding mode thau observer for actuator fault diagnosis of quadcopter uavs. *Applied Sciences*, 8(10):1893, 2018.

[10] Agus Hasan and Tor Arne Johansen. Model-based actuator fault diagnosis in multirotor uavs. In *International Conference on Unmanned Aircraft Systems, Dallas, USA*, 2018.

[11] M Hadi Amoozgar, Abbas Chamseddine, and Youmin Zhang. Experimental test of a two-stage kalman filter for actuator fault detection and diagnosis of an unmanned quadrotor helicopter. *Journal of Intelligent & Robotic Systems*, 70(1-4):107–117, 2013.

[12] Chen Li, Yu Zhang, and Ping Li. Extreme learning machine based actuator fault detection of a quadrotor helicopter. *Advances in Mechanical Engineering*, 9(6):1687814017705068, 2017.

[13] Elgiz Baskaya, Murat Bronz, and Daniel Delahaye. Fault detection & diagnosis for small uavs via machine learning. In *Digital Avionics Systems Conference (DASC), 2017 IEEE/AIAA 36th*, pages 1–6. IEEE, 2017.

[14] Qian Zhang, Xueyun Wang, Xiao Xiao, and Chaoying Pei. Design of a fault detection and diagnose system for intelligent unmanned aerial vehicle navigation system. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, page 0954406218780508, 2018.

[15] Tong Zhang. An introduction to support vector machines and other kernel-based learning methods. *AI Magazine*, 22(2):103, 2001.

[16] Zachary Dydek, Anuradha Annaswamy, and Eugene Lavretsky. Combined/composite adaptive control of a quadrotor uav in the presence of actuator uncertainty. In *AIAA Guidance, Navigation, and Control Conference*, page 7575, 2010.

[17] D Zachary, A Anuradha, and L Eugene. Adaptive control of quadrotor uavs in the presence of actuator uncertainties ser. In *Infotech@ Aerospace Conferences. AIAA*, 2010.

[18] Ngoc Nguyen and Sung Hong. Fault-tolerant control of quadcopter uavs using robust adaptive sliding mode approach. *Energies*, 12(1):95, 2019.

[19] Lénaïck Besnard, Yuri B Shtessel, and Brian Landrum. Quadrotor vehicle control via sliding mode controller driven by sliding mode disturbance observer. *Journal of the Franklin Institute*, 349(2):658–684, 2012.

[20] Univeristy of Cagliari. A quick introduction to sliding mode control and its applications, Accessed 01-2019.

[21] Jing Sun. *Model Reference Adaptive Control*, page 1–7. Springer London, 2014.

[22] Alessandro Freddi, Alexander Lanzon, and Sauro Longhi. A feedback linearization approach to fault tolerance in quadrotor vehicles. *IFAC proceedings volumes*, 44(1):5413–5418, 2011.

[23] Vincenzo Lippiello, Fabio Ruggiero, and Diana Serra. Emergency landing for a quadrotor in case of a propeller failure: A pid based approach. In *Safety, Security, and Rescue Robotics (SSRR), 2014 IEEE International Symposium on*, pages 1–7. IEEE, 2014.

[24] Mark W Mueller and Raffaello D'Andrea. Relaxed hover solutions for multicopters: Application to algorithmic redundancy and novel vehicles. *The International Journal of Robotics Research*, 35(8):873–889, 2016.

[25] Sihao Sun, Leon Sijbers, Xuerui Wang, and Coen de Visser. High-speed flight of quadrotor despite loss of single rotor. *IEEE Robotics and Automation Letters*, 3(4):3201–3207, 2018.

[26] Raymond E. King. Commercial aviation safety team and joint safety analysis teams, Accessed 01-2019.

[27] Commercial Aviation Safety Team. Airplane state awareness joint safety analysis team final report. *Washington, DC: CAST*, 2014.

[28] Christine M Belcastro, John V Foster, Gautam H Shah, Irene M Gregory, David E Cox, Dennis A Crider, Loren Groff, Richard L Newman, and David H Klyde. Aircraft loss of control problem analysis and research toward a holistic solution. *Journal of Guidance, Control, and Dynamics*, 40(4):733–775, 2017.

[29] James Wilborn and John Foster. Defining commercial transport loss-of-control: A quantitative approach. In *AIAA atmospheric flight mechanics conference and exhibit*, page 4811, 2004.

[30] Christine M Belcastro, Richard L Newman, Joni Evans, David H Klyde, Lawrence C Barr, and Ersin Ancel. Hazards identification and analysis for unmanned aircraft system operations. In *17th AIAA Aviation Technology, Integration, and Operations Conference*, page 3269, 2017.

[31] John V Foster and David Hartman. High-fidelity multi-rotor unmanned aircraft system (uas) simulation development for trajectory prediction under off-nominal flight dynamics. In *17th AIAA Aviation Technology, Integration, and Operations Conference*, page 3271, 2017.

[32] Christine M Belcastro and Steven R Jacobson. Future integrated system concepts for preventing aircraft loss-of-control accidents. In *AIAA Guidance, Navigation, and Control Conference*, pages 2–5, 2010.

[33] June Chongvisal, Nikolas Tekles, Enric Xargay, Don Talleur, Alex Kirlik, and Naira Hovakimyan. Loss-of-control prediction and prevention for nasa's transport class model. In *AIAA Guidance, Navigation, and Control Conference*, page 0784, 2014.

[34] Stefan Schuet, Thomas Lombaerts, Diana Acosta, John Kaneshige, Kevin Wheeler, and Kimberlee Shish. Autonomous flight envelope estimation for loss-of-control prevention. *Journal of Guidance, Control, and Dynamics*, 40(4):847–862, 2016.

[35] Nikolas Tekles, June Chongvisal, Enric Xargay, Ronald Choe, Donald A Talleur, Naira Hovakimyan, and Christine M Belcastro. Design of a flight envelope protection system for nasa's transport class model. *Journal of Guidance, Control, and Dynamics*, 40(4):863–877, 2016.

[36] Ye Zhang, Coen C de Visser, and Q Ping Chu. Online safe flight envelope prediction for damaged aircraft: A database-driven approach. In *AIAA Modeling and Simulation Technologies Conference*, page 1189, 2016.

[37] Remon van den Brandt and CC de Visser. Safe flight envelope uncertainty quantification using probabilistic reachability analysis. *IFAC-PapersOnLine*, 51(24):628–635, 2018.

[38] Mingzhou Yin. Envelope estimation and protection of innovative control effectors (ice) aircraft: A probabilistic approach. 2018.

[39] Ye Zhang, Coen C de Visser, and Q Ping Chu. Database building and interpolation for a safe flight envelope prediction system. In *2018 AIAA Information Systems-AIAA Infotech@ Aerospace*, page 1635, 2018.

[40] Sihao Sun and Coen C de Visser. Quadrotor safe flight envelope prediction in the high-speed regime: A monte-carlo approach. In *AIAA Scitech 2019 Forum*, page 0948, 2019.

[41] Kieran Kersbergen. Quantifying Loss-Of-Control of Quadrotors. Master's thesis, TU Delft, the Netherlands, 2018.

[42] Man Shan Kan, Andy CC Tan, and Joseph Mathew. A review on prognostic techniques for non-stationary and non-linear rotating systems. *Mechanical Systems and Signal Processing*, 62:1–20, 2015.

[43] Mark Schwabacher. A survey of data-driven prognostics. In *Infotech@ Aerospace*, page 7002. 2005.

[44] Yanfei Lu, Qing Li, and Steven Y Liang. Physics-based intelligent prognosis for rolling bearing with fault feature extraction. *The International Journal of Advanced Manufacturing Technology*, pages 1–10, 2018.

[45] Andrea De Martin, Giovanni Jacazio, and George Vachtsevanos. Anomaly detection and prognosis for primary flight control emas. In *3rd European Conference of the Prognostics and Healh Management Society*, pages 5–8, 2016.

[46] Soheila Mehrmolaei and Mohammad Reza Keyvanpourr. A brief survey on event prediction methods in time series. In *Artificial Intelligence Perspectives and Applications*, pages 235–246. Springer, 2015.

[47] Jyoti Soni, Ujma Ansari, Dipesh Sharma, and Sunita Soni. Predictive data mining for medical diagnosis: An overview of heart disease prediction. *International Journal of Computer Applications*, 17(8):43–48, 2011.

[48] Salvador Gabarda and Gabriel Cristóbal. Detection of events in seismic time series by time-frequency methods. *IET Signal Processing*, 4(4):413–420, 2010.

[49] Chaitanya Damle and Ali Yalcin. Flood prediction using time series data mining. *Journal of Hydrology*, 333(2-4):305–316, 2007.

[50] Emilio Florido, Francisco Martínez-Álvarez, Antonio Morales-Esteban, Jorge Reyes, and José Luis Aznarte-Mellado. Detecting precursory patterns to enhance earthquake prediction in chile. *Computers & geosciences*, 76:112–120, 2015.

[51] Saptarsi Goswami, Sanjay Chakraborty, Sanhita Ghosh, Amlan Chakrabarti, and Basabi Chakraborty. A review on application of data mining techniques to combat natural disasters. *Ain Shams Engineering Journal*, 9(3):365–378, 2018.

[52] Didier Sornette. *Why stock markets crash: critical events in complex financial systems*. Princeton University Press, 2017.

[53] Xiang-Bin Yan, Tao Lu, Yi-Jun Li, and Guang-Bin Cui. Research on event prediction in time-series data. In *Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on*, volume 5, pages 2874–2878. IEEE, 2004.

[54] Sergey Dolenko, Alexander Guzhva, Igor Persiantsev, and Julia Shugai. Multi-stage algorithm based on neural network committee for prediction and search for precursors in multi-dimensional time series. In *International Conference on Artificial Neural Networks*, pages 295–304. Springer, 2009.

[55] Eugine Kang. Time series: Check stationarity, 2017. Accessed on 02-2019.

[56] SA Dolenko, Yu V Orlov, IG Persiantsev, and Yu S Shugai. Neural network algorithms for analyzing multidimensional time series for predicting events and their application to study of sun-earth relations. *Pattern Recognition and Image Analysis*, 17(4):584–591, 2007.

[57] SA Dolenko, Yu V Orlov, IG Persiantsev, and Ju S Shugai. Neural network algorithm for events forecasting and its application to space physics data. In *International Conference on Artificial Neural Networks*, pages 527–532. Springer, 2005.

[58] Guokun Lai, Wei-Cheng Chang, Yiming Yang, and Hanxiao Liu. Modeling long-and short-term temporal patterns with deep neural networks. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pages 95–104. ACM, 2018.

[59] Gautam Das, King-Ip Lin, Heikki Mannila, Gopal Renganathan, and Padhraic Smyth. Rule discovery from time series. In *KDD*, volume 98, pages 16–22, 1998.

[60] Lexiang Ye and Eamonn Keogh. Time series shapelets: a new primitive for data mining. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 947–956. ACM, 2009.

[61] Mohamed F Ghalwash and Zoran Obradovic. Early classification of multivariate temporal observations by extraction of interpretable shapelets. *BMC bioinformatics*, 13(1):195, 2012.

[62] Guoliang He, Yong Duan, Guofu Zhou, and Lingling Wang. Early classification on multivariate time series with core features. In *International Conference on Database and Expert Systems Applications*, pages 410–422. Springer, 2014.

[63] Roberto Medico, Joeri Ruyssinck, Dirk Deschrijver, and Tom Dhaene. Learning multivariate shapelets with multi-layer neural networks. In *ACDL-Advanced Course on Data Science & Machine Learning*, 2018.

[64] Arvind Kumar Shekar, Marcus Pappik, Patricia Iglesias Sánchez, and Emmanuel Müller. Selection of relevant and non-redundant multivariate ordinal patterns for time series classification. In *International Conference on Discovery Science*, pages 224–240. Springer, 2018.

[65] Guoliang He, Yong Duan, Rong Peng, Xiaoyuan Jing, Tieyun Qian, and Lingling Wang. Early classification on multivariate time series. *Neurocomputing*, 149:777–787, 2015.

[66] Pang-Ning Tan et al. *Introduction to data mining*. Pearson Education India, 2007.

[67] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3):15, 2009.

[68] Christoffer Brax. *Anomaly detection in the surveillance domain*. PhD thesis, Örebro universitet, 2011.

[69] Rodney Martin and Santanu Das. Near real-time optimal prediction of adverse events in aviation data. In *AIAA Infotech@ Aerospace 2010*, page 3517. 2010.

[70] Rodney A Martin. An investigation of state-space model fidelity for ssme data. In *Prognostics and Health Management, 2008. PHM 2008. International Conference on*, pages 1–12. IEEE, 2008.

[71] Rodney A Martin. Unsupervised anomaly detection and diagnosis for liquid rocket engine propulsion. In *Aerospace Conference, 2007 IEEE*, pages 1–15. IEEE, 2007.

[72] Santanu Das, Lishuai Li, Ashok Srivastava, and R John Hansman. Comparison of algorithms for anomaly detection in flight recorder data of airline operations. In *12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, page 5593, 2012.

[73] Rodney A Martin, Das Santanu, Vijay Manikandan Janakiraman, and Stefan Hosein. Accept: Introduction of the adverse condition and critical event prediction toolbox. 2015.

[74] James Martens. Learning the linear dynamical system with asos. In *ICML*, pages 743–750. Citeseer, 2010.

[75] Rodney A Martin. A state-space approach to optimal level-crossing prediction for linear gaussian processes. *IEEE Transactions on Information Theory*, 56(10):5083–5096, 2010.

[76] Vijay Manikandan Janakiraman, Bryan Matthews, and Nikunj Oza. Discovery of precursors to adverse events using time series data. In *Proceedings of the 2016 SIAM International Conference on Data Mining*, pages 639–647. SIAM, 2016.

[77] Markus Goldstein and Seiichi Uchida. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PloS one*, 11(4):e0152173, 2016.

[78] Eliahu Khalastchi, Meir Kalech, Gal A Kaminka, and Raz Lin. Online data-driven anomaly detection in autonomous robots. *Knowledge and Information Systems*, 43(3):657–688, 2015.

[79] Marten Scheffer, Jordi Bascompte, William A Brock, Victor Brovkin, Stephen R Carpenter, Vasilis Dakos, Hermann Held, Egbert H Van Nes, Max Rietkerk, and George Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53, 2009.

[80] Vasilis Dakos, Stephen R Carpenter, Egbert H van Nes, and Marten Scheffer. Resilience indicators: prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1659):20130263, 2015.