

Data-driven analysis of wind power and power system dynamics via Koopman mode decomposition

Johan Fredrik Raak

2017

Abstract

Power systems are enormous dynamical systems and incredibly complex with a huge number of devices, generators and consumers connected to them. Power systems around the world are currently undergoing modernization to accommodate a larger share of renewable energy and to incorporate new technologies that enable better and faster monitoring and control. The importance of assessing the impact of new renewable generation such as solar and wind power on power system operation has widely been discussed and acknowledged. It has often been suggested that power systems are becoming more vulnerable due to new energy policies and higher demands, driving the systems to operate closer to their limits. In this dissertation, wind power is in focus because it is the fastest growing source of renewable energy in the world, and the largest contributor of new renewable energy.

To provide a steady supply of electricity to its customers, power systems are constantly monitored and regulated by balancing generation and consumption. In the case of severe faults such as a sudden loss of a major transmission line or a generator, undamped oscillations between generators can cause asynchrony between generators and the system, causing generators to disconnect, and potentially causing destabilization of the entire system and widespread blackouts. Blackouts have occurred numerous times in the last couple of decades, and are devastating to modern society. Protection strategies against cascading failures are imperative to avoid large scale blackouts. Quickly extracting relevant information of the system state from measurements and taking appropriate action in such cases are extremely important measures to ensure safe operation. The introduction of large-scale renewables inevitably creates more variable power system operating conditions. Since the dynamics depend on the current loading and the stress on transmission lines, it becomes even more vital to monitor the dynamics to ensure stable operation.

This dissertation focuses on data-driven analysis of power systems and wind power systems by utilizing a recently developed mathematical framework called Koopman Mode Analysis (KMA). KMA is based on spectral analysis of the so-called Koopman operator which governs the evolution of observables of nonlinear dynamical systems. KMA is related to the so-called Dynamic Mode Decomposition (DMD)—a promising technique for the analysis of complex multivariate data which was proposed a few years ago to analyze complex fluid flows. Since the advent of DMD, a great deal of interest has been directed towards this technique, and many new extensions and variants of DMD have been proposed. The purpose of these techniques is essentially to extract dynamic characteristics such as oscillatory structures or *modes* from complex, nonlinear systems, by utilizing only measurements, without a-priori knowledge of the system. This can be viewed as a model-

reduction technique, which aims at capturing the essential dynamic characteristics of a possibly unknown or ill-defined system, and enables analysis, prediction and control of the target system. Connections between DMD and the Koopman operator have been highlighted, and thus DMD has been called Koopman Mode Decomposition (KMD).

This dissertation demonstrates and proposes several new applications of KMD related to the analysis of wind power and power systems. In the first two chapters following the introduction, a thorough review of the Koopman operator and KMD is provided, including descriptions of the most common numerical techniques, and theorems connecting two different algorithms. Connections between Koopman and Perron-Frobenius operators are also discussed and numerically demonstrated with a simple system. Then, three KMD-algorithms are evaluated with several data sets and under the influence of various levels of noise. These results clarify the applicability and suitability of the algorithms to different types of data, and become invaluable to potential practitioners of KMD.

Following that, the bulk of this dissertation then discusses wind power fluctuations in Japan. Wind power fluctuations are analyzed on spatial scales ranging from hundreds of kilometers to hundreds of meters, and on timescales ranging from seconds to weeks. First, a statistical evaluation of short-term wind power fluctuations is conducted which shows that significant power output fluctuations from wind turbines are to a large extent not suppressed in the wind farm output, giving rise to large output fluctuations. The results imply that care should be taken about short-term correlations when installing large-scale wind power in a region. Then, a weather simulator which has not previously been used in wind power analysis is utilized, and an evaluation against real measurements from a wind farm in Japan is conducted. Following that, large-scale smoothing effects of wind power are investigated by incorporating data from weather simulations. The smoothing effects are determined by the correlation between wind farm outputs at various time-scales. A better understanding of the smoothing characteristics, e.g. understanding the size of the wind farm distribution area required to achieve a certain degree of smoothing in the aggregated wind power, would contribute to smarter planning and better utilization of equipment and natural energy resources, and greater reliability. To quantify the smoothing effects of wind power in Japan, a new index via KMD is proposed. The results from applying the proposed index highlight the importance of deliberately selecting sites for large-scale wind power production to more effectively smooth the aggregated power. Finally, it is demonstrated how very short-term wind speed fluctuations can be predicted a few seconds in the future by KMD.

This dissertation then looks at applications of KMD related to power system analysis and security. In particular, by applying KMD to power system dynamics, connections between so-called Koopman modes and spectral graph theory are discovered. This indicates that graph properties can be inferred from dynamics on networks via the spectrum of the Koopman operator. Also, a practical data-driven method utilizing KMD is outlined to partition a power system into disjoint parts, which is a vital component of a controlled islanding strategy aimed to save power systems from large-scale blackouts. This investigation indicates that data-driven methods such as KMD can be regarded as promising tools for effective monitoring and control of future, more variable, power systems.

Acknowledgments

First and foremost, I would like to extend my deepest gratitude to Prof. Takashi Hikihara for his constant support, encouragement, positivity, invaluable suggestions, and for always being an inspiration. I would also like to sincerely thank Prof. Yoshihiko Susuki for always looking out for me, supporting my stay in Japan, for an uncountable number of discussions, for always inspiring me to do better, and for introducing me to the academic world and to many excellent researchers. I would also like to thank Prof. Kazuhisa Tsuboki for discussions, collaboration and for teaching me about weather simulations, Prof. Hiroshi Uyeda for valuable discussions and help with weather simulation data, Mr. Masaya Kato for collaboration and hard work and support with weather simulations, Mr. Shinya Eguchi for collaboration and discussions about wind farm characteristics and operation, Prof. Igor Mezic for valuable suggestions and discussions and for inspiring me with his enthusiasm, Dr. Ryosuke Morita for helping me processing weather data, Mr. Harold Chamorro and Prof. Mehrdad Ghandhari for valuable discussions and suggestions, and Prof. Shinji Doi and Prof. Tetsuji Matsuo for useful suggestions, discussions, and for evaluating this work. I would also like to express my gratitude to Prof. Atsushi Ishigame who welcomed me to his lab at Osaka Prefecture University. Moreover, I would like to thank everyone in Hikihara lab and especially Mrs. Yoshiko Deguchi who has always been extremely helpful and friendly to me. I would also like to thank Dr. Takafumi Okuda for assisting me in getting this dissertation printed, and Dr. Shinya Nawata, Dr. Alexandros Kordonis, Dr. Hikaru Hoshino, Mr. Takuya Kajiyama, Mr. Yoshihiko Yamaguchi, Dr. Atsushi Yao, and Dr. Madoka Kubota for taking well care of me when I first came to Japan and Kyoto University. In particular, I would like to thank Dr. Shinya Nawata and Dr. Hikaru Hoshino for helping me with Japanese. I would also like to direct special thanks to Ms. Sepideh Shahmohammadi, Mr. Stefan Romare, Ms. Rutvika Manohar, Mr. Manuel Sánchez, and Dr. Joseph Scott for checking various parts of this dissertation for mistakes and typos. Furthermore, I am extremely thankful to Mitsubishi Corporation for granting me the Mitsubishi Corporation International Student Scholarship. I am also thankful for support from JST-CREST who in part supported the work presented in this dissertation. Finally, I am forever grateful to my family and Sepideh for always encouraging and supporting me.

Contents

Abstract	i
Acknowledgments	iii
Contents	v
Notations	ix
Acronyms	xi
1 Introduction	1
1.1 Koopman operator theory and KMA	3
1.2 Power systems	3
1.3 Wind power	4
1.4 Data-driven analysis of wind power and power system dynamics	5
1.5 Dissertation objectives and overview	8
2 On the Koopman operator and the Koopman mode decomposition	11
2.1 Introduction to the Koopman operator and the Koopman mode decomposition	13
2.2 KMD algorithms	14
2.2.1 Mathematical formulation	14
2.2.2 The Arnoldi-like algorithm	15
2.2.3 Dynamic mode decomposition	16
2.2.4 Equivalence between Arnoldi-KMD and DMD	17
2.2.5 Hankel-like matrices and Prony KMD	19
2.2.6 Extended DMD	21
2.3 The Perron-Frobenius (PF) operator	22
2.3.1 Perron-Frobenius via Ulam’s method and EDMD	23
2.4 Two illustrative examples	24
2.4.1 A linear system	24
2.4.2 A nonlinear system	25
2.5 Conclusions	28

3	Comparison and numerical evaluation of KMD-techniques	31
3.1	Introduction	31
3.2	Methodology	32
3.2.1	Sorting and selecting modes	32
3.2.2	Comparing eigenvalues	33
3.2.3	Comparison of modal frequency	33
3.2.4	Comparison via time-shifted snapshots	33
3.2.5	Comparison via future prediction	34
3.3	Description of test systems and data	34
3.3.1	Linear system and noise	34
3.3.2	Power system example	34
3.3.3	Weather prediction data	35
3.4	Results	35
3.4.1	Noise-corrupted data from a linear system	36
3.4.2	Power system example	39
3.4.3	Future prediction of power system dynamics	42
3.4.4	Future prediction of wind speeds	43
3.5	Conclusions	44
4	Modeling of wind power systems	47
4.1	Wind power and wind turbine modeling	48
4.1.1	Wind power	48
4.1.2	Mechanical model of wind turbines	49
4.1.3	Dynamics of mechanical system	50
4.2	Doubly-fed induction generator	51
4.3	Permanent magnet synchronous generator	52
4.4	Initialization and simulation of a wind farm	54
4.4.1	Initialization of wind turbines	55
4.4.2	Wind farm modeling	55
4.4.3	Simulation results	56
4.5	Conclusions	58
	Appendices	59
4.A	The dq -transformation	59
5	Assessment of short-term offshore wind power fluctuations in Japan	61
5.1	Introduction	61
5.1.1	Purpose and contributions	62
5.1.2	Outline	63
5.2	Description of target wind farm	63
5.3	Methods for statistical evaluation	64
5.4	Statistics of short-term wind speed and power	66
5.4.1	Deriving models for comparison	66
5.4.2	Statistical evaluation and comparison	69

5.5	Conclusions	75
6	Assessment of offshore wind power characteristics in Japan	77
6.1	Introduction	77
6.1.1	Literature review	78
6.1.2	Purpose and contributions	78
6.1.3	Outline	79
6.2	On measurements and data processing	80
6.3	CReSS model and outputs	81
6.4	Methods	84
6.4.1	Statistics for comparison	84
6.4.2	Power spectral density and spectral coherence	85
6.5	Evaluation of predicted wind and wind power characteristics	86
6.5.1	Wind speed	86
6.5.2	Wind power	91
6.5.3	Wind direction	94
6.5.4	Spectral coherence	97
6.6	Conclusions	98
7	Quantifying smoothing effects of wind power via KMD	101
7.1	Introduction	101
7.1.1	Literature review	101
7.1.2	Proposed idea	102
7.1.3	Outline	103
7.2	Statistical quantification of smoothing	103
7.3	Quantification of smoothing via PSD	104
7.4	Smoothing via Koopman modes	105
7.4.1	New smoothing indexes via KMD	106
7.5	Comparison between smoothing indexes	107
7.6	Test with large-scale wind prediction data	110
7.6.1	Data sectioning and mean smoothing with KMD	110
7.6.2	Smoothing of aggregated power in Honshu, Japan	111
7.7	Conclusions	116
8	Predicting wind speeds via KMD	117
8.1	Introduction	117
8.2	Motivation and problem formulation	118
8.3	Characterization of wind speed fluctuations	120
8.4	Prediction methods	125
8.4.1	Preliminaries and simple prediction methods	125
8.4.2	Nearest neighbor prediction	126
8.4.3	Autoregressive (AR) model	127
8.4.4	On standard DMD and AR	128

8.4.5	On Arnoldi-like KMD and AR	129
8.4.6	DMD and nearest neighbor prediction	130
8.4.7	Reshaped DMD via optimal hard threshold	131
8.5	Numerical test of wind speed prediction	131
8.6	Conclusions	133
	Appendices	137
8.A	Cao’s test for dimension and determinism	137
9	Partitioning power networks via KMD	139
9.1	Introduction	139
9.1.1	Proposed idea and contributions	140
9.2	Theoretical backgrounds	140
9.2.1	Modeling of power system dynamics	141
9.2.2	Slow coherency	141
9.2.3	Spectral graph theory for network partitioning	143
9.3	Proposed data-driven partitioning method	144
9.3.1	Coherency in Koopman modes	144
9.3.2	Outline of KMD-based partitioning strategy	145
9.4	Demonstration	148
9.4.1	Benchmark systems and simulation setting	148
9.4.2	Simple test system	149
9.4.3	An analytical evaluation	150
9.4.4	IEEE 118-bus test system	152
9.4.5	Comparison with slow-coherency-based islanding	156
9.4.6	Computational considerations	157
9.5	Number of partitions and multiple coherency	158
9.5.1	Demonstration of multiple coherency in KMs	160
9.5.2	Time-domain simulations	160
9.6	Conclusions	161
10	Conclusions and future outlook	165
10.1	Conclusions	166
10.2	Future work and outlook	168
	Bibliography	171
	List of Publications	189

Notations

$\frac{dx}{dt}$ (or $\frac{d\mathbf{x}}{dt}$)	Time-derivative of variable x (or variables contained in a vector $\mathbf{x}$)
$[\mathbf{x}]_i$ (or $[\mathbf{A}]_{ij}$)	i -th (or (i, j) -th) element of a vector $\mathbf{x}$ (or a matrix $\mathbf{A}$)
$\ \mathbf{x}\ $	Euclidean norm of $\mathbf{x}$
$\mathbf{A}^\dagger$	Moore-Penrose pseudo-inverse of the matrix $\mathbf{A}$
$\text{diag}(e_1, \dots, e_n)$	$n \times n$ diagonal matrix with elements $(e_1, \dots, e_n)$
$\mathbf{I}_n$	$n \times n$ identity matrix: $[\mathbf{I}_n]_{ij} = 1$ iff $i = j$, and otherwise $[\mathbf{I}_n]_{ij} = 0$
$\top$	Transpose operation of vectors or matrices
$\mathbf{z}^*$	Complex-conjugate transpose of $\mathbf{z} \in \mathbb{C}^m$
$\text{Re}(\mathbf{z})$	Real part of complex value $\mathbf{z} = a + ib$ such that $\text{Re}(\mathbf{z}) = a$
$\text{Im}(\mathbf{z})$	Imaginary part of complex value $\mathbf{z} = a + ib$ such that $\text{Im}(\mathbf{z}) = b$
$z = z \angle\phi$	Angle notation of z with modulus $ z $ and argument $\phi := \text{Im}[\ln z]$
$\mathcal{N}$ (or $\mathcal{K}$)	Set $\{1, \dots, N\}$ (or $\{1, \dots, K\}$)
$\mathbb{N}$	Set of natural numbers $\{1, 2, \dots, \infty\}$
$\mathbb{Z}$	Set of integers $\{-\infty, \dots, -1, 0, 1, \dots, \infty\}$
$\langle \cdot \rangle$	Mean of the elements of a vector $\langle \mathbf{x} \rangle := ([\mathbf{x}]_1 + \dots + [\mathbf{x}]_m)/m$ or sample mean of a random variable $\langle x \rangle := (x_1 + \dots + x_N)/N$
$\sigma(x)$	Standard deviation of samples of a random variable x
$\langle \mathbf{x}, \mathbf{y} \rangle$	Inner product of vectors: $\mathbf{x}^\top \mathbf{y}$
$\text{floor}(x)$	The integer part of a value x : e.g. $\text{floor}(3/2) = 1$

Acronyms

KMA	Koopman Mode Analysis
KMD	Koopman Mode Decomposition
KM	Koopman Mode
KE	Koopman Eigenvalue
PF	Perron-Frobenius
DMD	Dynamic Mode Decomposition
EDMD	Extended Dynamic Mode Decomposition
NWP	Numerical Weather Prediction (or Predictor)
CR_eSS	Cloud Resolving Storm Simulator
CFD	Computational Fluid Dynamics
WP	Wind Power
WT	Wind Turbine
WF	Wind Farm
DFIG	Doubly Fed Induction Generator
PMSG	Permanent Magnet Synchronous Generator
MPPT	Maximum Power Point Tracking
TSR	Tip Speed Ratio
PSAT	Power System Analysis Toolbox
IB	Infinite Bus
PMU	Phasor Measurement Unit

DAE	Differential Algebraic Equations
PDF	Probability Density Function
CDF	Cumulative Density Function
PSD	Power Spectral Density
RMSE	Root Mean Square of Errors
STDE	Standard Deviation of Errors
SDE	Stochastic Differential Equation
SNR	Signal to Noise Ratio

Chapter 1

Introduction

Many systems and phenomena we come in contact with in our daily lives can be described within the framework of dynamical systems theory. Such systems are formulated by a set (*system*) of differential equations that govern their time-evolution and interactions between different variables (*dynamics*). Early important theoretical contributions to the study of dynamical systems were made by Henri Poincaré at the end of the 19th century. Birkhoff, Lyapunov, Andronov and Arnold are recognized as some of the most influential contributors in the 20th century [1]. Common examples of dynamical systems are atmospheric systems, power systems, and mechanical systems (e.g. pendulum and celestial dynamics).

There are many different approaches to the analysis of such systems, and the most suitable choice depends on the phenomena of interest (e.g. long-term or transient behavior) and the nature of the target system (e.g. stochastic or deterministic, discrete or continuous, linear or nonlinear). Through characterization and understanding of complex dynamical systems, we are continuously becoming better at predicting their future states, controlling their behavior, and consequently engineering smarter and more robust systems. In particular, an inherent feature of many real-world examples of dynamical systems is very complex behavior (fluctuations, turbulence), and an extreme sensitivity to initial conditions. This is commonly known as chaos [2, 3], and probably the most famous examples of it are weather systems, which are notoriously difficult to predict. Despite that, even chaotic dynamical systems exhibit coherent structures and patterns [4, 5], which are important to our understanding and characterization of them. This dissertation will to a large extent focus on characterizing coherence (correlation in some sense) of wind power fluctuations and power system oscillations. For example, this dissertation looks at Wind Farm (WF) outputs which are to some extent correlated over long distances, and whose aggregated characteristics affect power system operation [6]. It also looks at power systems (in this case not exhibiting any chaotic behavior), which when excited by a large disturbance can exhibit oscillations between groups of generators that can be detrimental to their operation if not adequately damped [7].

For many real-world systems, there is only a limited knowledge of the governing equations—leaving us to heuristic models with various accuracies, and data analyses com-

pletely dependent on numerical methods and data quality. In the last couple of decades, we have seen a huge interest in the latter approach, as new, sophisticated algorithms are continuously being developed together with the ever increasing computational capabilities, allowing for fast, *data-driven analysis*. Moreover, we might only have access to some observations of the system, e.g. voltages in a power system, and not state variables. State variables are a minimum set of variables, which form differential equations, and completely characterize the dynamics that we observe in the system. In the absence of knowledge of the state variables and the equations governing them, we have to completely rely on data-driven analysis of measurement data.

As a relatively recent approach to the study of dynamical systems, the so-called transfer operator-theoretic view has gained substantial traction. Transfer operators are linear mathematical operators characterizing *the evolution* of a system. It has been shown that their action can be estimated purely from data [8–10], hence they are appealing tools in data-driven-analysis of complex dynamical systems. In particular, the so-called Perron-Frobenius (PF) operator [11] governs the evolution of densities, and the Koopman operator [11, 12] which is the left-adjoint of the PF operator, governs the evolution of observable functions of the system. It has been argued that analyses conducted with either of these operators are theoretically equivalent because of their connection—merely representing different points of view [13]. However, their equivalent suitability in practice seems to remain unclear, and might depend on the situation and system under consideration. In particular, the so-called Koopman Mode Analysis (KMA) or Koopman Mode Decomposition (KMD) (these acronyms are sometimes used interchangeably) was introduced by Mezić in [14] as the spectral decomposition of the Koopman operator. The interest towards the Koopman-operator-framework has since then grown substantially, has been accompanied by a number of efficient algorithms linked to it (e.g [15–18]), and has resulted in many demonstrations of applications. The algorithms extract dynamic features (e.g. spatio-temporal oscillatory structures also called modes) from data without any use of models, which characterize the evolution of the observables and provide finite-dimensional approximations of the action of the operator.

The recent advances in KMA and applications in power systems [19–22] and fluid dynamics [15, 23, 24] serve as motivation for this work towards characterization of aggregated wind power and new power system applications. This dissertation focuses on wind power fluctuations in Japan and power system dynamics by incorporating ideas from the Koopman-operator-framework. One of the main purposes of this study is to explore and introduce a new approach to the study of wind power fluctuations. This is a new and promising application of KMA, which has previously been shown to extract coherent dynamic features from complex fluid flows. Large-scale wind power in Japan is investigated by incorporating weather simulations of various temporal and spatial resolutions. The dissertation also discusses the numerical applicability to the study of problems where only a few measurement locations are available or utilized, by evaluating several related algorithms. It also provides a review of the current state-of-the-art procedure of numerical KMA. This study thus serves as a guideline for future studies applying KMD-related methods to wind power, and systems with only a few measurements. Large-scale wind

power fluctuations in Japan has not been extensively studied, despite Japan’s great potential for wind power [25]. However, potential impacts and characteristics of large-scale aggregated wind power have been studied for decades and are still an active area of research, and correlations over long distances have been demonstrated [26]. Characterizing the coherence of these fluctuations becomes important to assess the impact of large-scale integration of wind power into power systems [6, 27, 28].

1.1 Koopman operator theory and KMA

KMA is related to the so-called Dynamic Mode Decomposition (DMD) [16]—a promising technique for the analysis of complex multivariate data which was proposed a few years ago to analyze complex fluid flows. Since the advent of DMD, a great deal of interest has been directed towards this technique, and many new extensions and variants of DMD have been proposed: see e.g. [17, 18]. The purpose of these techniques is essentially to extract dynamic characteristics such as oscillatory structures or *modes* from complex, nonlinear systems, by utilizing only measurements, without a-priori knowledge of the system. This can be viewed as a model-reduction technique, which aims to capture the essential dynamic characteristics of a possibly unknown or ill-defined system, and enables analysis, prediction, and control of the target system. Connections between DMD and the Koopman operator have been highlighted [15], and consequently DMD has been called Koopman Mode Decomposition (KMD) [23].

1.2 Power systems

Power systems are enormous dynamical systems and incredibly complex with a huge number of devices, generators and consumers connected to them [7, 29]. In order to provide a steady supply of electricity to its customers, power systems are constantly closely monitored and regulated by balancing generation and consumption. In the case of severe faults such as a sudden loss of a major transmission line or generator, undamped oscillations between generators can cause asynchrony between generators and the system, causing generators to disconnect, and potentially causing destabilization of the entire system and widespread blackouts. In such an example of *cascading* failure, a stable operating point is lost, which causes a voltage collapse (blackout) as a result. Blackouts have occurred numerous times in the last couple of decades [30] and are devastating to modern society. Protection strategies against cascading failures are imperative to avoid large scale blackouts. Quickly extracting relevant information of the system’s state from measurements and taking appropriate action in such cases are extremely important measures to ensure safe operation [31].

Power systems are currently undergoing modernization to accommodate a larger share of renewable energy and to incorporate new technologies that enable better and faster monitoring and control. The impact of new renewable generation and, in particular, wind

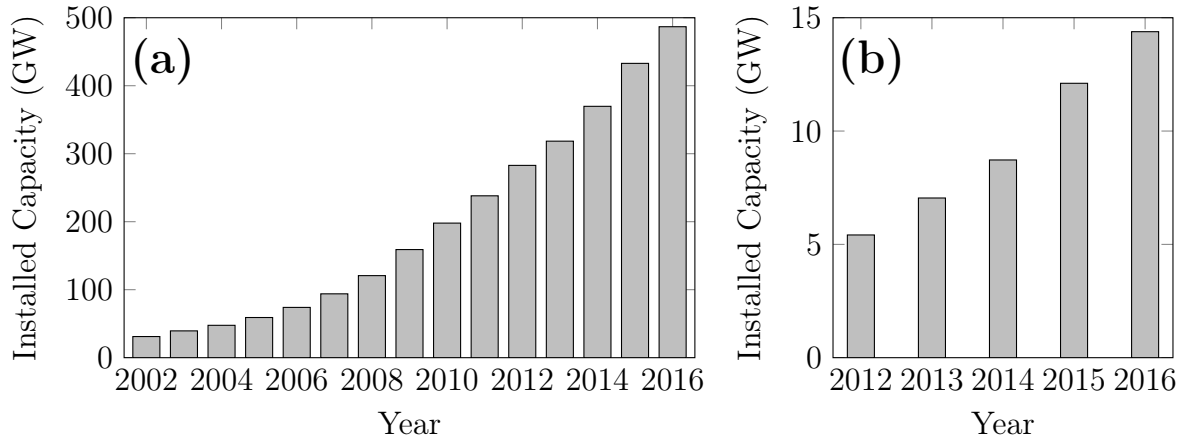


Figure 1.1: (a) Total installed wind power capacity worldwide. (b) Total installed offshore wind power capacity worldwide. The data are from [35].

power, on power system operation has widely been discussed [32, 33]. It has often been suggested that power systems are becoming more vulnerable due to higher demands and new energy policies, driving the systems to operate closer to its limits [7, 34]. Thus, improved power system monitoring and control, as well as a better understanding of the impact of the recently mainstream renewable generation such as solar and wind power, are required to ensure a continued safe operation of power systems.

1.3 Wind power

As one of the up-and-coming renewable energy sources expected to replace some of the environmentally unfriendly energy sources used today, wind power is on the rise all over the world [36]. Figure 1.1 shows histograms of the total installed wind power capacity over the last 15 years, and the total installed offshore wind power capacity in the last 5 years [35]. The figure shows that only a small fraction of the total installed capacity is installed offshore. Wind power is the fastest growing source of renewable energy [37, 38], and the largest contributor of new renewable energy (not including hydro), while solar power, which also is quickly increasing its total capacity worldwide, is ranked number two [39]. One of the biggest advantages of wind power is that the wind always blows (to some extent), giving it a higher *capacity factor* (defined as the ratio between the actual produced energy and its maximum possible energy production) than solar power [39]. On the other hand, solar power obviously only works during daytime and is not very attractive to regions where there are long periods without sunlight in the winter, when electricity is needed the most. Wind power only supports a relatively small part of the total power demand in most countries, but in countries such as Denmark with 42% of supplied demand, Germany with 60% (in four states), Spain with 18%, and Uruguay with 15%, wind power has proven itself to be a viable and reliable alternative to conventional

generation [39]. Denmark is targeting 100% generation from wind energy by 2035 [40].

Wind energy is still considered a greatly unexploited resource, and large on- and off-shore wind parks are currently planned or being constructed around the world [36, 39, 41, 42]. In Japan, the share of wind power is expected to increase in the coming years as well [43]. It is thus important to evaluate potential sites to place WFs by looking at the weather and wind conditions. The weather conditions around the islands of Japan are quite severe with seasonal typhoons and winter/summer monsoons, creating challenging conditions for WF operation, and causing variable output [44].

From a system engineering point of view, a significant wind power penetration raises several questions regarding operational and dynamical implications [45–47]. Several issues have been discussed in literature; apart from power and frequency fluctuations, the frequency response of the system changes when replacing synchronous generation with Wind Turbines (WT) asynchronously connected to the grid via converters [40, 48]. However, it has been argued recently that wind plants (or farms) can provide so-called synthetic inertia as effective as conventional generation [33]. Sufficient reactive power support is needed to ensure voltage stability [49], and new WT technology has shown to be up to par with conventional generation [50]. Furthermore, the well-known dynamic behavior of synchronous generators in power systems may change due to the introduction of wind generation [46]. This follows as a consequence to the necessary changes to the operation of power systems with large-scale renewable generation [33], which have to become more adaptable to accommodate the more variable generation. The transmission capability between different areas is a key factor in allowing this introduction [6].

The biggest challenge to a large-scale introduction of wind power in today’s systems has been variability on the scale of 1–6 hours, which impacts the frequency control that aims to balance the load and consumption [6]. For example, solar power has been argued to be less spatially variable than wind power [51, 52]; however, [53] reports on similar variability in aggregated solar and wind power. It can thus be suspected that the variability of both these renewables is highly dependent on the particular region considered. In particular, [51] reports an overall reduced variability by combining solar and wind power—indicating that solar and wind power are complementary. The dynamics of the power system depend on the loading and the stress on critical transmission lines [32], which may be affected in a highly variable system. Thus, system upgrades may be necessary [32], and the system state needs to be monitored carefully to ensure stability.

1.4 Data-driven analysis of wind power and power system dynamics

Here, in order to contextualize this research concerning KMD applications to wind power and power systems, some future challenges to power system operation posed by a potential large-scale introduction of *alternative* renewable energy sources such as solar and wind are discussed. The fundamental difference between these energy sources and conventional

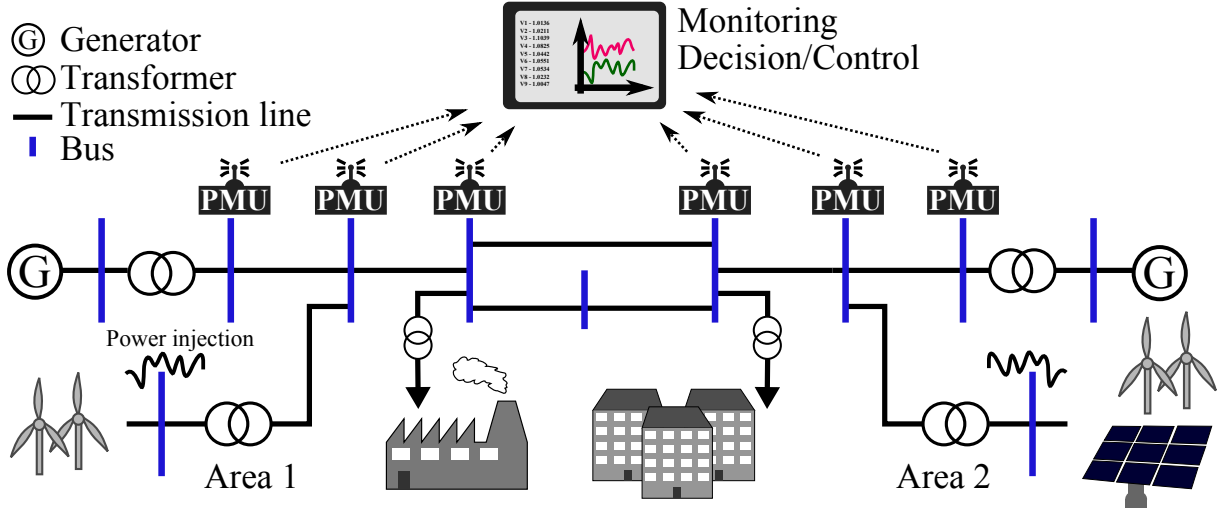


Figure 1.2: Schematic model of a power system, inspired by the two-area model [7], with Phasor Measurement Units (PMU), solar and wind power, and power system “loads” (or consumers): illustrated by residential buildings and a factory.

ones is that they are impossible to control and cannot be stored. However, wind and solar power can indeed be used to charge batteries, to power pumped hydro storages and so on [54], but so far these energy storage capabilities are miniscule compared to the total energy of the system [55].

The most fundamental requirement in power system operation is to maintain the balance between generation and consumption. Such an operational condition is called steady state. At steady state, the frequency maintains constant at the desired system frequency (50 or 60 Hz), and voltages are kept constant at nominal levels. As soon as the power consumption increases or decreases, the system frequency and voltages deviate from their nominal values. It can easily be imagined that turning on the light or brewing coffee only causes an insignificant load imbalance in any large power system. However, larger perturbations caused by, e.g., loss of generation due to unexpected faults, or loss of important transmission lines, have the potential to cause significant oscillations. Likewise, the introduction of large-scale solar and wind power creates more variable conditions [6], requiring careful monitoring to ensure safe operation. System oscillations need to be well-damped, and are in practice damped via appropriately tuned power electronics and synchronous generators [7]. Power systems are constantly subjected to various disturbances, and less damped oscillations can be identified by data analysis, and are important to monitor to ensure a stable operation [56]. Simplified power system models are used for understanding the fundamental behavior of power systems and for the operation and control of them. The basic equations describing the power system dynamics can be written as follows [7]:

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= \mathbf{f}(\mathbf{x}, \mathbf{y}), \\ \mathbf{0} &= \mathbf{g}(\mathbf{x}, \mathbf{y}), \end{aligned} \tag{1.1}$$

where $\mathbf{f}$ denotes the differential equations describing the dynamics of the system, $\mathbf{g}$ the algebraic equations governing the power balance at each bus, and $\mathbf{x}$ and $\mathbf{y}$, the state and algebraic variables, respectively. In particular, $\mathbf{g}$ comprises the following type of equations at each bus i :

$$0 = P_{\text{in},i}(\mathbf{x}, \mathbf{y}) - P_{\text{load},i}, \quad (1.2)$$

where $P_{\text{in},i}$ denotes the sum of power flows flowing into bus i , and $P_{\text{load},i}$ (simply represented by constants) the power consumed at bus i . Figure 1.2 provides a schematic view of a fictive power system, where the loads are illustrated by residential buildings and a factory. Transformers (interconnecting subsystems of different voltage levels) have been placed between the loads and the transmission system to clearly distinguish the transmission system where the major part of the power is produced and transmitted, from the subtransmission and distribution systems to which consumers and minor generation are connected. The most basic differential equation of (1.1) is the one describing the dynamics of angular frequencies ω_j of generators by their deviations from the system frequency:

$$M_j \frac{d\omega_j}{dt} = P_{\text{m},j} - P_{\text{e},j}(\mathbf{x}, \mathbf{y}), \quad (1.3)$$

where $P_{\text{m},j}$ is the mechanical input to the generator, $P_{\text{e},j}$ the electrical power injected to the system, ω_j the rotational frequency deviation of the generator, M_j an inertia constant, and j the index of a generator. At steady state operation, (1.2) holds for all i , mechanical and electric power will balance in (1.3) for all generators j , and thus $d\omega_j/dt = 0$, i.e. the rate of frequency change is zero (operation at the system's frequency). However, an imbalance at any of the buses i in (1.2) would require small changes to the generators' operating states to maintain the power balance at each bus. A large disruption to the power balance, e.g. caused by a fault, would result in transient responses of the generators which try to drive the system back to a steady state, followed by changes in $P_{\text{m},j}$ to restore the balance. Such perturbations can cause significant oscillations between the generators in the system. Knowledge of the power system's resilience to faults, and understanding of its dynamics (e.g. oscillatory behavior) following faults, can be obtained from analysis of the system of equations (1.1). A power system should be capable of withstanding any single large fault, called the N-1 criterion [57].

Measurements from power systems following faults or from normal operation conditions, e.g. supposedly provided by the so-called Phasor Measurement Units (PMU) [58] (see Fig. 1.2), can be used to extract the essential dynamic features in the power system dynamics, useful for monitoring and control (e.g. [19, 20, 59, 60]). PMUs provide synchronized high frequency measurements ideal for wide-area monitoring and control. The advantage of data-driven monitoring and control is mainly two-fold. First, a model such as (1.1) is likely a very rough estimation of the actual power system dynamics and cannot possibly capture all dynamic features that the system might exhibit without being infinitely detailed. Second, in realistic power systems, the number of equations in (1.1) becomes large, leading to a significant computational burden. Furthermore, a power system incorporating significant amounts of renewables becomes more variable [33], resulting in different dynamic characteristics during different operating conditions.

To reduce the overall power variability in a power system with significant amounts of renewables and to secure a reliable and steady supply of energy at all times, it is important to distribute the renewables well across the system. Recently, it was argued that a well distributed portfolio of alternative renewables (such as wind and solar power) together with hydro power could entirely support power systems as large as the Nordic system (Sweden, Norway, Denmark, and Finland) [61]. If we consider the hypothetical power system in Fig. 1.2, it can easily be imagined that it would be an operationally difficult task if the renewables in ‘Area 1’ and ‘Area 2’ were completely correlated. A sudden large increase or decrease in the generation would be challenging to the frequency regulation. If wind suddenly stops blowing, all power not delivered by the WFs need to be compensated by other generation. Without a large energy storage capacity, operation with high-level penetration of renewables would be impossible. Fortunately, depending on the distance between the WFs in the two areas, the WF outputs exhibit more or less correlation at different time-scales. If separated enough, they might be uncorrelated, which is a desirable property for power system operators.

Due to the highly intermittent and uncontrollable nature of renewable energy resources, and difficulties in predicting them, they create additional challenges for power system operation and planning. If these challenges are not well understood or dealt with adequately, unexpected changes in the production can potentially lead to instability if the power system is operating close to its limits.

1.5 Dissertation objectives and overview

This dissertation covers several topics on data-driven analysis of power systems and wind power systems via KMD, and addresses the following three main tasks:

- Analysis of smoothing effects (reduction of variability) of aggregated large-scale wind power production in Japan by utilizing highly-resolved weather predictions (Chapter 7).
- Short-term prediction of wind speeds (Chapter 8).
- Analysis of power system dynamics following disturbances using KMD, and proposal of an application related to power system protection (Chapter 9).

Also, this dissertation makes the following contributions to the study of dynamical systems via KMD:

- A unified interpretation of KMD-techniques is proposed which clarifies their relation and provides clues to the techniques’ numerical performance.
- Evaluation of KMD-techniques with various experimental data with support from the new interpretation. This is invaluable to potential practitioners of KMD.

- Several practical applications of KMD in power systems engineering are explored and demonstrated.

Chapter 2 reviews Koopman operator theory and KMD. KMD is based on spectral analysis of the Koopman operator that governs the evolution of observables of nonlinear dynamical systems. The operator is infinite dimensional but linear, and corresponds to the left-adjoint of the PF operator. KMD is related to the spectral decomposition of this operator in terms of eigenvalues, modes, and eigenfunctions. The numerical methods to approximate KMD all perform a spectral decomposition (the discrete spectrum) from data into a finite number of eigenvalues and modes, and are often called DMD. The chapter reviews the state-of-the-art of KMD by outlining and comparing different KMD and DMD algorithms and by clarifying their relation.

Chapter 3 evaluates the numerical performance of three KMD/DMD algorithms by various types of experimental data. For example, it is shown how the sampling frequency and temporal length of the data acquisition window affect the performances. The results are interpreted with support from a new unified interpretation of the algorithms. It is shown that the so-called vector Prony analysis [62] works well for a small spatial dimension, where standard DMD is not suitable.

Chapter 4 explains the fundamentals of wind power generation and modeling. It outlines and explains the basic modeling of WTs and WFs that are utilized in the rest of the dissertation. It also provides an example of simulation of an offshore WF incorporating weather simulation data.

Chapter 5 evaluates short-term wind power fluctuations from WTs with real data from a WF located in Ibaraki Prefecture, Japan. By incorporating a dynamic model which is shown to reproduce much of the output fluctuations, the nature and origin of the fluctuations are clarified by a statistical evaluation. Furthermore, it is clarified to what extent the aggregated WT fluctuations are mitigated on a WF-scale.

Chapter 6 deals with larger scale wind power fluctuations, that are analyzed by utilizing a weather prediction model called Cloud Resolving Storm Simulator (CReSS) [63], which is also described here. Both coarser 2-km simulations encompassing the whole of Japan and extremely detailed 200-m resolution simulations of a smaller domain are evaluated. The predictions are compared against real measurements. It is shown that wind power variability—here in terms of increments [27, 64]—can be assessed on the time-scale of several hours. Improvements in the prediction of wind power increments are demonstrated for hourly averaged power, which suggests CReSS’s ability to predict wind power ramps on an hourly scale. The measured and predicted wind spectra are found similar in both the long- and short-term. The results for the detailed simulations indicate that CReSS captures the so-called spectral coherence of wind on the scale of one to a few kilometers.

Chapter 7 investigates smoothing effects of wind power in Japan by incorporating wind speed predictions from CReSS. Wind power has increased rapidly worldwide in recent years. As a consequence, it becomes important for power system and WF operators to grasp the amount of smoothing achieved in the aggregated power, and to grasp at what temporal and spatial scales smoothing is achieved. A new smoothing index for

wind power based on KMD is proposed, which can be regarded as a generalization of a previously proposed index based on power spectral densities. WFs around six regions in northern Honshu, Japan, are considered as a test case for the evaluation. By applying the index to simulated wind power, it is shown how the smoothing improves by distributing WFs over different regions. The results indicate that by distributing WFs over only one to three regions, smoothing results vary considerably depending on the choice of regions. However, as the number of considered regions increases, the smoothing improves, and the particular choice of regions matters less for smoothing effects at the investigated time-scales. These results highlight the importance of deliberately selecting sites for large-scale wind power production to more effectively smooth the aggregated power.

Chapter 8 discusses KMD with applications to short-term wind speed prediction and optimal power extraction of WTs. The most fundamental and important task of a WT is to extract as much power as possible and feed the power into the power grid. To do so, MPPT (Maximum Power Point Tracking) [65] techniques have been developed to control the turbine’s rotor speed as close to its optimal value as possible. The optimal rotor speed depends on the ratio between the speed at the tip of the turbine’s blades and the incoming wind, called Tip Speed Ratio (TSR) [66]. However, the wind speed changes quickly, and a turbine cannot change its speed as fast as the wind due to its inertia. This will render the turbine unable to follow the fast changes, and consequently unable to perform at its optimal TSR. In this chapter, it is demonstrated how KMD can be used to predict the wind speed a few seconds in the future which has applications in MPPT control, and how KMD relates to commonly used prediction methods.

Chapter 9 discusses applications of KMD related to power system security. KMD is applied to data from test systems and the capability of partitioning a power network—i.e. to determine the points of separation in a so-called controlled islanding strategy in power system emergency control—is demonstrated. A practical data-driven algorithm incorporating KMD is proposed for network partitioning. Comparisons are made with two techniques previously applied to the network partitioning problem: spectral graph theory [67], which is based on the eigenstructure of the graph Laplacian [68,69], and slow-coherency [70], which identifies coherent groups of generators for a specified number of linear low-frequency modes. The partitioning results share common features with results obtained both with graph theory and slow-coherency-based techniques. The suggested partitioning method is evaluated with two test systems, and similarities between Koopman modes and Laplacian eigenvectors are shown numerically and elaborated theoretically.

Chapter 10 summarizes and concludes this dissertation, which explores several different aspects of data-driven analysis and control of wind power and power systems via KMD. Through a number of applications, a comprehensive theoretical background, and an evaluation of KMD-techniques, the applicability of KMD to various problems and types of data was clarified. It is hoped that this will stimulate further research efforts using KMD to improve the monitoring and security of power systems, and increase our understanding of the characteristics of aggregated wind power. This chapter also discusses several direct extensions of this work related to wind power and power system applications.

Chapter 2

On the Koopman operator and the Koopman mode decomposition

The Koopman operator [12] is closely related to the so-called transfer operator, often called Perron-Frobenius (PF) operator [11], which has been used to characterize the behavior and evolution of dynamical systems [9–11]. Until recently, the PF operator approach has been dominant in comparison to the dual view provided by the Koopman operator. The Koopman operator is the left adjoint of the PF operator, hence closely related, mathematically speaking. Consequently, they can be regarded as two different but equivalent viewpoints of the evolution of dynamical systems [13]. In particular, the Koopman operator governs the evolution of observable functions of the dynamical system. This is different to the standard description of evolution of dynamical systems according to a map acting on the *state variables*, or a *flow* describing the time-evolution from an initial state. An example of an observable function is the voltage at a bus in a power system, which is a nonlinear function of a number of variables such as the voltages of adjacent buses and state variables of generators that determine the power flow in the system. One can also construct appropriate observables as functions of measurements. This has shown to be an effective approach for estimating properties of the Koopman operator, and which will be discussed in this chapter. Despite the system being nonlinear and finite dimensional, the Koopman operators are linear but infinite dimensional. Linearity is an extremely attractive property since it allows for direct use of all the well-known and efficient tools developed in linear algebra, operator theory, and scientific computing.

In the last few years, a great deal of work has been done towards data-derived approximations of the Koopman operator, and a substantial amount of new theory and mathematical tools have been presented. However, obtaining an explicit representation of the operator, which in general is infinite-dimensional, is not always feasible or even desired. In the data-driven setting, a finite-dimensional approximation of the action of the operator is defined and estimated on the subspace spanned by observables (measurements or functions of measurements). Many publications have shown that these finite-dimensional approximations are often capable of capturing dominant features of the Koopman operator, sometimes relying only on very simple and efficient mathematical tools. The recent

intensified academic interest in Koopman operator theory, numerical approximation, and applications can be attributed to the seminal paper [14] by Mezić, which derived the so-called Koopman Mode Decomposition (KMD), i.e. a spectral decomposition of the operator in terms of eigenvalues, eigenfunctions, and eigenvectors (modes) (definitions of these are given in the forthcoming sections). A few years after that in [15], a method called Dynamic Mode Decomposition (DMD) was discovered as a numerical method to approximate the KMD. Related numerical methods proposed afterwards have since then most often adopted the name DMD. The early review paper [13] has been instrumental in the KMD-development by collecting and describing the early applications of KMD, and by providing a solid theoretical background for other researchers to build on. The Koopman-operator-approach to the analysis and control of dynamical systems has shown to be a timely and appealing framework in the current world of ever increasing computational capability to store and process “big data”. Applications of DMD have ranged from fluid mechanics [15,24], thermal dynamics in buildings [71], power systems [19–21], neural activity [72], and molecular kinetics [73], to name a few.

One might wonder why one should focus on the Koopman operator instead of the PF operator, which governs the evolution of densities (probabilities), instead of observable functions of the system—in particular since these two approaches are *arguably* theoretically equivalent. There are several reasons for this. First, it has been believed that approximating the PF operator would require many short simulations of a system with different initial conditions, while an estimation of the Koopman operator or KMD only relies on one longer run [13]. An example of such a simple method to approximate the PF operator is Ulam’s method [74,75], which is described in Section 2.3.1. Second, the evolution of dynamical systems in terms of observable functions governed by the Koopman operator is close to the standard mathematical setting in analysis and control of linear systems. Hence, all the well-established tools in control of linear systems have the potential of being transferred to the Koopman framework. Third, a series of numerically efficient algorithms have been linked to the Koopman operator in recent years. Interestingly, it has been shown recently in [75] that a slight modification to the state-of-the-art procedure of DMD indeed produces approximations of the eigenfunctions and eigenvalues of the PF operator.

This dissertation investigates KMD/DMD applied to data from power systems and wind power systems, and in contrast to applications such as fluid dynamics which often involve huge spatio-temporal data sets, often only a quite limited number of time-series or measurement locations is available. In this setting, the right choice of DMD-method is shown to be crucial. An evaluation of the numerical performance of a few related DMD algorithms are provided in the next chapter. This chapter, on the other hand, provides the theoretical background of the Koopman operator and KMD, and discusses the ideas underlying all numerical KMD approximations following [13]. The theory is augmented with descriptions of several different variants of DMD, including *Extended* DMD (EDMD). Additionally, the PF operator and its connection to the Koopman operator are briefly discussed. Finally, we provide two examples exemplifying the strengths and challenges in numerical KMD, and show how Koopman and PF can both be calculated from data following [75].

The remainder of the chapter is organized as follows. Section 2.1 explains the ideas of the Koopman operator and introduces KMD. Section 2.2 outlines the ideas underlying the numerical approximation of KMD and provide descriptions of standard DMD-algorithms. The PF operator and its relation to the Koopman operator are described in Section 2.3. In Section 2.4, we illustrate the core ideas of the Koopman approach by applying the KMD algorithms to a linear system and a simple nonlinear system inspired by [17, 76]. The connection between the Koopman and the PF operators is also discussed. Finally, conclusions are given in Section 2.5.

2.1 Introduction to the Koopman operator and the Koopman mode decomposition

This section will introduce the Koopman operator and the KMD. Now, consider a discrete-time dynamical system evolving on a finite-dimensional state space $\mathbb{X}$: for $\mathbf{x} \in \mathbb{X}$,

$$\mathbf{x}_{k+1} = \mathbf{T}(\mathbf{x}_k), \quad (2.1)$$

where $\mathbf{T} : \mathbb{X} \rightarrow \mathbb{X}$ is a nonlinear map governing the evolution of $\mathbf{x}$ at each discrete time $k \in \mathbb{Z}$, to its next state $\mathbf{x}_{k+1}$. The so-called *observable* f is now introduced as a scalar-valued function defined on the state space $\mathbb{X}$, namely

$$f : \mathbb{X} \rightarrow \mathbb{C}.$$

The observable f is a mathematical formulation of a measurement of the system dynamics as a function of its state: i.e. $f(\mathbf{x})$, and can e.g. be viewed as an *output* function in control systems theory. Now, we denote the space of observables by $\mathcal{F}$. For an observable $f \in \mathcal{F}$, the *Koopman operator* $\mathbf{U}$ for the discrete-time system (2.1) maps f into a new function, by *composing* it with the map T as follows:

$$\mathbf{U}f := f \circ \mathbf{T}.$$

That is, the Koopman operator $\mathbf{U}$ governs the evolution of the observable f along the state's dynamics $\mathbf{x}_k \rightarrow \mathbf{x}_{k+1}$:

$$(\mathbf{U}f)(\mathbf{x}_k) = (f \circ \mathbf{T})(\mathbf{x}_k) = f(\mathbf{T}(\mathbf{x}_k)) = f(\mathbf{x}_{k+1}).$$

Although the dynamical system (2.1) can be nonlinear and evolves in the finite-dimensional space $\mathbb{X}$, the Koopman operators are *linear* but *infinite-dimensional*.

KMD is the analysis of the time evolution of observables under iteration of the nonlinear system (2.1) through the linear operator $\mathbf{U}$. Since $\mathbf{U}$ is linear, eigenvalues $\lambda_j \in \mathbb{C}$ and eigenfunctions $\phi_j \in \mathcal{F}$ can be defined as follows:

$$\mathbf{U}\phi_j = \lambda_j\phi_j, \quad j = 1, 2, \dots$$

where the functions ϕ_j are referred to as *Koopman eigenfunctions*, and λ_j as the associated *Koopman Eigenvalues* (KEs). Now consider a vector-valued observable $\mathbf{f} = [f_1, \dots, f_m]^\top : \mathbb{X} \rightarrow \mathbb{C}^m$ for $f_j \in \mathcal{F}$. If each of the m components of $\mathbf{f}$ lies within the span of Koopman eigenfunctions ϕ_j , as in [15], we may expand $\mathbf{f}$ in terms of the eigenfunctions as

$$\mathbf{f} = \sum_{j=1}^{\infty} \phi_j \mathbf{V}_j,$$

where $\mathbf{V}_j$ are vector-valued coefficients in the expansion. The time evolution $\mathbf{f}(\mathbf{x}_k)$ from $\mathbf{f}(\mathbf{x}_0)$ is then exactly given in terms of $\{\lambda_j, \phi_j, \mathbf{V}_j\}$ by

$$\mathbf{f}(\mathbf{x}_k) = (\underline{\mathbf{U}}^k \mathbf{f})(\mathbf{x}_0) = \sum_{j=1}^{\infty} \lambda_j^k \phi_j(\mathbf{x}_0) \mathbf{V}_j, \quad (2.2)$$

where

$$\underline{\mathbf{U}}^k \mathbf{f} := [\mathbf{U}^k f_1, \dots, \mathbf{U}^k f_m]^\top.$$

The vector $\mathbf{V}_j$ is called the j -th *Koopman Mode* (KM) with respect to the observable $\mathbf{f}$. Thus, the spectral decomposition (2.2) of the Koopman operator in terms of eigenvalues, eigenfunctions, and KMs is called the KMD [23].

2.2 KMD algorithms

Here we review the mathematical formulation of KMD based on [13], and the two original algorithms of KMD: Arnoldi-like KMD [15], and standard/exact Dynamic Mode Decomposition (DMD) [16, 17, 77] (the “exact” DMD is an extension the original algorithm). An overview of a recently proposed extension of DMD, called EDMD [18], is also provided. In particular, we show that the underlying idea of these algorithms is the way the Koopman operator $\mathbf{U}$ is approximated by a matrix on a finite-dimensional subspace constructed directly from data.

2.2.1 Mathematical formulation

The objective of all KMD algorithms is to find a finite-dimensional approximation of the action of the Koopman operator directly from data. The description that follows in this subsection is based on [13]. Now consider a fixed vector-valued observable $\mathbf{f} = [f_1, \dots, f_m]^\top \in \mathcal{F}^m := \underbrace{\mathcal{F} \times \dots \times \mathcal{F}}_m$ and associated finite-time data (snapshots) as

$$\{\mathbf{y}_k = \mathbf{f}(\mathbf{x}_k); k = 0, \dots, N-1, N < \infty\}. \quad (2.3)$$

Then, the following function space (called the *Krylov* subspace in [13]) $\mathcal{K}_N$ is sufficient for analysis of the finite-time data:

$$\mathcal{K}_N := \text{span}\{\underline{\mathbf{U}}^k \mathbf{f}; k = 0, \dots, N-1\}. \quad (2.4)$$

In the following, we will assume that $\{\underline{\mathbf{U}}^k \mathbf{f}; k = 0, \dots, N-1\}$ is a linearly independent set so that these functions form a basis for $\mathcal{K}_N$. Here, let $\mathbf{P}_N : \mathcal{F}^m \rightarrow \mathcal{K}_N$ be a projection from the space of vector-valued observables onto $\mathcal{K}_N$. Thus, the following operator

$$\mathbf{P}_N \underline{\mathbf{U}} : \mathcal{K}_N \rightarrow \mathcal{K}_N, \quad (2.5)$$

is finite-dimensional and has a matrix representation, $\mathbf{C}_N : \mathbb{C}^N \rightarrow \mathbb{C}^N$, in terms of the $\{\underline{\mathbf{U}}^k \mathbf{f}\}$ -basis. Let $(\lambda, \mathbf{V}_\lambda)$ be a pair of eigenvalue and associated eigenvector of $\mathbf{C}_N$, where $\mathbf{V}_\lambda = [V_{\lambda,0}, \dots, V_{\lambda,N-1}]^\top \in \mathbb{C}^N$, then λ is an eigenvalue of $\mathbf{P}_N \underline{\mathbf{U}}$ and $\Phi_\lambda := \sum_{j=0}^{N-1} V_{\lambda,j} (\underline{\mathbf{U}}^j \mathbf{f})$ the associated eigenfunction [13]. By constructing a finite-dimensional subspace $\mathcal{K}_N$ by the vector-valued observable $\mathbf{f}$, the problem of finding eigenvalues and modes of the Koopman operator has been reformulated into the problem of finding the matrix $\mathbf{C}_N$ and calculating its eigenvalues and eigenvectors. The matrix $\mathbf{C}_N$ can be considered as a finite-dimensional approximation of the action of the Koopman operator on the subspace $\mathcal{K}_N$. All KMD algorithms are procedures for computing approximations of the action of the Koopman operator directly from data (e.g. via the matrix $\mathbf{C}_N$).

2.2.2 The Arnoldi-like algorithm

Now, we look at how the so-called Arnoldi-like algorithm [15] constructs $\mathbf{C}_N$ from data. We define a subspace or matrix of input data as follows:

$$\mathbf{K}_{N-1} := [\mathbf{y}_0, \dots, \mathbf{y}_{N-2}], \quad (2.6)$$

constructed by snapshots $\mathbf{y}_k \in \mathbb{R}^m$, where $\mathbf{K}_{N-1} \in \mathbb{R}^{m \times (N-1)}$. The dimension m is called the *spatial dimension* here and denotes the number of measurement locations. A vector of constants, $\mathbf{c} = [c_1, \dots, c_{N-1}]^\top \in \mathbb{R}^{N-1}$, is then defined in the following formula:

$$\mathbf{y}_{N-1} = \mathbf{K}_{N-1} \mathbf{c} + \mathbf{r}, \quad \mathbf{r} \perp \text{span}\{\mathbf{K}_{N-1}\}, \quad (2.7)$$

where $\mathbf{r} \in \mathbb{R}^m$ is a residual vector. If $\mathbf{y}_{N-1}$ is in the span of columns of $\mathbf{K}_{N-1}$, there exist constants c_i such that $\mathbf{r}$ vanishes. Otherwise, $\mathbf{c}$ is determined such that the $\mathbb{C}^m$ -norm of $\mathbf{r}$ is minimized, which coincides with the choice of the projection operator $\mathbf{P}_{N-1}$ such that $\mathbf{P}_{N-1} \underline{\mathbf{U}}^{N-1} \mathbf{f}(\mathbf{x}_0)$ is the least-squares approximation to $\underline{\mathbf{U}}^{N-1} \mathbf{f}(\mathbf{x}_0)$ [13]. It should be noted that for data with low spatial dimension, namely $m \ll N$, the rank of $\mathbf{K}_{N-1}$ is at most m , which inevitably results in a rank-deficiency when computing $\mathbf{c}$. This implies no uniqueness of $\mathbf{c}$, and that the derived decomposition and dynamic quantities (frequencies, modes, etc.) and their physical interpretation should be carefully cross-checked against other methods. One of the solutions is $\mathbf{c} = \mathbf{K}_{N-1}^\dagger \mathbf{y}_{N-1}$ [78], where $\dagger$ stands for the Moore-Penrose pseudo-inverse. This computation outputs the vector $\mathbf{c}$ which is the optimal choice in a least squares sense. Then, denote the *companion matrix* [15] by $\mathbf{C}_{N-1}$,

$$\mathbf{C}_{N-1} := \begin{bmatrix} 0 & 0 & \dots & 0 & c_0 \\ 1 & 0 & & 0 & c_1 \\ 0 & 1 & & 0 & c_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & c_{N-1} \end{bmatrix},$$

and locate its $N - 1$ eigenvalues, also called empirical Ritz values, $\tilde{\lambda}_1, \dots, \tilde{\lambda}_{N-1}$. Finally, we construct the so-called *Vandermonde* matrix [15] with $\tilde{\lambda}_j$,

$$\mathbf{T} := \begin{bmatrix} 1 & \tilde{\lambda}_1 & \tilde{\lambda}_1^2 & \dots & \tilde{\lambda}_1^{N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \tilde{\lambda}_{N-1} & \tilde{\lambda}_{N-1}^2 & \dots & \tilde{\lambda}_{N-1}^{N-2} \end{bmatrix}.$$

The empirical Ritz vectors $\tilde{\mathbf{V}}_j$ are the columns of $\mathbf{V}_{\mathbf{T}} := \mathbf{K}_{N-1} \mathbf{T}^{-1}$. The data $\mathbf{y}_k$ can then be decomposed as

$$\left. \begin{aligned} \mathbf{y}_k &= \sum_{i=1}^N \tilde{\lambda}_i^k \tilde{\mathbf{V}}_i, \quad k = 0, \dots, N-2, \\ \mathbf{y}_{N-1} &= \sum_{i=1}^N \tilde{\lambda}_i^N \tilde{\mathbf{V}}_i + \mathbf{r}. \end{aligned} \right\} \quad (2.8)$$

Here, although not mathematically rigorous, we have

$$\begin{aligned} \underline{\mathbf{U}} \mathbf{K}_{N-1} &= [\mathbf{y}_1, \dots, \mathbf{y}_{N-1}] \\ &= \mathbf{K}_{N-1} \mathbf{C}_{N-1} + \mathbf{r}[0, \dots, 0, 1], \end{aligned} \quad (2.9)$$

which can more simply be written as:

$$[\mathbf{y}_1, \dots, \mathbf{y}_{N-1}] = \mathbf{K}_{N-1} \mathbf{C}_{N-1} + \mathbf{r}[0, \dots, 0, 1], \quad (2.10)$$

or with the following definitions:

$$\mathbf{X} := \mathbf{K}_{N-1} = [\mathbf{y}_0, \dots, \mathbf{y}_{N-2}], \quad \mathbf{Y} := [\mathbf{y}_1, \dots, \mathbf{y}_{N-1}], \quad (2.11)$$

we have

$$\mathbf{Y} = \mathbf{X} \mathbf{C}_{N-1} + \mathbf{r}[0, \dots, 0, 1]. \quad (2.12)$$

This implies that if $\mathbf{r}$ is small, then $\mathbf{C}_{N-1}$ is thought of as an approximation of the action of the Koopman operator on the associated finite-dimensional space $\mathcal{K}_N$.

2.2.3 Dynamic mode decomposition

Here, we look at probably the most widely used KMD-algorithm: DMD. In [16] the author considers the matrices of snapshots in (2.11), and conducts the Singular Value Decomposition (SVD) [79] of $\mathbf{X} = \mathbf{K}_{N-1}$, given by

$$\mathbf{X} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^*,$$

where $*$ denotes the complex conjugate transpose, $\mathbf{U}$ is an $m \times m$ matrix with orthonormal columns, $\mathbf{\Sigma} := \text{diag}(\sigma_1, \dots, \sigma_m)$ an m -th order diagonal matrix with r non-zero singular values σ , and $\mathbf{V}$ an $(N - 1) \times m$ matrix with orthonormal columns. Note that $r :=$

$\text{rank}(\mathbf{X}) \leq m$. We have $\mathbf{U} = \mathbf{X}\mathbf{V}\Sigma^{-1}$ from the SVD if $r = m$. Furthermore, from (2.9) and (2.10) we have that

$$\mathbf{U}^* \underline{\mathbf{U}} \mathbf{U} = \mathbf{U}^* \underline{\mathbf{U}} \mathbf{X} \mathbf{V} \Sigma^{-1} = \mathbf{U}^* \mathbf{Y} \mathbf{V} \Sigma^{-1}, \quad (2.13)$$

where we have made use of $\underline{\mathbf{U}} \mathbf{X} = \mathbf{Y}$. The right-hand side of (2.13) can be computed directly from data and is hence thought of as an approximation of the action of the Koopman operator in the space projected with $\mathbf{U}$, that is, in m -dimensional space. However, for data $\{\mathbf{y}_k\}_{k=1}^N$ with low spatial dimension ($m \ll N$), the approximation might not be enough to extract the dynamical characteristics (frequencies, growth rates, and modes), in comparison with the Arnoldi-like KMD.

Following [17, 77], we approximate the evolution of the nonlinear system (2.1) in terms of our snapshots as follows:

$$\mathbf{Y} = \bar{\mathbf{A}} \mathbf{X}, \quad (2.14)$$

where $\bar{\mathbf{A}} : \mathbb{C}^m \rightarrow \mathbb{C}^m$. The so-called exact DMD [17] is then the solution $\bar{\mathbf{A}} = \mathbf{Y} \mathbf{X}^\dagger$, where the pseudo-inverse is given by $\mathbf{X}^\dagger = \mathbf{V} \Sigma^{-1} \mathbf{U}^*$. Eigenanalysis of $\bar{\mathbf{A}}$ yields the sought decomposition, and its time-evolution is given by:

$$\mathbf{y}_k = \sum_{i=1}^m (\lambda'_i)^k (\hat{\phi}_i \mathbf{y}_0) \phi_i = \sum_{i=1}^m (\lambda'_i)^k \varphi_i(\mathbf{y}_0) \phi_i, \quad (2.15)$$

where $\lambda'_i \in \mathbb{C}$ are eigenvalues of $\bar{\mathbf{A}}$, and $\phi_i \in \mathbb{C}^m$ the so-called DMD modes defined as follows: $\bar{\mathbf{A}} \phi_i = \lambda'_i \phi_i$. $\hat{\phi}_i$ are the adjoint modes defined as follows: if $\Phi = [\phi_1, \dots, \phi_m]$, then $\hat{\Phi} = \Phi^{-1} = [\hat{\phi}_1^\top, \dots, \hat{\phi}_m^\top]^\top$. The eigenfunctions $\varphi_i(\mathbf{y}) := \hat{\phi}_i \mathbf{y}$ can be thought of as approximations of the Koopman eigenfunctions (ϕ_j , see (2.2)) associated with eigenvalue λ'_i . For large m , a more computationally effective, reduced-order DMD is described in [17], which calculates the matrix $\tilde{\mathbf{A}}$, defined as the right-hand side of (2.13), i.e. $\tilde{\mathbf{A}} = \mathbf{U}^* \mathbf{A} \mathbf{U} = \mathbf{U}^* \mathbf{Y} \mathbf{V} \Sigma^{-1}$, and whose eigenvectors are related to $\bar{\mathbf{A}}$'s via a linear transformation [17]. More precisely, the iteration of data in the reduced DMD is given by

$$\tilde{\mathbf{y}}_{k+1} = \tilde{\mathbf{A}} \tilde{\mathbf{y}}_k, \quad (2.16)$$

where $\tilde{\mathbf{y}}_k \approx \mathbf{U}_r^* \mathbf{y}_k$ is the relation between projected snapshots $\tilde{\mathbf{y}}_k$ and the original snapshots. The following matrices: $\mathbf{U}_r \in \mathbb{C}^{m \times r}$, $\mathbf{V}_r \in \mathbb{C}^{(N-1) \times r}$ corresponding to the r first columns of $\mathbf{U}$ and $\mathbf{V}$, and $\Sigma_r = \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{C}^{r \times r}$ have been utilized in (2.13) for the computation of $\tilde{\mathbf{A}}$, where r corresponds to the most dominant (large) singular values. The choice of r has been suggested [77] to be determined via the optimal hard threshold method in [80]. This procedure can also be viewed as a filtering of the input data according to: $\tilde{\mathbf{X}} = \mathbf{U}_r \Sigma_r \mathbf{V}_r^*$, with $\tilde{\mathbf{X}}$ denoting the filtered input data. Since this dissertation deals with considerably small m , the exact DMD is used, if not explicitly stated otherwise.

2.2.4 Equivalence between Arnoldi-KMD and DMD

In the following, two theorems on the equivalence between Arnoldi-KMD and DMD are provided. Note that the SVD is a so-called “economy-sized” SVD, which can e.g. be

computed with the command $[U, \Sigma, V] = \text{svd}(X, 'econ')$ in MATLAB, where X is the data matrix.

Definition 2.2.1 Two data matrices of snapshots are defined as: $X = [\mathbf{y}_1, \dots, \mathbf{y}_N] \in \mathbb{R}^{m \times N}$, and $Y = [\mathbf{y}_2, \dots, \mathbf{y}_{N+1}] \in \mathbb{R}^{m \times N}$. Then, we assume the following: A is given by $A = YX^\dagger$, such that we exactly have $Y = AX$ by (2.14); and $Y = XC$ by (2.10) exactly holds. There are two pseudoinverses of X : right $X_r^\dagger$, and left $X_l^\dagger$. $X_r^\dagger$ is non-singular if $m \leq N$, and there are m independent rows. $X_l^\dagger$ is non-singular if $N \leq m$, and there are N independent columns. Multiplying X with $(X_r^\dagger, X_l^\dagger)$ from the left and right side, respectively, gives the identity matrix I_n of dimension n as follows: $XX_r^\dagger = I_m$, and $X_l^\dagger X = I_N$.

Theorem 2.2.1 The eigenvalues of A and C are identical if $m = N$ and X is of full rank.

Proof 2.2.1 From the definition, we have $AX = XC$, which can be rewritten as $A = XCX_r^\dagger$. Eigenvalues λ_i and eigenvectors $\mathbf{v}_i$ of A are given by: $A\mathbf{v}_i = XCX_r^\dagger \mathbf{v}_i = \lambda_i \mathbf{v}_i$, and hence $CX_l^\dagger \mathbf{v}_i = \lambda_i X_l^\dagger \mathbf{v}_i$. X_r and X_l are equal if and only if $m = N$, and X is of full rank. The eigenvectors of C associated with λ_i are thus given by $X_l^\dagger \mathbf{v}_i = X_r^\dagger \mathbf{v}_i$, and the eigenvalues of C and A are identical. ■

While the above theorem concerns a rather special case, a more general theorem is proven below, for the case that the spatial dimension m exceeds the number of snapshots N in the matrices X and Y .

Theorem 2.2.2 The eigenvalues of C are eigenvalues of A if $m > N$ and $\text{rank}(X) = N$. The remaining $m - N$ eigenvalues of A are zero.

Proof 2.2.2 Because $m > N$, we have that $\text{rank}(C) = N < m$, and we cannot assume that the residual $\mathbf{r}$ vanishes. Thus, in contrast to above, we now have

$$AX = XC + R, \quad (2.17)$$

where $R = \mathbf{r}[0, \dots, 1]$ is a matrix containing the residual vector $\mathbf{r}$ in its last column. Further, $A \in \mathbb{R}^{m \times m}$ is of rank N , same as the rank of the reduced matrix $\tilde{A} = U^*AU$. Multiplying (2.17) with $X_l^\dagger = V\Sigma^{-1}U^*$ from the left, we have

$$V\Sigma^{-1}U^*AX = C + V\Sigma^{-1}U^*R,$$

where $U^*R = 0$ ($N \times N$ matrix of zeros) due to $\mathbf{r} \perp X$. Then, by multiplication from left with U^*Y we get

$$\tilde{A}U^*Y = U^*YC,$$

where we have utilized that $\tilde{A} = U^*YV\Sigma^{-1}$ and $Y = AX$, and thus we have

$$\tilde{A}U^*Y\mathbf{v}_i = U^*YC\mathbf{v}_i = \lambda_i U^*Y\mathbf{v}_i,$$

where $C\mathbf{v}_i = \lambda_i \mathbf{v}_i$, implying that the eigenvalues λ_i ($i = 1, \dots, N$) are common between C and the reduced DMD-matrix $\tilde{A}$. Since $\tilde{A} = U^*AU = U^*YV\Sigma^{-1}$ holds, but $A \neq U\tilde{A}U^*$ in this case since $\text{rank}(U) = N$ leading to $UU^* \neq U^*U = I_N$, the N non-zero eigenvalues of A are associated with the singular values σ_i ($i = 1, \dots, N$), and are retained in $\tilde{A}$. ■

These theorems theoretically clarify connections between Arnoldi-KMD and DMD, and provide insight and support to the results of applying these methods to various data. Thus, they are valuable to potential practitioners of KMD/DMD.

2.2.5 Hankel-like matrices and Prony KMD

Consider the following two matrices constructed by a scalar time series $y_i \in \mathbb{R}$ ($i = 0, \dots, 2N - 1$):

$$\mathbf{H}_x = \begin{bmatrix} y_0 & y_1 & \cdots & y_{N-1} \\ y_1 & y_2 & \cdots & y_N \\ \vdots & \vdots & \ddots & \vdots \\ y_{N-1} & y_N & \cdots & y_{2N-2} \end{bmatrix}, \quad \mathbf{H}_y = \begin{bmatrix} y_1 & y_2 & \cdots & y_N \\ y_2 & y_3 & \cdots & y_{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_N & y_{N+1} & \cdots & y_{2N-1} \end{bmatrix}. \quad (2.18)$$

The matrices $\mathbf{H}_x, \mathbf{H}_y \in \mathbb{R}^{N \times N}$ are so-called Hankel matrices. We can think of the columns of $\mathbf{H}_x, \mathbf{H}_y$ as vector valued snapshots $\mathbf{y}_k \in \mathbb{R}^N$ constructed by time-shifted measurements. By the introduction of time-shifted measurements, we have constructed a larger subspace $\mathcal{K}_N$ on which the Koopman operator is projected. In fact, if y_i are measurements of the impulse response of the linear system, DMD applied to (2.18), i.e. solving $\mathbf{H}_y = \mathbf{A}\mathbf{H}_x$, is the so-called Eigensystem Realization Algorithm (ERA), a well-known technique in system identification [17].

Related to the Hankel matrix, the so-called vector Prony analysis (or KMD) based on [62] is discussed in the following. Instead of a scalar-time series, suppose now that the $2N$ vector-valued snapshots $\mathbf{y}_k \in \mathbb{R}^m$ are available as

$$\{\mathbf{y}_k = \mathbf{f}(\mathbf{x}_k); k = 0, \dots, 2N - 1, N < \infty\}.$$

Then, we define two so-called vector Hankel matrices as follows:

$$\mathbf{A}_x := \begin{bmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \cdots & \mathbf{y}_{N-1} \\ \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_N \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_{N-1} & \mathbf{y}_N & \cdots & \mathbf{y}_{2N-2} \end{bmatrix}, \quad \mathbf{A}_y := \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_N \\ \mathbf{y}_2 & \mathbf{y}_3 & \cdots & \mathbf{y}_{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_N & \mathbf{y}_{N+1} & \cdots & \mathbf{y}_{2N-1} \end{bmatrix}. \quad (2.19)$$

In the same manner as for Arnoldi-like KMD: see (2.7), a vector of constants $\mathbf{p} \in \mathbb{R}^N$ is defined in the following formula:

$$\mathbf{b} = \mathbf{A}_x \mathbf{p} + \mathbf{r}, \quad (2.20)$$

where

$$\mathbf{b} := [\mathbf{y}_N^\top, \dots, \mathbf{y}_{2N-1}^\top]^\top \in \mathbb{R}^{m \cdot N}, \quad \mathbf{r} \perp \text{span}\{\mathbf{A}_x\}.$$

The vector $\mathbf{r} \in \mathbb{R}^{m \cdot N}$ is still a residual. One solution is $\mathbf{p} = \mathbf{A}_x^\dagger \mathbf{b}$. If $\mathbf{A}_x$ is of full column rank, $\mathbf{p}$ can be uniquely solved as $\mathbf{p} = (\mathbf{A}_x^\top \mathbf{A}_x)^{-1} \mathbf{A}_x^\top \mathbf{b}$. The constants of $\mathbf{p}$ are then used in the same way as $\mathbf{c}$ in Arnoldi-like algorithm in order to derive the companion matrix

$\mathbf{C}'_N$ and associated Vandermonde matrix. Note that the iteration of the input data can be written as

$$\mathbf{A}_y = \mathbf{A}_x \mathbf{C}'_N + \mathbf{r}[0, \dots, 0, 1]. \quad (2.21)$$

Also, the following decomposition of data holds:

$$\mathbf{y}_k = \sum_{i=1}^N \hat{\lambda}_i^k \hat{\mathbf{V}}_i, \quad k = 0, \dots, N-1, \quad (2.22)$$

where $\hat{\lambda}_i \in \mathbb{C}$ is called the j -th *Prony value* and $\hat{\mathbf{V}}_i$ the associated *Prony vector*.

As with the Arnoldi-like algorithm, let us look at the geometric interpretation of vector Prony analysis. To do this, define the function space $\mathcal{K}'_N$ generated by the operation of *time-shifting* as

$$\mathcal{K}'_N := \text{span} \left\{ \begin{bmatrix} \underline{\mathbf{U}}^k \mathbf{f} \\ \underline{\mathbf{U}}^{k+1} \mathbf{f} \\ \vdots \\ \underline{\mathbf{U}}^{k+N-1} \mathbf{f} \end{bmatrix} ; k = 0, \dots, N-1 \right\}.$$

Then, by letting $\mathbf{P}'_N : \mathcal{F}^m \rightarrow \mathcal{K}'_N$ be a projection from the original space $\mathcal{F}^m$ onto $\mathcal{K}'_N$, the following operator

$$\mathbf{P}'_N \underline{\mathbf{U}} : \mathcal{K}'_N \rightarrow \mathcal{K}'_N,$$

is finite-dimensional and has a matrix representation with the same dimension of $\mathbf{C}_N$ in the $\{[(\underline{\mathbf{U}}^k \mathbf{f})^\top, \dots, (\underline{\mathbf{U}}^{k+N-1} \mathbf{f})^\top]^\top\}$ -basis, which we have denoted by $\mathbf{C}'_N : \mathbb{C}^N \rightarrow \mathbb{C}^N$. Needless to say, it has been implicitly assumed that the N new functions (“fictive” vector-valued observables) are linearly independent; which would be more beneficial for constructing $\mathcal{K}_N$ than with a single $\underline{\mathbf{U}}^k \mathbf{f}$. Hence, we are able to state that in vector Prony analysis, $\mathbf{p}$ is chosen to minimize the $\mathbb{C}^{m \cdot N}$ -norm of $\mathbf{r}$, corresponding to choosing the projection operator $\mathbf{P}'_N$ so that $\mathbf{P}'_N \underline{\mathbf{U}}^N \mathbf{f}$ is the least-squares approximation to $\underline{\mathbf{U}}^N \mathbf{f}$ at $\mathbf{x}_0 \in \mathbb{X}$. In this way, $\mathbf{A}_x$ in (2.20) can be denoted as $\mathbf{K}'_N \in \mathbb{R}^{(m \cdot N) \times N}$:

$$\mathbf{b} = \mathbf{K}'_N \mathbf{p} + \mathbf{r}, \quad \mathbf{r} \perp \text{span}\{\mathbf{K}'_N\}.$$

In comparison with $\mathbf{K}_N \in \mathbb{R}^{m \times N}$, the rank-deficiency of $\mathbf{K}'_N$ when determining $\mathbf{p}$ can be avoided even in the case of data with low spatial dimension ($m \ll N$). Thus, a unique solution of $\mathbf{p}$, in other words, an unique decomposition of $\{\mathbf{y}_k\}_{k=0}^{2N-1}$ can be derived. The possibility of avoiding the rank-deficiency is a result of the use of time-shifting which enable us to utilize more data for computing $\mathbf{C}'_N$. The matrix $\mathbf{C}'_N$ is thought of as an approximation of the action of the Koopman operator on the finite-dimensional subspace $\mathcal{K}'_N$.

In summary, it can be stated that the difference between the Arnoldi-like algorithm and vector Prony analysis for KMD is the choice of finite-dimensional function spaces on which the original Koopman operator is projected. Both the derived matrices are thought of as the least-squares approximations of the corresponding action of the Koopman operator.

2.2.6 Extended DMD

The so-called extended DMD (EDMD) [18] was proposed to overcome a limitation of standard DMD, namely that DMD only relies on raw “linear” measurements $\mathbf{y}_k$, which might not be enough to fully characterize the dynamics of the underlying nonlinear system. Following [18], input data consisting of pairs of snapshots $(\mathbf{x}_k, \mathbf{y}_k)$ are introduced, which are stacked into two matrices similar to what we had previously in (2.11),

$$\mathbf{X}' := [\mathbf{x}_0, \dots, \mathbf{x}_{N-1}], \quad \mathbf{Y}' := [\mathbf{y}_0, \dots, \mathbf{y}_{N-1}]. \quad (2.23)$$

The key idea of EDMD is to introduce a new vector-valued observable function:

$$\mathbf{\Psi}(\mathbf{z}) = [\psi_1(\mathbf{z}), \dots, \psi_K(\mathbf{z})], \quad (2.24)$$

where the functions $\psi_i(\cdot)$ are called dictionary elements, and represent some *appropriately* chosen functions of the measurements. Then, one can consider the action of the Koopman operator in this richer subspace constructed with $\mathbf{\Psi}(\mathbf{z})$. The idea is that if the functions $\psi_i(\cdot)$ are appropriately chosen, then the evolution in this augmented subspace becomes linear or close to linear, and thus becomes described well by a finite dimensional approximation of the Koopman operator. Obviously, if the measurements $\mathbf{y}_k$ already are “rich” enough to capture the dynamics, there is no need to introduce (2.24), and it can be shown that EDMD then reduces to standard DMD [18]. Suggested choices for the dictionary elements ψ_i are monomials, polynomials, radial basis functions, or spectral elements [18]. Now, similar to (2.14) and (2.10) in the conventional DMD-algorithms, we construct an approximation of the Koopman operator $\mathbf{K}$ by:

$$\mathbf{K} := \mathbf{G}^\dagger \mathbf{Q}, \quad (2.25)$$

where the matrices $\mathbf{G} \in \mathbb{C}^{K \times K}$ and $\mathbf{Q} \in \mathbb{C}^{K \times K}$ are constructed as follows by utilizing (2.24):

$$\begin{aligned} \mathbf{G} &= \frac{1}{N} \sum_{i=0}^{N-1} \mathbf{\Psi}(\mathbf{x}_i)^* \mathbf{\Psi}(\mathbf{x}_i), \\ \mathbf{Q} &= \frac{1}{N} \sum_{i=0}^{N-1} \mathbf{\Psi}(\mathbf{x}_i)^* \mathbf{\Psi}(\mathbf{y}_i). \end{aligned} \quad (2.26)$$

The eigenvalues and eigenvectors of $\mathbf{K} \in \mathbb{C}^{K \times K}$ are denoted by μ'_i and $\boldsymbol{\xi}_i$, respectively, where we as usual have $\mathbf{K}\boldsymbol{\xi}_i = \mu'_i \boldsymbol{\xi}_i$. Two matrices of eigenvectors of $\mathbf{K}$ are defined by $\Xi = [\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_K]$, and $\mathbf{W}^* = \Xi^{-1} = [\mathbf{w}_1, \dots, \mathbf{w}_K]^*$, where $\mathbf{w}_i$ are the left eigenvectors of $\mathbf{K}$, associated with the eigenvalue μ'_i . The eigenfunctions $\varphi'_i(\mathbf{x})$ of $\mathbf{K}$ are approximated as $\varphi'_i(\mathbf{x}) = \mathbf{\Psi}(\mathbf{x})\boldsymbol{\xi}_i$.

Now, we explicitly express our observables as follows: $\mathbf{x} = \mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), \dots, f_m(\mathbf{x})]^\top$, where $f_i : \mathbb{X} \rightarrow \mathbb{R}$ are scalar-valued functions. Then, if all f_i lies within the span of our constructed observable functions (dictionary elements) ψ_i , we can express our scalar

observables as linear combinations of ψ_i :

$$f_i(\mathbf{x}) = \sum_{l=1}^K \psi_l(\mathbf{x}) b_{i,l} = \boldsymbol{\psi}(\mathbf{x}) \mathbf{b}_i, \quad (2.27)$$

where $\mathbf{b}_i \in \mathbb{R}^K$ is a vector of coefficients. Then, a matrix of coefficients $\mathbf{B} := [\mathbf{b}_1, \dots, \mathbf{b}_K] \in \mathbb{C}^{K \times K}$ can be constructed which relates the snapshots $\mathbf{x}$ to (2.24) as follows:

$$\mathbf{x} = (\boldsymbol{\Psi}(\mathbf{x}) \mathbf{B})^\top. \quad (2.28)$$

By defining, $\mathbf{X}_\Psi := [\boldsymbol{\Psi}(\mathbf{x}_0)^\top, \dots, \boldsymbol{\Psi}(\mathbf{x}_{N-1})^\top]^\top$, one can approximate $\mathbf{B}$ as follows:

$$\mathbf{B} = \mathbf{X}_\Psi^\dagger \mathbf{X}^\top. \quad (2.29)$$

Then, following [18], we can define approximations of the KMs $\mathbf{v}_j$ as the columns of the following matrix:

$$\mathbf{V} = [\mathbf{v}'_1, \dots, \mathbf{v}'_K] = (\mathbf{W}^* \mathbf{B})^\top. \quad (2.30)$$

Now, similar to the previously discussed KMD/DMD algorithms, the spectral decomposition of the original data is expressed in terms of μ'_i , φ'_j , and $\mathbf{v}'_j$:

$$\mathbf{x}_k = \sum_{j=1}^K (\mu'_j)^k \mathbf{v}'_j \varphi'_j(\mathbf{x}). \quad (2.31)$$

This method can be computationally demanding for large amounts of data. However, a faster version of this method, utilizing the so-called kernel trick, is described in [81].

2.3 The Perron-Frobenius (PF) operator

This section introduces the PF operator following [11, 75]. As previously mentioned, the Koopman operator $\mathbf{U}$ is the left adjoint of the PF operator $\mathbf{P}$. To define the action of $\mathbf{P}$, we need to introduce the so-called measure space given by the tuple $(\mathbb{X}, \mathcal{A}, \mu)$, associated with the nonlinear system (2.1), where $\mathcal{A}$ is a so-called Borel σ -algebra of subsets of $\mathbb{X}$, and μ a measure which assigns some value to $\mathcal{A}$. Further, we introduce f and g as functions defined on $\mathbb{X}$. The relation between $\mathbf{P}$ and $\mathbf{U}$ is by definition [11] given by:

$$\langle \mathbf{P}f, g \rangle = \langle f, \mathbf{U}g \rangle, \quad (2.32)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. Numerical methods connecting these operators have not extensively been discussed or demonstrated until recently [75]. By their definition and relation, it should not matter which one of them that is used to study dynamical systems [13, 75], and thus one can expect that the numerical methods used to approximate either one of them are viable.

To characterize the distribution, or density of states in $\mathbb{X}$ iterated by $\mathbf{x}_{k+1} = \mathbf{T}(\mathbf{x}_k)$ ($\mathbf{x} \in \mathbb{X}$) from (2.1), we introduce two functions (densities of states) as $f_0, f_1 \in L^1(\mathbb{X}, \mathcal{A}, \mu)^1$, or more simply written as $f_0, f_1 \in L^1(\mathbb{X})$. If $\mathbf{T}$ is nonsingular with respect to μ , then we can write the evolution from an initial density f_0 to f_1 by action of the map $\mathbf{T}$ as follows:

$$\int_{\mathcal{B}} f_1 d\mu = \int_{\mathbf{T}^{-1}(\mathcal{B})} f_0 d\mu, \quad (2.33)$$

where $\mathcal{B} \in \mathcal{A}$, e.g. an interval or “box” of state space $\mathbb{X}$. In particular, the so-called PF operator $\mathbf{P} : L^1(\mathbb{X}) \rightarrow L^1(\mathbb{X})$, is the name of the operator governing the evolution by application to f_0 as follows $f_1 = \mathbf{P}f_0$. With this definition, we can rewrite (2.33) as follows:

$$\int_{\mathcal{B}} \mathbf{P}f_0 d\mu = \int_{\mathbf{T}^{-1}(\mathcal{B})} f_0 d\mu, \quad \mathcal{B} \in \mathcal{A}, \quad (2.34)$$

which is the standard definition of the PF operator [11]. In the next subsection, a simple method is outlined which has been used to approximate $\mathbf{P}$ as a finite-dimensional matrix.

2.3.1 Perron-Frobenius via Ulam’s method and EDMD

Ulam’s method [74, 75] is based on the idea of partitioning the state space $\mathbb{X}$ into a finite number of measurable boxes $\mathbb{B}_i \subset \mathcal{A}$. First, the indicator function is introduced as follows

$$\mathbb{1}_{\mathbb{B}_i}(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \mathbb{B}_i, \\ 0, & \text{otherwise.} \end{cases} \quad (2.35)$$

Then, it can be shown that we can approximate the entries $p_{ij} = [\mathbf{P}]_{ij}$ of a matrix $\mathbf{P}$ which approximates the action of the PF operator, as follows:

$$p_{ij} = \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{\mathbb{B}_j} \left(\mathbf{T} \left(x_i^{(k)} \right) \right). \quad (2.36)$$

$\mathbf{P}$ describes the probability of the mapping from a box $\mathbb{B}_i$ to another one $\mathbb{B}_j$ by $\mathbf{T}$. The calculation (2.36) is very simple: take $k = 1, \dots, n$ points in each box $\mathbb{B}_i$, then count how many times $\mathbf{T}(x_i^{(k)})$ ends up in $\mathbb{B}_j$. Another method of approximating $\mathbf{P}$ was noted in [75] by means of EDMD. It can be shown that eigenfunctions of the PF operator can be estimated by computing the left eigenvectors of the following matrix:

$$\mathbf{P} = \mathbf{Q}^* \mathbf{G}^\dagger, \quad (2.37)$$

where $\mathbf{Q}$ and $\mathbf{G}$ are given by (2.25). The connection between the matrices $\mathbf{K}$ and $\mathbf{P}$ and the Koopman and PF operators can easily be seen via simple matrix manipulation by

¹For $f : \mathbb{X} \rightarrow \mathbb{R}$, the so-called L^p -space is formally defined as the space where $\int |f(x)|^p \mu(dx) < \infty$ [11].

plugging them into $\langle \mathbf{P}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{K}\mathbf{y} \rangle$ according to (2.32) as follows:

$$\begin{aligned}\langle (\mathbf{Q}^* \mathbf{G}^\dagger) \mathbf{x}, \mathbf{y} \rangle &= \langle \mathbf{x}, (\mathbf{G}^\dagger \mathbf{Q}) \mathbf{y} \rangle, \\ (\mathbf{Q}^* \mathbf{G}^\dagger \mathbf{x})^\top \mathbf{y} &= \mathbf{x}^\top (\mathbf{G}^\dagger \mathbf{Q}) \mathbf{y}, \\ \mathbf{x}^\top ((\mathbf{G}^\dagger)^\top \mathbf{Q}) \mathbf{y} &= \mathbf{x}^\top (\mathbf{G}^\dagger \mathbf{Q}) \mathbf{y},\end{aligned}$$

where $(\mathbf{G}^\dagger)^\top = \mathbf{G}^\dagger$ by definition since $\mathbf{G}$ is a symmetric matrix (see (2.25)), and $\mathbf{x}, \mathbf{y}$ are arbitrary vectors of appropriate lengths. Finally, the transpose of the approximation of the PF operator via Ulam's method: $\mathbf{P}^\top$ can also be shown to provide an approximation of $\mathbf{K}$ [75]. This will be experimentally verified in the next section, and numerically establishes the connection between Koopman and PF operators.

2.4 Two illustrative examples

2.4.1 A linear system

Here, we will consider the discrete-time linear dynamical system given by

$$\mathbf{x}_{k+1} = \mathbf{D}\mathbf{x}_k, \quad (2.38)$$

where $\mathbf{x} = [x_1, x_2]^\top$, and the matrix $\mathbf{D}$ is defined as follows:

$$\mathbf{D} = \begin{bmatrix} 0.9 & -0.35 \\ 0.35 & 0.97 \end{bmatrix}. \quad (2.39)$$

The two eigenvalues of $\mathbf{D}$ are $\lambda_1 \approx 0.935 + i0.348$ and $\lambda_2 \approx 0.935 - i0.348$, and $|\lambda_1| \approx 0.998$. Thus, we can understand that the dynamics are oscillatory and lightly damped. Now, we first define the initial condition as $\mathbf{x}_0 = [2, -2]^\top$. Then (2.38) is iterated, giving the following input data (showed rounded off to two digits following decimal point):

$$\mathbf{X} = \begin{bmatrix} 2 & 2.5 & 2.68 & 2.53 & 2.06 \\ -2 & 1.24 & -0.33 & 0.62 & 1.49 \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} 2.5 & 2.68 & 2.53 & 2.06 & 1.33 \\ 1.24 & -0.33 & 0.62 & 1.49 & 2.16 \end{bmatrix}. \quad (2.40)$$

The dynamics are shown in Fig. 2.1 for the first five iterations. DMD, Arnoldi-KMD, and vector Prony KMD are applied to the data (2.40). The real and imaginary parts of the identified eigenvalues are shown in Fig. 2.2 and plotted against the norm of their modes $\|\mathbf{v}\|$, with $\mathbf{v}$ represented by $\tilde{\mathbf{V}}$, $\hat{\mathbf{V}}$, and $\overline{\mathbf{V}} := (\hat{\phi}\mathbf{y}_0)\phi$ from Arnoldi KMD, Prony KMD, and DMD, respectively. Despite the fact that Arnoldi-KMD and Prony identify a larger number of eigenvalues than two: in this case $N - 1 = 5$ for Arnoldi and $N/2 = 3$ for Prony, the norm of the modes not associated with λ_1 or λ_2 are zero, thus not contributing to the decompositions. From Fig. 2.2 we see that all algorithms produce the same spectral decomposition. Via DMD, we can calculate the eigenfunction corresponding to λ_1 , given by $\varphi_1(\mathbf{y}) = \hat{\phi}_1 \mathbf{y}$, which is equal to the eigenfunction of (2.38) associated with λ_1 (defined

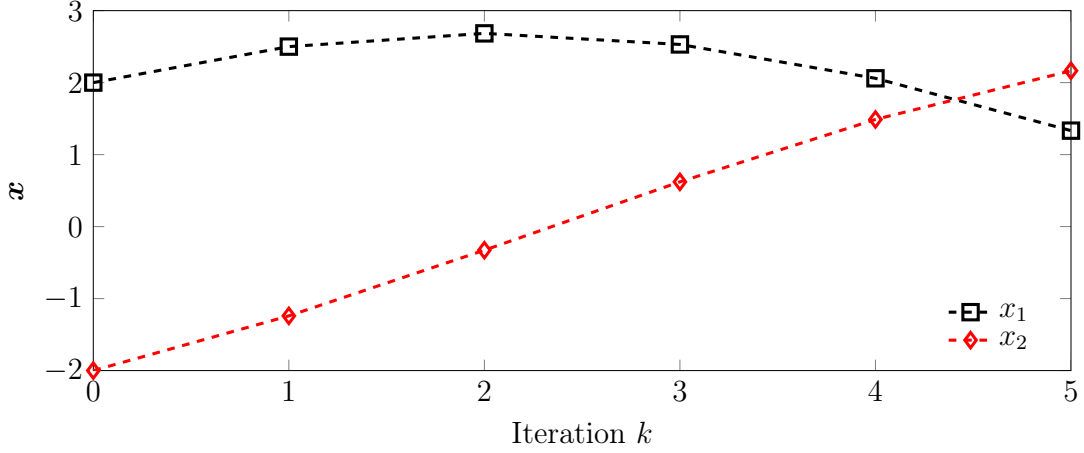


Figure 2.1: Iterations of the linear dynamical system given by (2.38) and (2.39).

as $\varphi_1(\mathbf{y}) = \mathbf{v}_{1,1}\mathbf{y}$, where $\mathbf{v}_{1,1}$ is the *left* eigenvector of $\mathbf{D}$ associated with λ_1). This is because $\bar{\mathbf{A}} = \mathbf{Y}\mathbf{X}^\dagger = \mathbf{D}$ from (2.14). Now, Ulam's method is also applied to the dynamics to approximate $\mathbf{P}^\top$ (an approximation of the action of the Koopman operator), by dividing the 2D state space into 100 equally sized squares of unit area with exactly n number of initial conditions in each square: i.e. we consider a $(-5, 5) \times (-5, 5)$ domain of the state space. The squares are *not closed*, i.e. they do not contain the boundaries between them. However, in the calculations here, no point starts or ends up at any boundary. An eigendecomposition of $\mathbf{P}^\top$ provides an approximation of λ_1 as $\hat{\lambda}_1 = 0.928 + i0.342$, i.e. close to λ_1 . The absolute values of the eigenfunctions from EDMD, the linear system ((2.38) and (2.39)), and $\mathbf{P}^\top$, associated with λ_1 , are shown in Fig. 2.3, normalized by their largest values $\max_i |\varphi_1(\mathbf{x}_i)|$. From the figure we clearly see that EDMD well approximates the original eigenfunction, while Ulam's method only gives a rough estimate, despite utilizing $n = 50$ initial conditions in each one of the 100 squares. This means that 5000 snapshots (or initial conditions $\mathbf{x}_0$) are utilized to construct $\mathbf{P}$, compared to 6 for DMD and Arnoldi-KMD. A finer partition of the state space leads to $\mathbf{P}$ becoming singular, and thus we are only able to achieve a rough estimation of the eigenfunction. In this example, EDMD was computed with 100 randomly chosen initial conditions, and the dictionary elements ψ_i were chosen as Hermite polynomials of up to fourth order according to [18].

2.4.2 A nonlinear system

Let us here discuss a toy example of a nonlinear dynamical system similar to [17, 76]. The discrete-time nonlinear system is defined as follows:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{k+1} = \begin{bmatrix} a & 0 \\ a[x_1]_k & 2a^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_k, \quad (2.41)$$

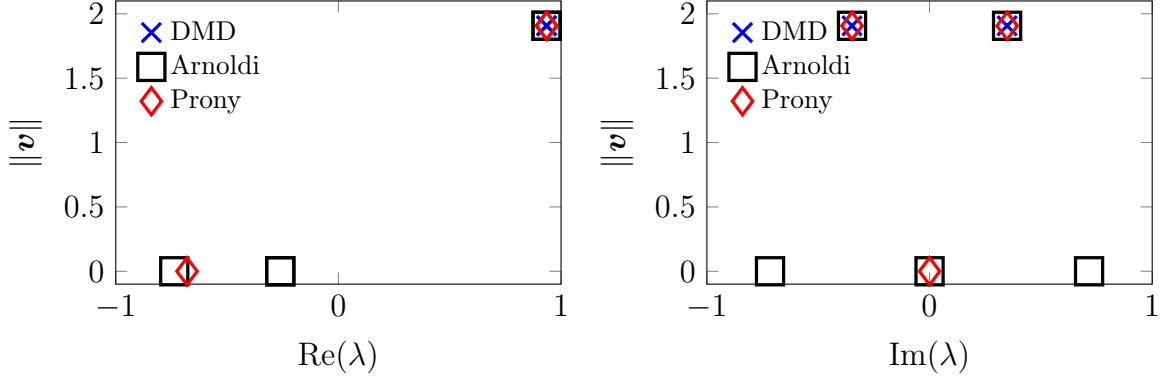


Figure 2.2: Norms of modes $\|\mathbf{v}\|$ plotted against real and imaginary parts of associated eigenvalues from DMD, Arnoldi-like KMD, and vector Prony KMD. Note that the two eigenvalues of $\mathbf{D}$ in (2.39) are $\lambda_1 = 0.935 + i0.348$ and $\lambda_2 = 0.935 - i0.348$.

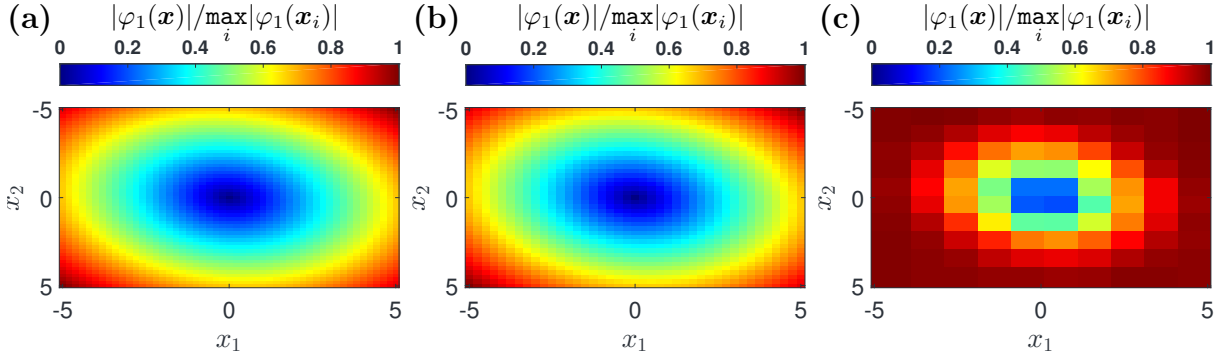


Figure 2.3: Absolute values of approximations of the eigenfunction $|\varphi_1(\mathbf{x})|$ associated with λ_1 calculated from: (a) the linear dynamical system (2.39); (b) EDMD; (c) Ulam's method. The plots show $|\varphi_1(\mathbf{x})|$ normalized by $\max_i |\varphi_1(\mathbf{x}_i)|$.

where a is a scalar parameter. Following [17, 76], with the change of variables: $\mathbf{y} = [y_1, y_2, y_3]^\top := [x_1, x_2, x_1^2]^\top$ we get the following *linear* dynamical system:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_{k+1} = \begin{bmatrix} a & 0 & 0 \\ 0 & 2a^2 & a \\ 0 & 0 & a^2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_k, \quad (2.42)$$

in terms of our new vector valued observable $\mathbf{y}$, since we have that

$$\begin{aligned} [y_1]_{k+1} &= a[x_1]_k = a[y_1]_k, \\ [y_2]_{k+1} &= a[x_1]_k^2 + 2a^2[x_2]_k = a[y_3]_k + 2a^2[y_2]_k, \\ [y_3]_{k+1} &= ([x_1]_{k+1})^2 = a^2([x_1]_k)^2 = a^2[y_3]_k. \end{aligned} \quad (2.43)$$

This is a simple example of how a nonlinear system can be transformed into a linear system by an appropriate choice of vector valued observable. The eigenvalues of the

system matrix in (2.42) are $\lambda_1 = a$, $\lambda_2 = 2a^2$, and $\lambda_3 = a^2$. By applying DMD/KMD to data using the observable $\mathbf{y} = [x_1, x_2, x_1^2]^\top$, the linear dynamical system (2.42) is hence captured and results as accurate as in the previous section are obtained. Further, the right and left eigenvectors of system matrix in (2.42) are contained in the following two matrices:

$$\mathbf{V}_{\text{right}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1/a \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{V}_{\text{left}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1/a \\ 0 & 0 & 1 \end{bmatrix}$$

where $\mathbf{V}_{\text{right}}\mathbf{V}_{\text{left}} = \mathbf{I}_3$, and $\mathbf{V}_{\text{right}} := [\mathbf{v}_{r,1}, \mathbf{v}_{r,2}, \mathbf{v}_{r,3}]$ and $\mathbf{V}_{\text{left}} := [\mathbf{v}_{l,1}^\top, \mathbf{v}_{l,2}^\top, \mathbf{v}_{l,3}^\top]^\top$, where $\mathbf{v}_{r,i}$ and $\mathbf{v}_{l,i}$ denote the i -th right and left eigenvectors, respectively. Defining $\varphi_i(\mathbf{x}) = \mathbf{v}_{l,i}\mathbf{y}$ as the Koopman eigenfunctions, we have

$$\begin{aligned} \varphi_1(\mathbf{x}) &:= \mathbf{v}_{l,1}\mathbf{y} = x_1, \\ \varphi_2(\mathbf{x}) &:= \mathbf{v}_{l,2}\mathbf{y} = x_2 + x_1^2/a, \\ \varphi_3(\mathbf{x}) &:= \mathbf{v}_{l,3}\mathbf{y} = x_1^2. \end{aligned}$$

Then, we can explicitly express the evolution of the nonlinear system (2.41) in terms of our observable $\mathbf{y}$ by KMD as follows:

$$\mathbf{y}_k = \mathbf{g}(\mathbf{x}_k) = (\mathcal{U}^k \mathbf{g})(\mathbf{x}_0) = \sum_{j=1}^3 \lambda_j^k \varphi_j(\mathbf{x}_0) \mathbf{v}_{r,j} = \begin{bmatrix} \lambda_1^k \cdot x_{1,0} \\ \lambda_2^k \cdot (x_{2,0} + \frac{x_{1,0}^2}{a}) - \lambda_3^k \cdot \frac{x_{1,0}^2}{a} \\ \lambda_3^k \cdot x_{1,0}^2 \end{bmatrix}, \quad (2.44)$$

where the KMs are *linear*, the eigenfunctions $\varphi_i(\mathbf{x})$ are *nonlinear* functions of $\mathbf{x}$, and ‘ $\cdot$ ’ denotes *scalar* multiplication. In this simple example, the evolution (2.44) depends only on the initial state $\mathbf{x}_0 := [x_{1,0}, x_{2,0}]^\top$ and the parameter a (the eigenvalues λ_i are in this example also functions of a).

In the following, we look at a numerical example which demonstrates the potential benefits of using a nonlinear observable and time-shifted measurements. We define the *scalar* observable $f(\mathbf{x})$ as follows:

$$f(\mathbf{x}) = x_1 + x_2 + x_1^2, \quad (2.45)$$

and define the snapshots $\mathbf{z}_k$ containing three time-shifted measurements as:

$$\mathbf{z}_k := [f(\mathbf{x}_k), f(\mathbf{x}_{k+1}), f(\mathbf{x}_{k+2})]^\top = [z_{1,k}, z_{2,k}, z_{3,k}]^\top. \quad (2.46)$$

With $a = -0.7$ and $\mathbf{x}_0 := [0.5, 0.5]^\top$, the dynamics of (2.41) and the time-series of the scalar observable $f(\mathbf{x})$ are shown in Figs. 2.4 (a) and (b), respectively. Figure 2.4 (c) shows the dynamics of the time-shifted measurements z_1 and z_2 of the vector-valued observable (2.46). Clearly, the dynamics in Fig. 2.4 (c) resembles Fig. 2.4 (a). By application of any of the DMD/KMD methods to the data $\{\mathbf{z}_k\}_{k=0}^N$ we find the eigenvalues $\lambda_{1,2,3}$ of the original nonlinear system. Furthermore, we also realize that there exists a map $\mathbf{M}_{zy}$ such that

$$\mathbf{y}_k = \mathbf{M}_{zy} \mathbf{z}_k,$$

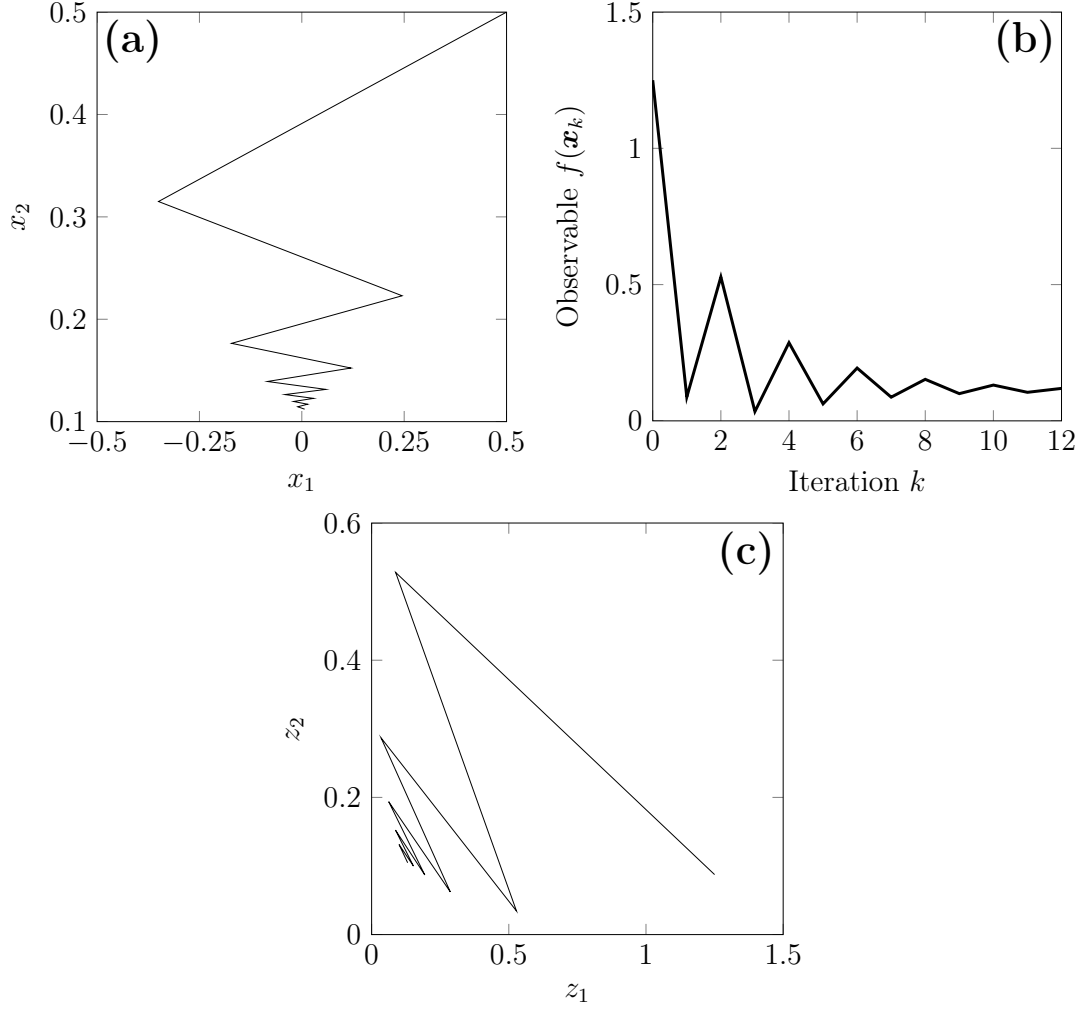


Figure 2.4: (a) Dynamics of the nonlinear system (2.41). (b) Scalar observable function (2.45). (c) Time-shifted measurements z_1 and z_2 of the vector-valued observable (2.46).

holds exactly, which takes us from the space of time-delayed measurements to $\mathbf{y}$ (and therefore also to $\mathbf{x}$), because all components of $\mathbf{y}$ are included in our scalar observable $\mathbf{z}$.

This was as an illustrative example of how the evolution of a simple nonlinear dynamical system can be described completely by a finite dimensional spectral decomposition (KMD) by an appropriate choice of observable functions.

2.5 Conclusions

This chapter introduced the Koopman operator, the Koopman Mode Decomposition (KMD), and discussed several numerical algorithms for the approximation of KMD, commonly called Dynamic Mode Decomposition (DMD). The effectiveness of several methods

was demonstrated by two simple examples. Two proofs were also provided which unified two commonly utilized KMD-algorithms under certain assumptions. Utilizing time-shifted measurements, an inherent feature of the so-called vector Prony decomposition, has the potential to substantially enlarge the subspace spanned by sampled data, on which the decomposition is defined, consequently enabling a better posed least squares problem. Further, connections between the Koopman and Frobenius operators were both discussed and numerically demonstrated. From this chapter, we understand that an appropriate choice of KMD-algorithm, and appropriately chosen observables are crucial to the application of KMD to a certain problem or set of data. Related to this, in the next chapter, a numerical evaluation of three similar KMD-algorithms is conducted.

Chapter 3

Comparison and numerical evaluation of KMD-techniques

3.1 Introduction

Chapter 2 introduced Koopman operator theory and explained the theoretical ideas underlying KMD algorithms. In addition, several algorithms were described, often called DMD, which can be viewed as numerical approximations of the theoretical KMD. Since the so-called standard DMD was proposed and demonstrated in [16], several new DMD-algorithms and extensions have been proposed: see e.g. [17, 18, 62, 77, 78, 82]. All these algorithms essentially have the same purpose: to extract the dominant features in the observed dynamics in terms of a spectral decomposition of the Koopman operator, which can be viewed as a type of reduced-order model. They indeed produce the exact same decomposition under certain conditions, which was theoretically and numerically demonstrated in Chapter 2, and will also be demonstrated in this chapter by numerical experiments. However, applying these methods directly to data from a nonlinear system, for a case where the number of measurement locations, called the *spatial dimension* here, is smaller than the number of temporal snapshots, the algorithms differ by definition since the number of modes is not equal. One could, however, imagine a case where some modes are zero, and the rest of the modes are equal, thus resulting in the same decomposition, despite the total number of modes being different. This is exactly the case when companion-matrix based KMD is applied to data from a linear system: see Section 2.4. In practice, noise is always present in some form, and the performance of the algorithms when applied to data contaminated by noise is crucial for practical applications to real measurements.

The purpose of this chapter is to study applications of KMD to data of low spatial dimension. In particular, the three closely related algorithms outlined in Chapter 2: Arnoldi-like KMD (called Arnoldi KMD from now), standard DMD, and Prony KMD are evaluated here. These are all based on “raw” measurements without use of any additional observable function such as in EDMD. Note, however, that Prony KMD utilizes so-called time-shifting or time-delayed measurements, with the effect of expanding the

subspace on which the Koopman operator is defined. A numerical evaluation is conducted by application of these algorithms to three fundamentally different sets of data: data produced by a linear dynamical system, weakly nonlinear power system oscillations, and lastly highly nonlinear wind speed data. Depending on the data, we introduce noise; vary the ratio between the number of temporal snapshots and the spatial dimension by adjusting the temporal length, sampling frequency, and by constructing snapshots containing time-shifted measurements. Furthermore, their performance is evaluated by their ability to predict future values. The numerical investigation conducted here sheds some light on the applicability of these techniques to various types of data with low spatial dimension—useful for potential practitioners of them.

The remainder of the chapter is organized as follows. Section 3.2 explains and outlines the methodology adopted here to perform the numerical evaluation. Section 3.3 describes the considered data and test systems. Section 3.4 presents and discusses the experimental results. Finally, conclusions are given in Section 3.5.

3.2 Methodology

This section explains the methodology adopted in the comparison of the three considered KMD-algorithms. For example, how eigenvalues, modes and modal frequencies (given by the eigenvalues) are compared.

3.2.1 Sorting and selecting modes

The input data consist of a set of snapshots $\mathbf{y}_k \in \mathbb{R}^m$ as previously:

$$\{\mathbf{y}_0, \dots, \mathbf{y}_{2N-1}\}. \quad (3.1)$$

By application of the DMD/KMD algorithms to (3.1), Arnoldi KMD outputs $2N - 1$ modes and vector Prony N modes, while DMD outputs m modes. It is possible that only a few, dominant modes are enough to describe the dynamics well (depending on the data and the “success of the decomposition”), while the other ones are redundant in our analysis. For consistency in the comparison of modes and modal frequencies generated by the algorithms, we follow two rules:

- (i) Arnoldi KMD modes and Prony modes are ranked based on $|\lambda_i|^{N_{\max}} \|\mathbf{v}_i\|$, where $N_{\max}$ is $2N - 1$ for Arnoldi and $N - 1$ for Prony, and λ_i and $\mathbf{v}_i$ represent the eigenvalues and modes generated by the respective algorithm.
- (ii) The m or N highest ranked modes are chosen, whichever is lower, and are called *dominant modes*. Most often we have $m < N$ or $m \ll N$. When $N < m$, only the N highest DMD modes ranked by $|\lambda'_i|$ are considered.

3.2.2 Comparing eigenvalues

Linear dynamics measured in noise will be considered. The eigenvalues of the matrix governing the dynamics are thus exactly known. The performance of the algorithms can then be evaluated by measuring the distance d_λ between a known eigenvalue λ_1 and an estimated eigenvalue $\tilde{\lambda}_1$ from the algorithms simply by computing the euclidean distance as follows:

$$d_\lambda := \sqrt{\operatorname{Re}(\lambda_1 - \tilde{\lambda}_1)^2 + \operatorname{Im}(\lambda_1 - \tilde{\lambda}_1)^2}, \quad (3.2)$$

where $\operatorname{Re}(z)$ takes the real part and $\operatorname{Im}(z)$ the imaginary part of a complex value z .

3.2.3 Comparison of modal frequency

The ability of the algorithms to capture the modal frequencies (assuming that the target dynamics are linear or exhibit a fixed frequency f_i) is quantified by:

$$\Delta f_i = (f_i - f_z)/f_i, \quad (3.3)$$

where f_i is a frequency identified from eigenanalysis of the linearized system equations in the case of power system data, and f_z the frequency closest to f_i among the dominant modes for each KMD algorithm. For example, for a mode $\mathbf{v}_z$ associated with the eigenvalue λ_z the modal frequency is calculated as follows: $f_z = \operatorname{Im}(\ln(\lambda_z))/(2\pi T_s)$, where T_s is the sampling period.

3.2.4 Comparison via time-shifted snapshots

As previously, we construct a Hankel-like matrix from a scalar time-series y_i ($i = 1, \dots, N$):

$$\mathbf{H}(y_i) = \begin{bmatrix} y_1 & y_2 & \cdots & y_{N-m+1} \\ y_2 & y_3 & \cdots & y_{N-m+2} \\ \vdots & \vdots & \ddots & \vdots \\ y_m & y_{m+1} & \cdots & y_N \end{bmatrix} := [\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2, \dots, \tilde{\mathbf{y}}_{N-m+1}]. \quad (3.4)$$

We can view $\tilde{\mathbf{y}}_i$ as “fictive” snapshots, generally defined by

$$\tilde{\mathbf{y}}_i := [y_i, y_{i+d}, \dots, y_{i+(m-1)d}]^\top, \quad (3.5)$$

where d is a time-shift ($d = 1$ in (3.4)), m the desired spatial dimension, and $N_s = N - m + 1$ denote the number of snapshots. The parameters of (3.5) are adjusted in this evaluation to achieve different values of m and N_s . By applying the DMD algorithms to (3.4), each component of the modes $\mathbf{v} \in \mathbb{C}^m$ ($m > 1$) describes how an oscillation varies in magnitude and phase with time. The time-shifted snapshots provide a way of expanding the subspace K on which the Koopman operator is projected (see (2.4) and (2.5)), if the snapshots $\tilde{\mathbf{y}}_i$ are linearly independent.

3.2.5 Comparison via future prediction

As a last comparison, we will look at the accuracy of future predictions a few samples ahead. This serves as an indication of the algorithm's ability to capture the underlying dynamics. For a scalar time-series y_i ($i = 1, \dots, 2N-1$), the considered algorithms iterate the data n steps forward in time as follows:

$$\begin{aligned}\hat{\mathbf{y}}_{i+n} &:= \bar{\mathbf{A}}^n \mathbf{y}_i, \\ \hat{\mathbf{Y}}_{i+n} &:= \mathbf{Y} \mathbf{C}_{2N-1}^n, \\ \hat{\mathbf{A}}_{y+n} &:= \mathbf{A}_y \mathbf{C}_N^n.\end{aligned}\tag{3.6}$$

Refer to (2.10), (2.14), and (2.21) for their definitions. The predicted values are to be found in the vector or matrices on the left hand side in (3.6).

3.3 Description of test systems and data

This section describes the test systems used to generate data for the evaluation. Data from weather simulations of Japan are used as an additional comparison.

3.3.1 Linear system and noise

We consider a discrete-time, linear, dynamical system $\mathbf{x}_{k+1} = \mathbf{D}\mathbf{x}_k$, where $\mathbf{D}$ is same as in (2.39). Here, we consider noisy measurements where each snapshot $\mathbf{y}_k \in \mathbb{R}^2$ is the state vector $\mathbf{x}_k \in \mathbb{R}^2$ corrupted by noise as follows:

$$\mathbf{y}_k = \mathbf{x}_k + \boldsymbol{\gamma}_k,\tag{3.7}$$

where $\boldsymbol{\gamma}_k \in \mathbb{R}^2$ is a vector of independent and identically distributed (i.i.d.) Gaussian noise, added with a specified Signal to Noise Ratio (SNR).

3.3.2 Power system example

A slightly modified version of Kundur's two area test system [7], shown in Fig. 3.1 is used, which comprises seven buses, and three generators, each denoted by G_i . The nominal system and angular frequencies are $f_{\text{sys}} = 60$ Hz and $\omega_s = 2\pi f_{\text{sys}}$ rad/s. The system has one constant impedance load (see e.g. [7] for a definition of this). The dynamics are described by the differential algebraic equations (DAE) in (1.1). The state variables $\mathbf{x}$ are generator internal bus angles δ and deviations from nominal frequency ω_i . Initial steady state values for w_i are $\boldsymbol{\omega}_0 = [\omega_{1,0}, \omega_{2,0}, \omega_{3,0}]^\top = [0, 0, 0]^\top$. The dynamics of each generator i are given by the following differential equations:

$$\frac{d\delta_i}{dt} = \omega_i, \quad M_i \frac{d\omega_i}{dt} = P_{m,i} - P_{e,i} - D_i \omega_i,\tag{3.8}$$

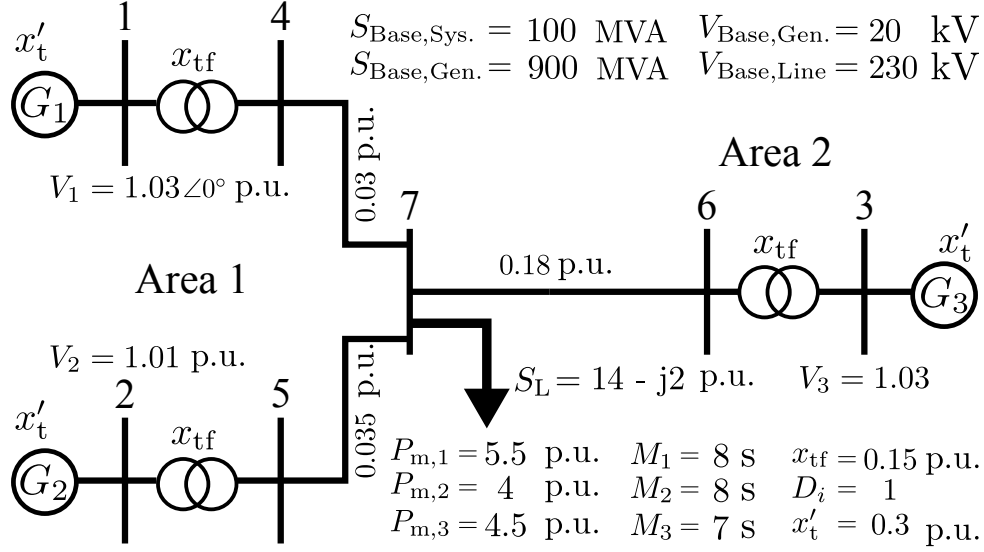


Figure 3.1: Test system with seven buses, three synchronous generators, and a load at bus 7.

where $D_i = D$ is damping, assumed same for all generators, P_m and P_e are the mechanical (assumed constant) and electrical power, and M_i an inertia constant. The electrical power produced by the i -th generator $P_{e,i}$, is given by

$$P_{e,i} = \frac{E_i V_i}{x'_{t,i}} \sin(\delta_i - \theta_i),$$

where $V_i \angle \theta_i$ is the bus voltage, and $E_i \angle \delta_i$ the generator's voltage connected to $V_i \angle \theta_i$ via a so-called transient reactance $x'_{t,i}$: see Fig. 3.1 for a schematic depiction of the test system.

3.3.3 Weather prediction data

In contrast to the power system, which generates dynamics well-described by so-called small-signal analysis (eigenanalysis of the linearized DAEs), we also use highly nonlinear wind speed predictions. The data are outputs from weather predictions of the Cloud Resolving Storm Simulator (CReSS) [63], discussed in detail in Chapter 6, solved for a domain encompassing the whole of Japan ($\mathcal{D}_1$ in Fig. 6.2, Chapter 6). The spatial resolution (distance between nearest grid points) is 2 km, and the time-resolution is 1 h.

3.4 Results

This section presents and discusses the results of applying the three considered algorithms to noise-corrupted measurements from a linear system, power system dynamics, and wind speed predictions of Japan.

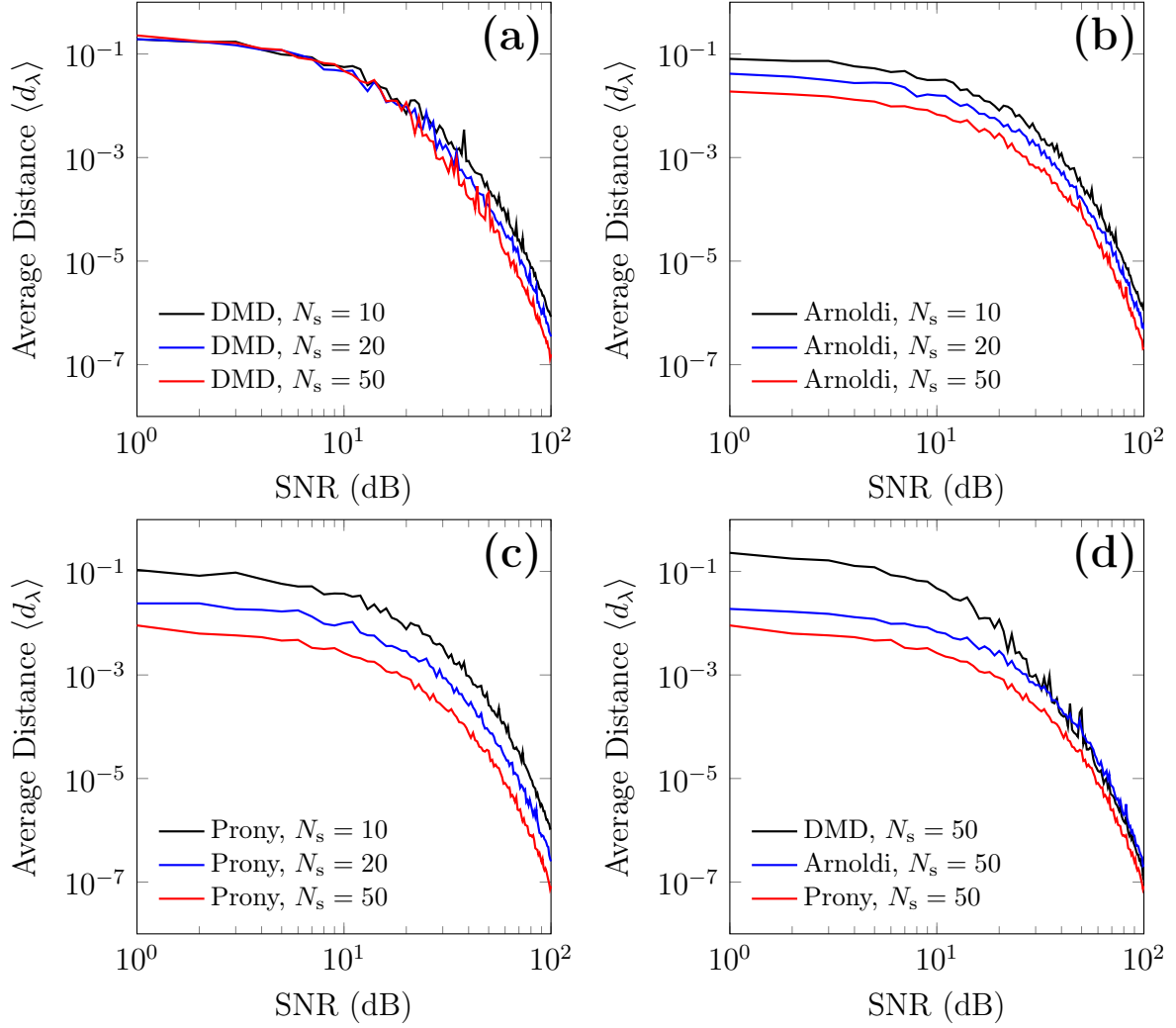


Figure 3.2: Average distance d_λ to the eigenvalue λ_1 based on 200 runs at every Signal to Noise Ratio (SNR) with noise applied to *all snapshots*, for (a) DMD; (b) Arnoldi KMD; (c) vector Prony KMD; (d) comparison between algorithms for number of snapshots $N_s = 50$.

3.4.1 Noise-corrupted data from a linear system

Here, the algorithms are applied to noise-corrupted measurement such as in (3.7). The effects of noise in DMD have been discussed in several papers, see e.g. [16, 17, 78, 83–86]. In [16, 78] it is stated that standard DMD is better suited for noisy data than Arnoldi KMD. This is mainly due to two reasons. First, Arnoldi KMD involves fitting constants such that the last snapshot, say, $\mathbf{y}_N \in \mathbb{R}^m$, can be expressed as linear combinations of all previous snapshots, and thus noise has a large impact on $\mathbf{y}_N$ [78]. Second, in DMD, often a reduced matrix $\tilde{\mathbf{A}}$ (see (2.16)) is utilized when processing large data sets, and is constructed from the r dominant singular values and vectors from the SVD of matrix of input data $\mathbf{X} = [\mathbf{y}_1, \dots, \mathbf{y}_{N-1}]$. This can be viewed as a type of filtering able to

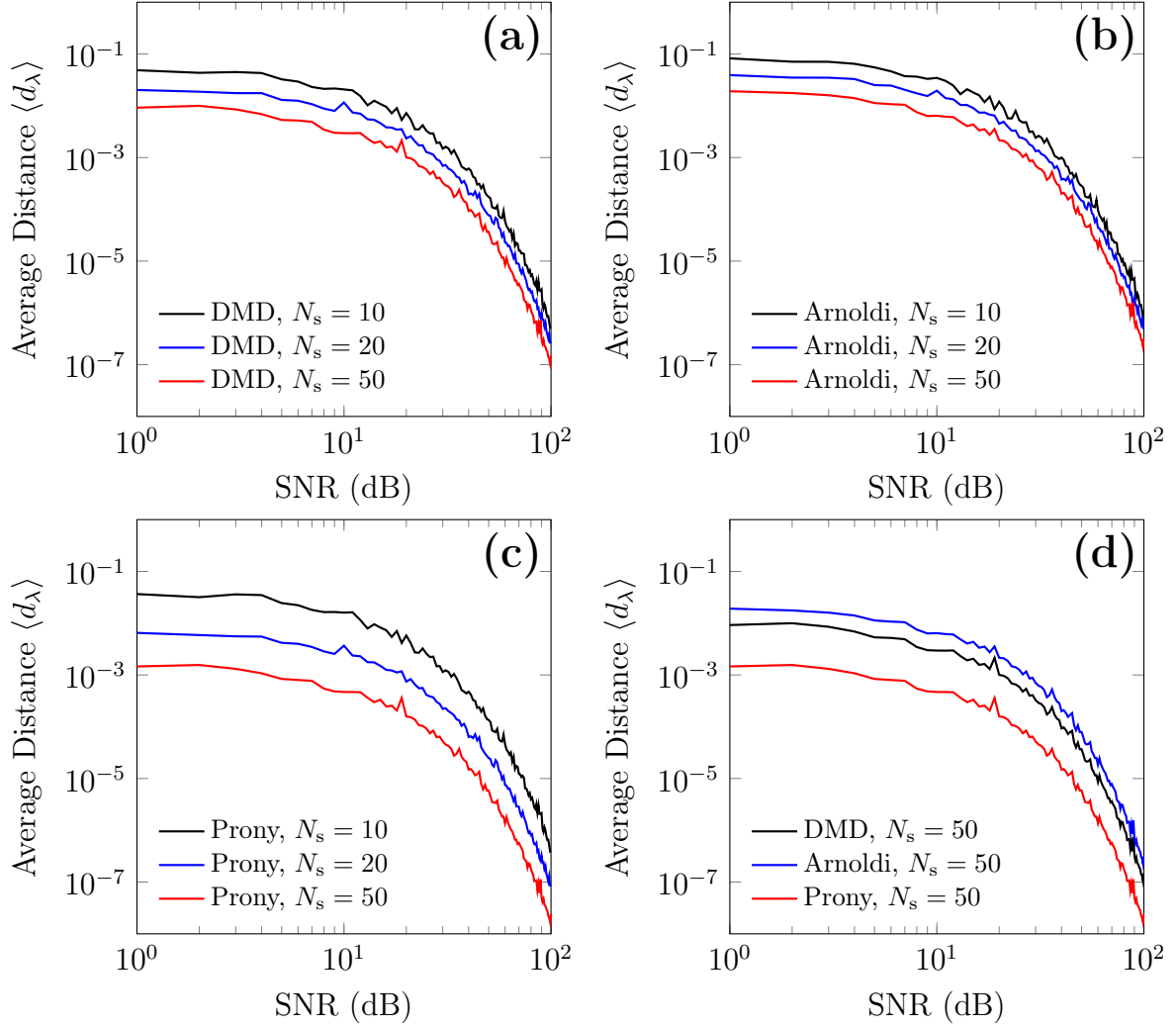


Figure 3.3: Average distance d_λ to the eigenvalue λ_1 based on 200 runs at every Signal to Noise Ratio (SNR) with noise applied to *only the last snapshot*, for (a) DMD; (b) Arnoldi KMD; (c) vector Prony KMD; (d) comparison between algorithms for $N_s = 50$.

remove noise [80]. However, in cases where the spatial dimension m is not large compared to the dimensionality of the dynamical system, it becomes obvious that the reduced matrix approach is not feasible without also possibly removing relevant dynamic features. Consequently, non-reduced, or exact DMD, has to be utilized when m is small.

In the following, the three considered algorithms are applied to noise-contaminated snapshots, and their performances are compared by calculating the distances between their closest eigenvalues and a known eigenvalue λ_1 , according to (3.2). We apply Gaussian noise with SNR varied from 0 to 100 dB. At each noise-level, the results of 200 applications are averaged to smooth the variations of individual results, each run with a randomly chosen initial condition $\mathbf{x}_0$. The noise is added to each $\mathbf{x}_i$ (see (3.7)) with the function

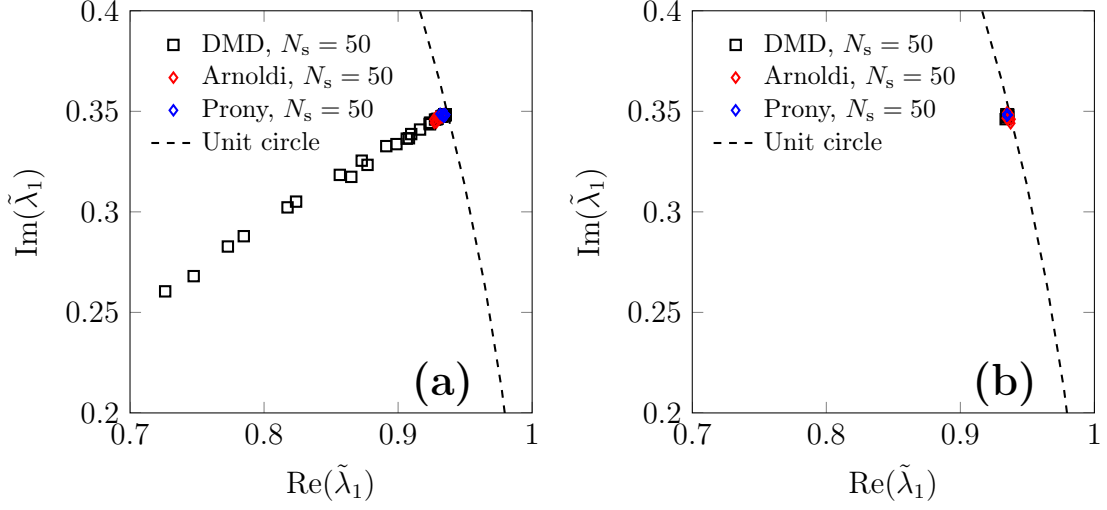


Figure 3.4: Identified eigenvalues for different SNR and with $N_s = 50$ snapshots in (a) for noise added to all snapshots; (b) noise added to only the last snapshot.

`awgn` in MATLAB. Three different cases of the number of snapshots N_s are considered: 10, 20, and 50, to investigate the potential influence on performance by increasing the amount of data. In particular, for the case that $m > N_s$, which is often the case in fluid experiments, [85] derived an analytical expression of the noise-induced bias to the DMD matrix $\tilde{\mathbf{A}}$, as a function of N_s and the noise variance σ^2 . Right now, we always have that $m \ll N_s$ and $m = 2$. The average distances d_λ to the eigenvalue λ_1 based on 200 runs, denoted $\langle d_\lambda \rangle$, for all algorithms are shown in Fig. 3.2, and plotted vs. the SNR. All algorithms show improvements for both a greater number of snapshots, and with increased SNR. The best results, and also the largest improvements with more snapshots, are seen for vector Prony KMD, and the reason is discussed below.

To further investigate the influence of noise, another test case is considered. Now, noise is applied to *only* the last snapshot $\mathbf{y}_N$. The results are shown in Fig. 3.3. Interestingly, the performance of Arnoldi KMD remains almost identical, indicating a strong dependence only on the last snapshot, which was discussed earlier. On the other hand, DMD and vector Prony KMD experience substantial improvements in performance. In Fig. 3.3 (d) we see that DMD is ranked second best after Prony. Prony KMD utilizes the so-called vector Hankel matrix, and even though noise is partly introduced in the regression problem, its effects have been mitigated since the entire last column is not affected: see (2.19). This construction did not only work better with partially introduced noise as in this example, but also in the previous example. In other words, noise reduction effects are obtained by restructuring the problem with the Hankel matrix. It was already proven that the eigenvalues of Arnoldi KMD and DMD are exactly the same if $m \geq N_s - 1$, provided certain conditions hold. In fact, by utilizing the vector Hankel matrices (2.19) as input to Arnoldi KMD and DMD too, the exact same results as for Prony KMD are achieved. In this case, the number of rows $\#rows$ and columns $\#col$ of the matrices

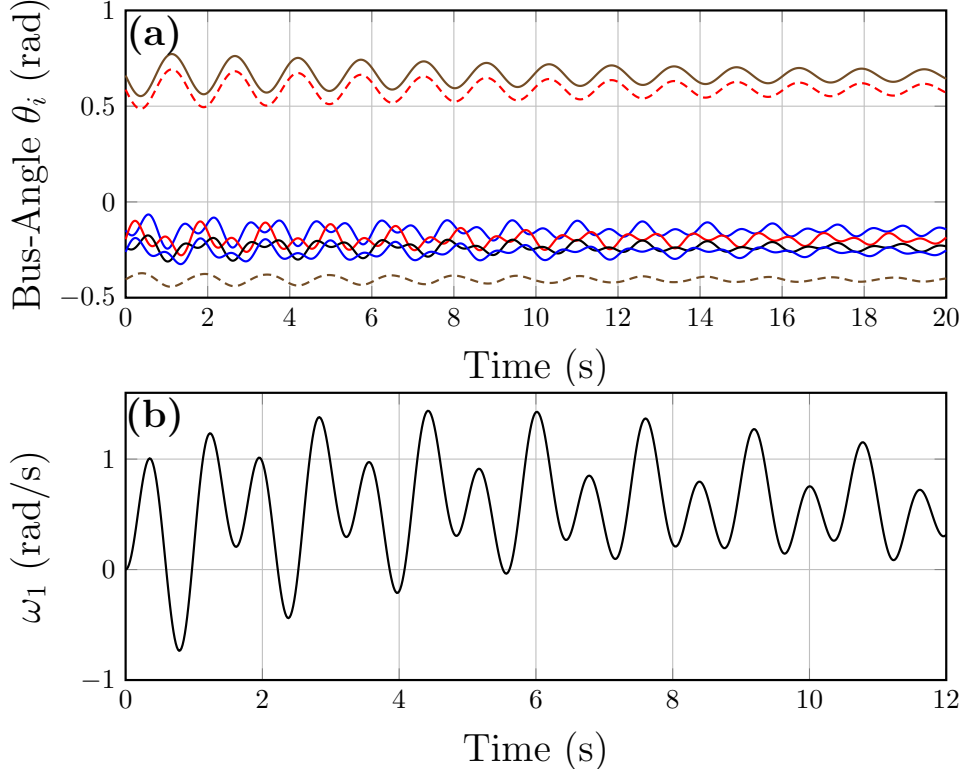


Figure 3.5: (a) Bus-angle oscillations and (b) angular frequency deviation ω_1 for G_1 following the disturbance.

(2.19) are $m(N_s/2)$ and $N_s/2$, i.e. $\#rows \geq \#col$. Thus, first, this numerically confirms Theorems 2.2.1 and 2.2.2, and second, connects DMD, Arnoldi KMD and Prony KMD.

The averaged eigenvalues identified in the two discussed cases are shown in Fig. 3.4. In Fig. 3.4 (a) we clearly see the influence of the so-called bias error for DMD [85], depending on the SNR. The largest discrepancies correspond to the highest noise levels (low SNR). In Fig. 3.4 (b) it is evident that all algorithms identify the eigenvalue λ_1 accurately regardless of the noise added to the last snapshot.

3.4.2 Power system example

Here, an evaluation is conducted by utilizing data from simulation of the power system shown in Fig. 3.1. Two modes with frequencies $f_1 = 1.2541$ Hz and $f_2 = 0.6566$ Hz are found from eigenanalysis of the linearized DAEs (1.1) which govern the dynamics. The first one is a so-called local mode characterizing an oscillation between G_1 and G_2 , and the second mode the inter-area mode between G_1, G_2 in Area 1 and G_3 in Area 2. Now we set $\omega_{2,0} = 4 \times 10^{-3} \omega_s$ to perturb the system, giving the bus angle dynamics as shown in Fig. 3.5 (a), which are used as test data for the evaluation. Later on, the data in Fig. 3.5 (b) are used as well.

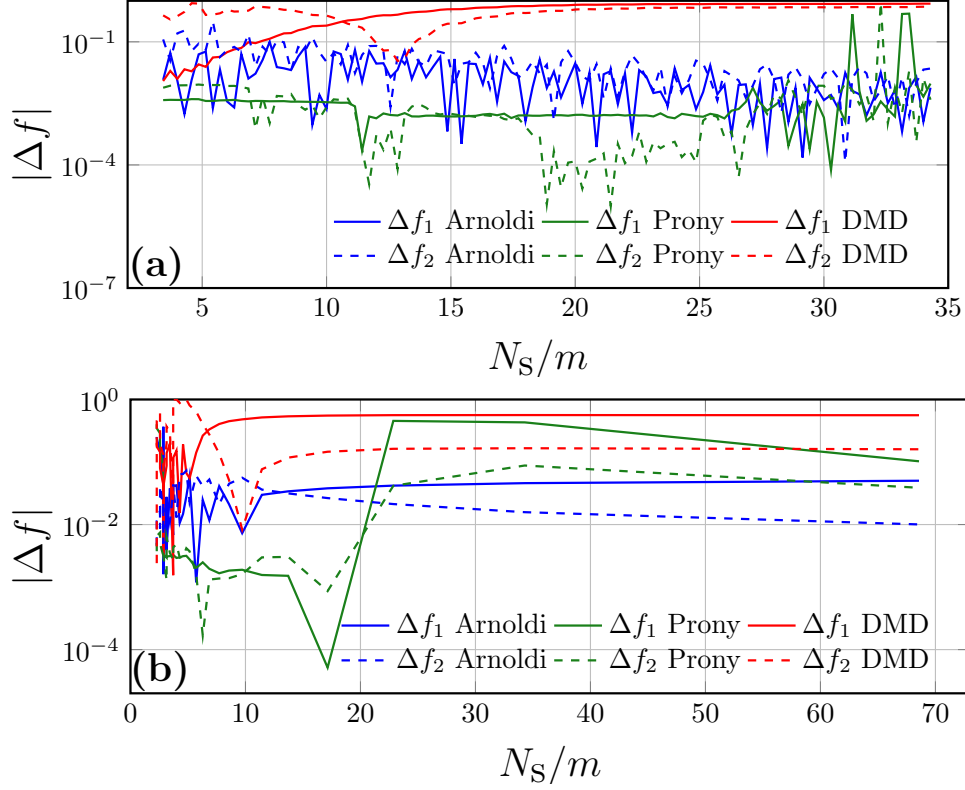


Figure 3.6: Deviation in frequency from linear modes $|\Delta f|$ for the data shown in Fig. 3.5 (a) for (a) with window time-length successively increased from 2 to 20 s, and f_s fixed at 12 Hz; and (b) with window time-length 8 s and f_s varied between 2 and 60 Hz.

Now, we compare the performance of the algorithms in identifying the modal frequencies identified from the linearization. The results on $|\Delta f|$ are given in Fig. 3.6 (a) for a window time-length increasing from 2 to 20 s. The sampling frequency f_s is fixed at 12 Hz. The number of snapshots is denoted by $N_s = 2N$, and the number of measurement locations by $m = 7$. It can be seen that DMD performs unsatisfactory compared with Arnoldi KMD and Prony. In particular, $|\Delta f|$ for DMD seems to diverge as N_s/m becomes large, although $|\Delta f_2|$ is initially large and shows an improvement around $N_s/m = 13$. Arnoldi KMD fluctuates more than Prony for $N_s/m < 25$, but as N_s grows further, $|\Delta f|$ for Prony starts increasing and the result fluctuates. It can be speculated that the advantage of vector Prony here stems from the expanded subspace on which the modes are defined. However, when the number of snapshots N_s increases, no new information is provided since the signals are basically repetitive and only contain two frequencies, and the increased amount of data seems to deteriorate the effectiveness of the algorithm.

Now, we look at $|\Delta f|$ as f_s varies between 2 and 60 Hz, thereby increasing the ratio N_s/m . The temporal length of the data is set to 8 s, and the results are shown in Fig. 3.6 (b). It is noted that the Prony algorithm performs better than the other ones when

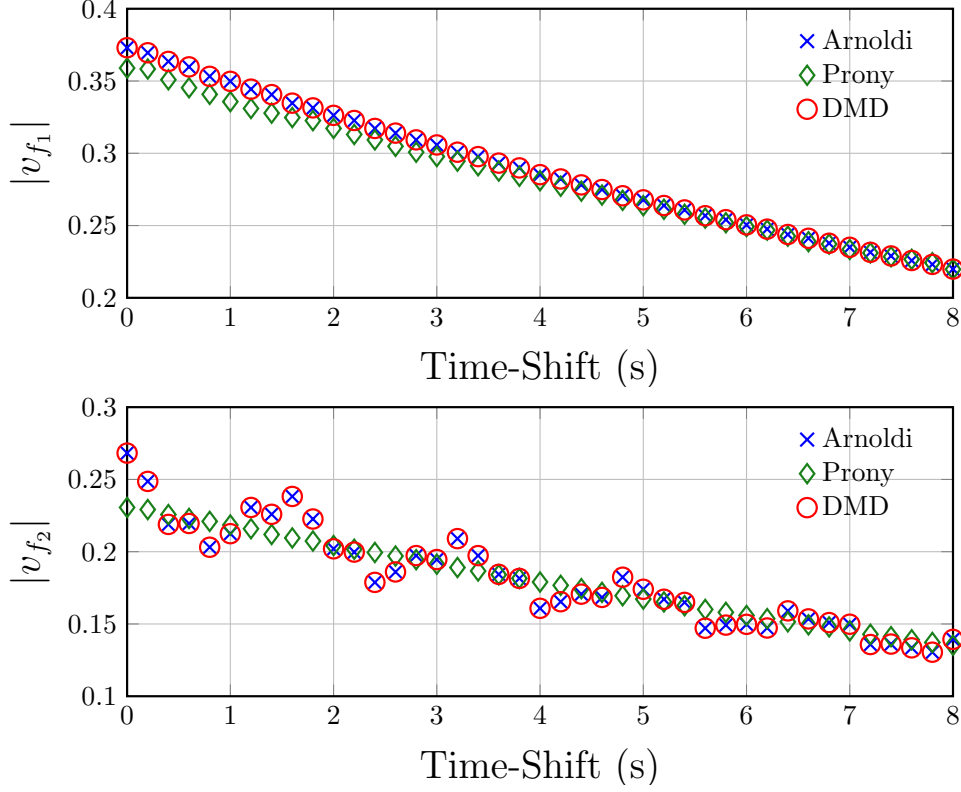


Figure 3.7: Components of $\mathbf{v}_{f_i}$, each corresponding to a time-shifted measurement of the data in Fig. 3.5 (b). f_i is either f_1 or f_2 from eigenanalysis of DAEs.

f_s is low and hence the number of snapshots, and experiences a peak around $N_s/m \approx 23$ and levels out. DMD does not seem at all successful in identifying the frequencies except for sporadically at low N_s/m . Both Arnoldi and Prony identify the frequencies well even for larger f_s ; $|\Delta f_1|$ for Prony and Arnoldi stops at 0.10 and 0.05, respectively.

Let $\mathbf{v}_{f_i}$ denote a vector (or mode) which is one of $\tilde{\mathbf{V}}_i$, $\hat{\mathbf{V}}_i$, or $\mathbf{c}_i := \phi_i(\hat{\phi}_i \mathbf{y}_0)$, see Section 2.2, oscillating with a frequency determined by associated eigenvalue λ_i , close to one of the identified f_i from eigenanalysis. Time-shifted snapshots (3.5) are now constructed from the data in Fig. 3.5 (b), such that $m = 41$ and $N_s = 40$. Arnoldi KMD and DMD generate the exact same decomposition in this case. In Fig. 3.7, $|\mathbf{v}_{f_i}|$'s are shown against the time-shift corresponding to the time-shifts of components of $\mathbf{v}_{f_i}$. For f_1 , the results of all algorithms become very close, and a steady decay of $|\mathbf{v}_{f_1,m}|$ as m increases is evident, where $\mathbf{v}_{f_1,m}$ denotes the m -th component of $\mathbf{v}_{f_1}$, time-shifted 0.2 seconds from the previous component. For f_2 , the results are still close but Arnoldi KMD/DMD modes fluctuate around the Prony mode which maintains a stable decrease with increased time-shift. These results show that the attenuations of the oscillations occur in a close to linear manner. However, the actual damping of oscillations occurs in an exponential manner if the power system dynamics are well described by small-signal analysis close to the equilibrium point. The results in Fig. 3.7 thus show that for the time-interval and data

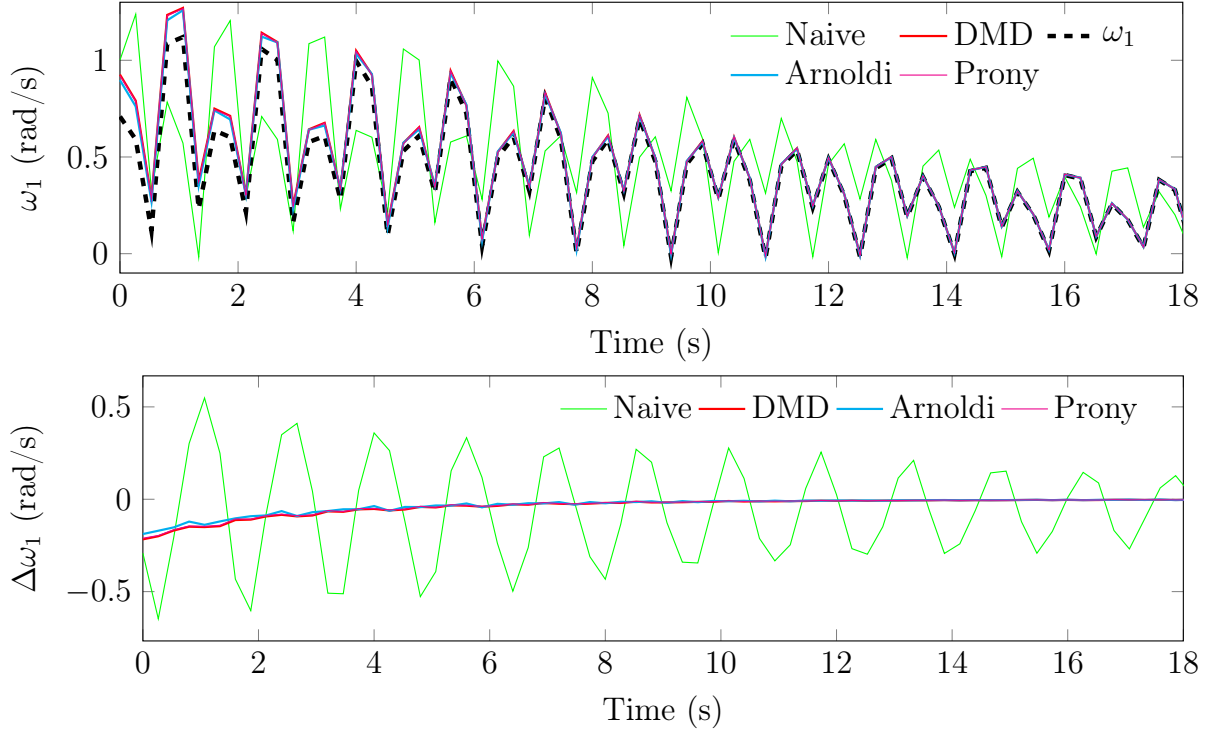


Figure 3.8: Predictions of ω_1 (see Fig. 3.5 (b)) via KMD/DMD and prediction errors.

considered, the exponential decrease can be approximated well by a linear trend line. This example clearly illustrates that all the considered techniques basically generate the same result when the subspace constructed by the snapshots is appropriately expanded.

3.4.3 Future prediction of power system dynamics

Now, the signal in Fig. 3.5 (b) is considered again. The same KMD/DMD algorithms will be applied and iterated nine times forward in time according to (3.6). An example of future predictions is shown in Fig. 3.8, together with associated prediction errors. Note that the “naive” predictor predicts a future value τ seconds later equal to what was last observed: $\hat{y}_{i+\tau} := y_i$. In the shown example, all algorithms predict the signal close to equally well, with the larger errors occurring in the beginning, due to larger and more nonlinear oscillations close to the time of the fault. Now, a more thorough evaluation of all algorithms is conducted for different values of spatial dimension m and number of snapshots N_s . The results are shown in the upper row of Fig. 3.9. The color indicates the Root Mean Square Error (RMSE) as a percentage of the RMSE of the naive predictor. The results clearly indicate smaller errors for Arnoldi and Prony KMD, particularly for small spatial dimensions. Clear improvements in prediction when the spatial dimension exceeds 4 for DMD, and 3 for Arnoldi, are also noted. This is closely related to the dimensionality of the power system dynamics which involves six state variables, and two

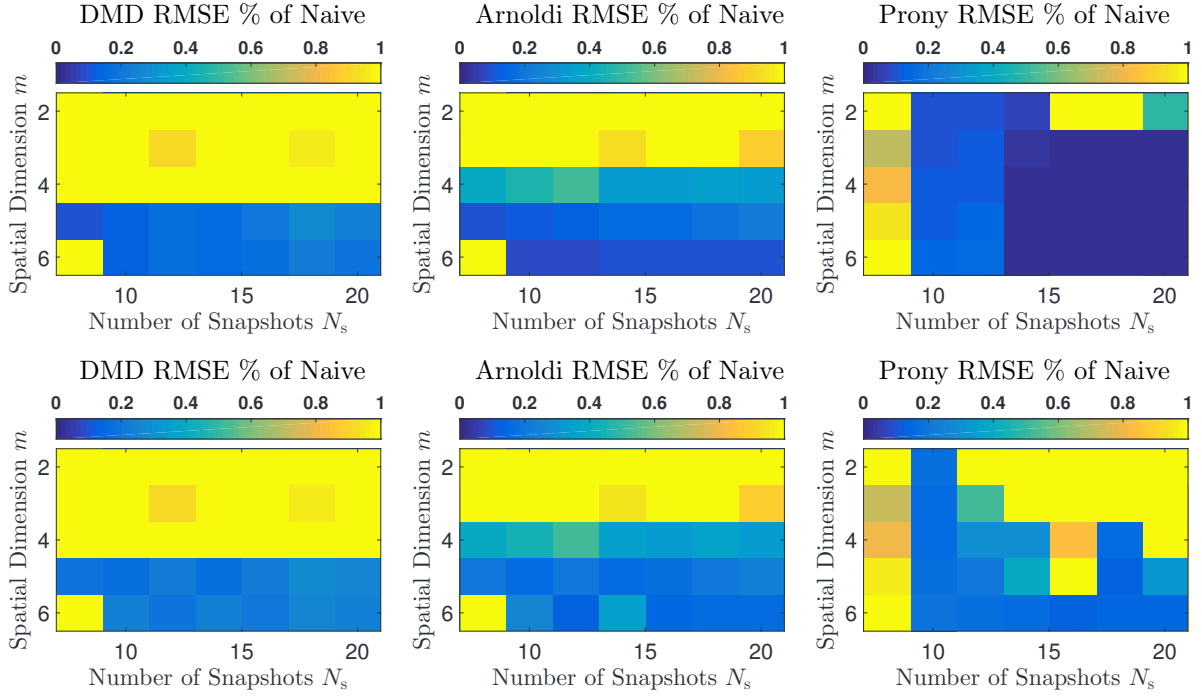


Figure 3.9: Reductions in RMSE for predictions of ω_1 (Fig. 3.5 (b)) compared to the naive predictor. Upper row without noise, and lower row with Gaussian noise, SNR = 40 dB.

complex conjugate pairs of eigenvalues are identified from linearization, as discussed in Section 3.4.2. We can thus expect that we would need a spatial dimension of around $m = 4$ to appropriately capture these dynamics. Now, we run the exact same experiment again with added noise. The noise is added with a SNR of 40 dB. The added noise has less effect on DMD and Arnoldi KMD, while the Prony KMD results have worsened. Previously, the benefit of expanding the subspace on which the Koopman operator is projected by using a Hankel-like matrix was discussed. These results show that when the subspace is already expanded enough, no additional benefit might be achieved by trying to stack more time-shifted measurements or vectors, and it might even deteriorate the performance of doing so, in particular if noise is present.

3.4.4 Future prediction of wind speeds

Here, we use wind speed predictions in Japan from CReSS [63], which can be considered as highly nonlinear and complex data. The performances of the algorithms are again evaluated by predicting future values of the time-series. Figure 3.10 shows wind speed fluctuations from one location in Japan, as deviations from the monthly mean, re-sampled with a sampling period of 12 h. The re-sampling substantially reduced autocorrelation of the time-series. Now, with the same methodology as in the previous section, we predict future values with different values on spatial dimension m and number of snapshots N_s . Again, the prediction results are compared against the naive predictor. In the upper and lower rows of Fig 3.11, we show the results for one and two step-ahead predictions,

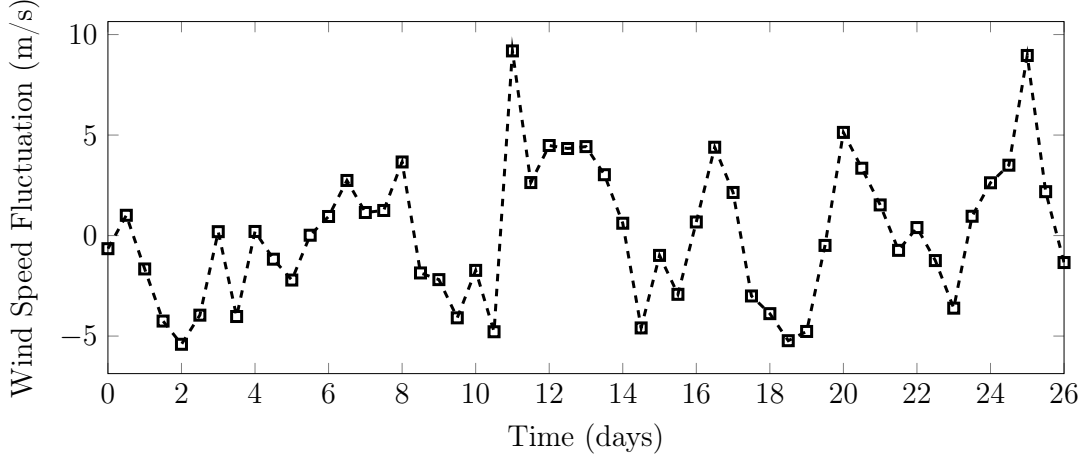


Figure 3.10: Predicted wind speed fluctuations from CReSS.

respectively. They correspond to 12- and 24-h ahead predictions. In the case of 12-h prediction, we achieve a maximum reduction in RMSE from the naive predictor of about 15–20% for DMD and Arnoldi, while Prony does not show any improvement. In the case of 24-h prediction by two iterations, where larger variations in the wind speed time-series are observed, larger improvements with KMD/DMD compared to naive predictions are obtained. This indicates that some dominant characteristics in the dynamics are captured. Improvements of up to 35% are achieved with DMD and Arnoldi KMD, and Prony achieves about 30% reduction at best. Again, we see that a larger spatial dimension is not always the best, as is achieved with Prony. This example further stressed the need for appropriate adjustments in spatial dimension, number of snapshots, choice of algorithm, and identification of number of dynamically relevant modes, when applying these types of algorithms to experimental data.

3.5 Conclusions

This chapter performed a numerical evaluation of three algorithms related to KMD, by applications to noisy data from a linear system, power system oscillations, and wind speed fluctuations. Their performance was evaluated by their accuracy in identification of eigenvalues, modal frequencies, modes, and by their ability to predict future values. The results were discussed with support from a new unified interpretation of the algorithms provided in Chapter 2. Utilizing time-shifted measurements, an inherent feature of the vector Prony KMD has the potential to substantially increase the subspace spanned by sampled data, on which the decomposition is defined, which allows for a more accurate decomposition. For example, in the case of data from linear dynamics with added noise, the spatial dimension was identified as a dominant factor in reducing noise. In fact, it was possible to make all algorithms produce the same result if the input data to them were

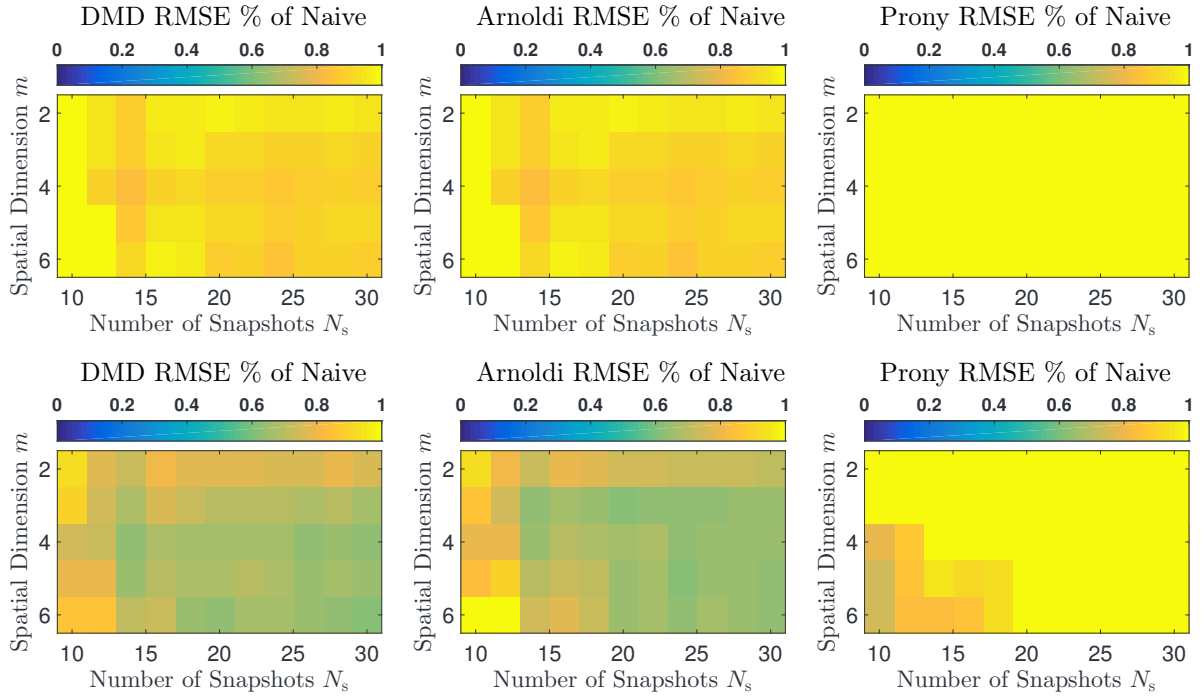


Figure 3.11: Reductions in RMSE from the naive predictor for predictions of the wind speeds in Fig. 3.10. Upper row: 12-h predictions. Lower row: 24-h predictions.

Table 3.1: A summary of numerical results by grading the performance of the three algorithms applied to different data sets. Performance is graded from ‘Excellent’ to ‘Poor’.

Data type	Algorithm		
	DMD	Arnoldi	Prony
Linear	Excellent	Excellent	Excellent
Linear + noise	Good	Good	Excellent
Weakly nonlinear (power system)	Good	Good	Good
Highly nonlinear (wind speed)	OK	OK	Poor

constructed according to the Hankel matrix, due to a larger spatial dimension than the number of snapshots. This is a consequence of Theorem 2.2.2. Also, it was shown that vector Prony works well in frequency identification of modes when the ratio between the number of snapshots and spatial dimension is on the lower side, for almost linear systems (weakly nonlinear). For data with high sampling frequency or consisting of many snapshots, the algorithm seems to perform worse. DMD is not recommended for application to data directly when the ratio between temporal and spatial dimension is large. Advantages of standard DMD can be leveraged when there is a substantial amount of data, where DMD modes can be projected on a smaller subspace retaining the dominant dynamic characteristics, which also has the effect of reducing measurement noise. In Table 3.1, an attempt has been made to summarize the numerical results by grading the performance of the algorithms applied to the different data sets.

Chapter 4

Modeling of wind power systems

The purposes of this chapter are to explain the fundamental ideas of wind power generation and to provide basic models of Wind Turbines (WT) and Wind Farms (WF), which are used in the forthcoming chapters. Modern WTs are often so-called variable-speed type WTs, capable of adjusting their rotational speed independently of the wind speed [66]. This feature is very attractive since optimal power extraction depends on the ratio between the speeds of the WT and the incoming wind, called the *Tip Speed Ratio* (TSR) [66]. WTs achieve optimal power extraction below rated wind speed by keeping the TSR constant by controlling their rotational speed [66]. Above rated wind speed, output power can be limited to rated value by adjusting the so-called *pitch angle* [87] of the blades. Thus, it can be concluded that the two most important controls in modern WTs for the power generation are speed and pitch angle controls. Additionally, a large number of other controls are required: yaw-control to align the turbine with the wind direction, voltage control, and controls to reduce structural loads [87–90], to name a few. Furthermore, modern WTs are capable of ancillary services such as participation with inertia (often called synthetic inertia) to the system which helps reducing frequency drops in case of disturbances, and reactive power support [33, 50]. There are mainly two types of variable-speed WTs, divided into two categories based on generator type: Doubly Fed Induction Generators (DFIG), and synchronous generators connected to the grid via AC/AC converters [91, 92]. In recent years, variable-speed WTs rated at several megawatts have become standard. Currently, the largest commercial WTs are rated at around 8 MW, and WT ratings are expected to reach 15–20 MW by 2020 [93]. Thus, with sophisticated control capabilities and high power ratings modern WTs are capable of replacing conventional generation.

WTs are very complex systems incorporating a large number of electrical and mechanical subsystems, controls, and engage in complex interactions with turbulent wind fields. Thus, there are many aspects to the study of WTs and wind power generation. This dissertation mainly focuses on the power conversion and output of WTs. In this case, the simplest model of a WT is given by a nonlinear curve relating the wind speed magnitude to the power output, and is commonly known as a *power curve* [66]. An example of a power curve derived from a simple model will be shown in Fig. 4.3 (b), which has a cer-

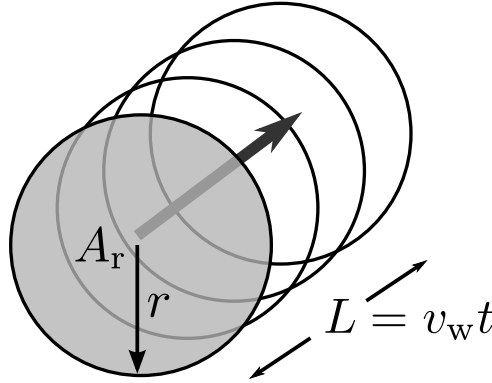


Figure 4.1: Depiction of air moving through a cross-section A_r thought of as the swept area of a turbine, forming a hypothetical cylinder with a volume determined by A_r , the wind speed v_w , and a time-duration t .

tain cut-in wind speed where the WT starts producing power, and a cut-off wind speed where the WT is shut off to protect it from damage. The cut-off wind speed is a design parameter, and here it was set to 25 m/s (typical cut-off wind speed is 20–25 m/s [94]). This type of model is useful in rough estimations of the *long-term output*, since it represents the *long-term average*. In the short-term, on the scale of seconds, characterized by fast wind speed fluctuations, a simple power curve is, however, not sufficient due to the effect of the WT's inertia and control systems in the power conversion [64]. To assess the short-term variability of WT and WF outputs, physical models including the mechanical system and standard controls are thus needed. In this dissertation, both power curves and more detailed dynamical models of WTs are utilized, depending on the time-scale of interest.

This chapter is organized as follows. The basic modeling of the mechanical systems of WTs is explained in Section 4.1. The models used to simulate the DFIG and Permanent Magnet Synchronous Generator (PMSG) WTs are outlined and explained in Sections 4.2 and 4.3, respectively. Modeling and initialization of simulation of a WF are explained in detail in Section 4.4, and simulation results are presented. The chapter is concluded in Section 4.5.

4.1 Wind power and wind turbine modeling

4.1.1 Wind power

The WT blades capture kinetic energy in the wind and convert it to mechanical torque, which rotates the blades around their axis causing the shaft to rotate. The shaft transfers the mechanical power to a generator inside the WT nacelle, and it then gets converted into electric power which is fed into the grid. According to the definition of kinetic energy,

we can express the wind's kinetic energy E_{wind} as follows:

$$E_{\text{wind}} = \frac{mv_w^2}{2},$$

where v_w is the the wind speed (m/s) and m (kg) the mass. The mass m depends on the volume V (m³) and the air density ρ (kg/m³) giving $m = V\rho$ (kg). The air density depends on the atmospheric pressure p_a (kg/(ms²)) and is approximately expressed as follows (assuming ideal gas): $\rho = p_a/(RT)$ where R (J/(K·mol)) is the gas constant and T the temperature measured in Kelvin (K). Now, considering air moving through a cross-section thought of as the swept area $A_r = \pi r^2$ of a WT, where r (m) is the radius of the WT, and forming a cylinder as depicted in Fig. 4.1, we can express the volume by $V = A_r L = A_r v_w t$, with L and t denoting the length and time-duration, respectively. With these definitions, the wind power is given as follows:

$$P_{\text{wind}} = \frac{E_{\text{wind}}}{t} = \frac{1}{2} \rho A_r v_w^3. \quad (4.1)$$

4.1.2 Mechanical model of wind turbines

A WT is only able to extract a certain percentage of the power available in the wind given by the previously derived (4.1). The exact amount is determined by the so-called power coefficient of the WT denoted by c_p ($= P_w/P_{\text{wind}}$) [87], which is a function of the TSR λ and pitch angle β , and P_w denotes the WT's mechanical power. The power coefficient describes the aerodynamic efficiency of the WT, and its highest theoretical value is exactly 16/27, which is the so-called Betz's limit [87]. With c_p , we can express the mechanical power extracted from wind with a WT in accordance with [92] as follows:

$$P_w = \frac{1}{2} c_p(\lambda, \beta) \rho A_r v_w^3, \quad (4.2)$$

where the TSR λ is calculated as $\lambda := (\omega_t r)/v_w$, where ω_t (rad/s) is the mechanical angular velocity of the turbine rotor. Power curves for different WTs are generally quite similar and c_p can be approximated by fixing a set of constants in the following equation [92]:

$$c_p = c_1 \left(\frac{c_2}{\lambda_i} - c_3 \beta - c_4 \beta^{c_5} - c_6 \right) \exp \left(-\frac{c_7}{\lambda_i} \right), \quad (4.3)$$

with λ_i defined as

$$\frac{1}{\lambda_i} = \frac{1}{\lambda + c_8 \beta} - \frac{c_9}{\beta^3 + 1}. \quad (4.4)$$

Figure 4.2 shows a plot of c_p as a function of β and λ by setting the constants c_i same as in [95]. In this case, the maximum power extraction is obtained for $\lambda \approx 7$ and $\beta = 0^\circ$. An optimal relation between P_w and ω_t , $P_w^*(\omega_t)$ is derived by tracking the maximum value of (4.2) for different wind/rotor speeds, called Maximum Power Point Tracking (MPPT):

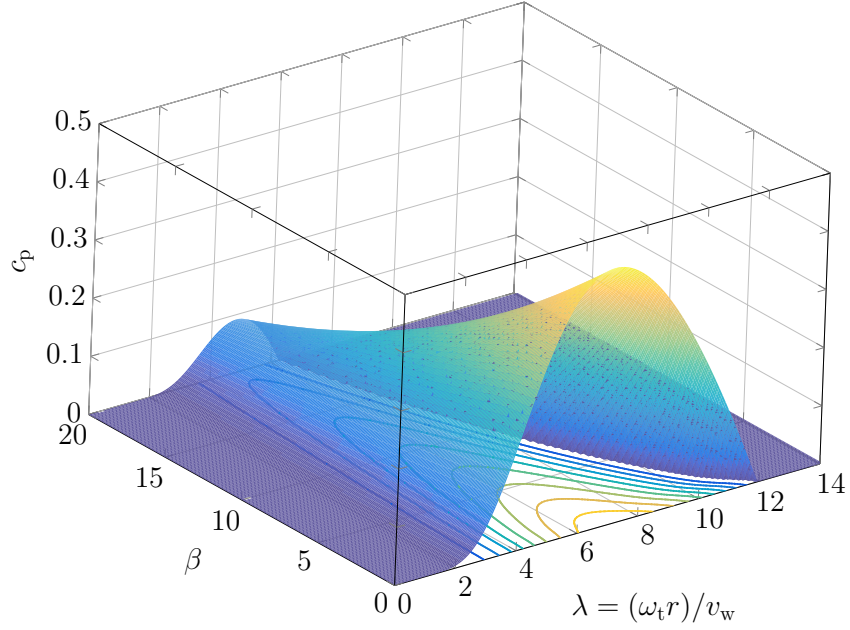


Figure 4.2: Surface plot of the power coefficient c_p as a function of the pitch angle β and the tip speed ratio λ . It is easy to see that maximum power extraction is obtained for $\lambda \approx 7$ and $\beta = 0$.

see Fig. 4.3 (a). $P_w^*(\omega_t)$ is used to control the turbine at optimal power. The power curve $P_w(v_w)$ is now derived by solving

$$P_w(\lambda(v_w, \omega_t)) - P_w^*(\omega_t) = 0, \quad (4.5)$$

for a number of wind speeds v_w , and is shown in Fig. 4.3 (b).

4.1.3 Dynamics of mechanical system

Dynamics of the electro-mechanical system of WTs are often described by a two-mass model with one mass representing the turbine's rotor and the other one the generator, e.g. in [96] given by

$$\begin{aligned} \frac{d\omega_t}{dt} &= \frac{1}{2H_t}(T_t - T_g), \\ \frac{d\omega_g}{dt} &= \frac{1}{2H_g}(T_g - T_e), \\ \frac{d\gamma}{dt} &= \omega_b(\omega_t - \omega_g), \end{aligned} \quad (4.6)$$

where $T_t = P_w/\omega_t$ is the turbine torque, $T_g = k_s\gamma + D\frac{d\gamma}{dt}$ the generator's mechanical torque, H_g, H_t inertia constants, k_s shaft stiffness, D a damping constant, γ shaft twist angle, $\omega_b = 2\pi f_{\text{sys}}$, and T_e the electrical torque. f_{sys} (Hz) is the system frequency. The model

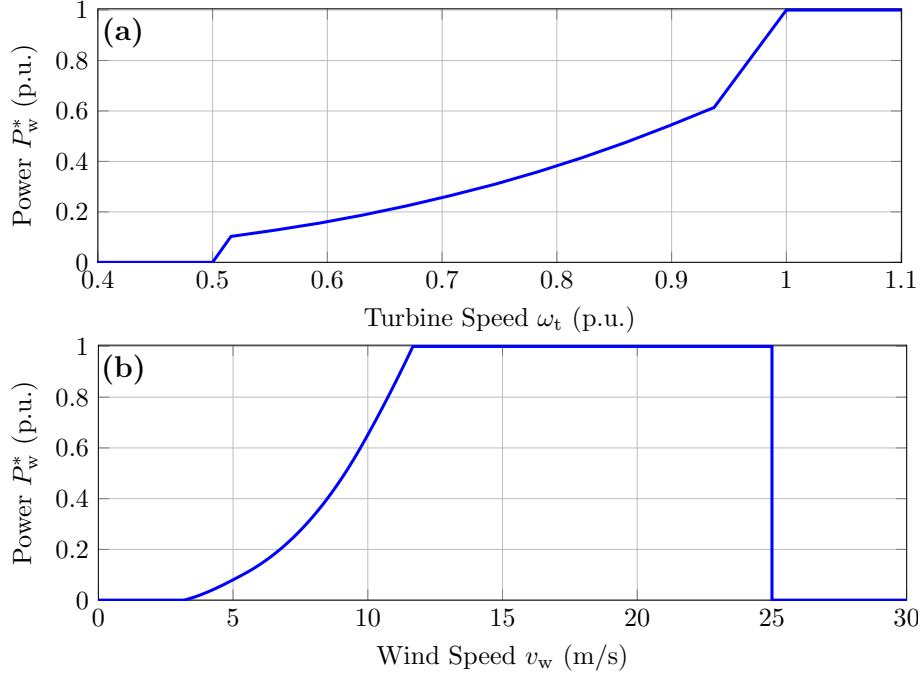


Figure 4.3: Examples of (a) optimal power vs. speed curve and (b) derived power curve.

(4.6) is often used for DFIG generators, and in particular when looking at protection following grid faults [92]. Otherwise, it is often sufficient to use a one-mass model [92]:

$$\frac{d\omega_m}{dt} = \frac{1}{2H_m}(T_m - T_e), \quad (4.7)$$

with T_m denoting the mechanical torque, T_e the electrical torque, and H_m the lumped inertia constant, given in [92] as $H_m = H_t/\eta^2 + H_g$, where η is the gear-ratio.

4.2 Doubly-fed induction generator

The following model of a Doubly-Fed Induction Generator (DFIG) is based on [97] and is the same model used in [98]. The model incorporates speed control to keep the WT close to optimal power extraction, voltage control, and pitch angle control.

The real and reactive power outputs from the DFIG are in the so-called dq reference frame (d stands for direct and q for quadratic, see Section 4.A) given by

$$\begin{aligned} P_e &= v_{ds}i_{ds} + v_{qs}i_{qs} + v_{dr}i_{dr} + v_{qr}i_{qr}, \\ Q_e &= -\frac{x_m V_t i_{dr}}{x_s + x_m} - \frac{V_t^2}{x_m}, \end{aligned} \quad (4.8)$$

where (v_{ds}, v_{qs}) and (v_{dr}, v_{qr}) are the direct and quadratic voltages of the stator and rotor, respectively, (i_{ds}, i_{qs}) and (i_{dr}, i_{qr}) the direct and quadratic currents of the stator

and rotor, x_s and x_m the stator and magnetizing reactances in p.u., and V_t the terminal bus voltage (see Fig. 4.4 (a)). The electric torque is given by

$$T_e = x_m(i_{qr}i_{ds} - i_{dr}i_{qs}). \quad (4.9)$$

With (4.7) governing the dynamics of the mechanical systems, the differential equations used to simulate the DFIG are in accordance with [97] the following:

$$\begin{aligned} \frac{dv_w}{dt} &= \frac{(v_{in} - v_w)}{\tau}, \\ \frac{d\omega_m}{dt} &= \frac{1}{2H_m}(T_m - T_e), \\ \frac{di_{qr}}{dt} &= \left(\frac{-(x_s + x_m)}{x_m V} P_w^*(\omega_m)/\omega_m - \frac{di_{qr}}{dt} \right) \frac{1}{T_\epsilon}, \\ \frac{di_{dr}}{dt} &= K_v(V_t - V_{ref}) - \frac{V_t}{x_m} - i_{dr}, \\ \frac{d\beta}{dt} &= (K_p \alpha(\omega_m - \omega_{ref}) - \beta) \frac{1}{T_p}. \end{aligned} \quad (4.10)$$

The first equation filters the input wind speed v_{in} through a low pass filter with time-constant τ set at 4 s [92] to give the “effective” wind speed v_w . This acts like the natural filtering of high frequency components over the rotor surface [92]. The quadratic rotor current i_{qr} is used to control rotor speed according to an optimal power vs. speed curve as shown in Fig. 4.3 (a), and T_ϵ is a small time constant. The direct rotor current i_{dr} is used to control V_t at reference value V_{ref} . The pitch angle β is controlled with the proportional controller with gain K_p depicted in Fig. 4.5 which incorporates a low-pass filter with a time constant T_p , and the function $\alpha(\Delta\omega) = \text{round}(\Delta\omega \times 10^p) \times 10^{-p}$, i.e. $\Delta\omega = \omega_m - \omega_{ref}$ rounded to p decimals, where $\text{round}(\cdot)$ rounds a value to the closest integer.

4.3 Permanent magnet synchronous generator

Here, the other common variable speed type WT will be described: PMSG. It incorporates the same type of speed and pitch angle controls as the DFIG, and the same mechanical system. There are two advantages of the PMSG. First, usage of full scale AC/AC converters increases efficiency [93]. Second, if a multiple pole synchronous generator is used, it can be operated without a gearbox, called *direct drive*, and which reduces maintenance costs [99], and is more robust against power system faults [93]. Reduced maintenance is attractive for offshore wind power where maintenance is more expensive [93]. It will therefore be used in an example of a numerical simulation of a hypothetical offshore WF in the next section.

The electric output power of the generator in the dq reference frame is given by

$$\begin{aligned} P_s &= v_{ds}i_{ds} + v_{qs}i_{qs}, \\ Q_s &= v_{qs}i_{ds} - v_{ds}i_{qs}, \end{aligned} \quad (4.11)$$

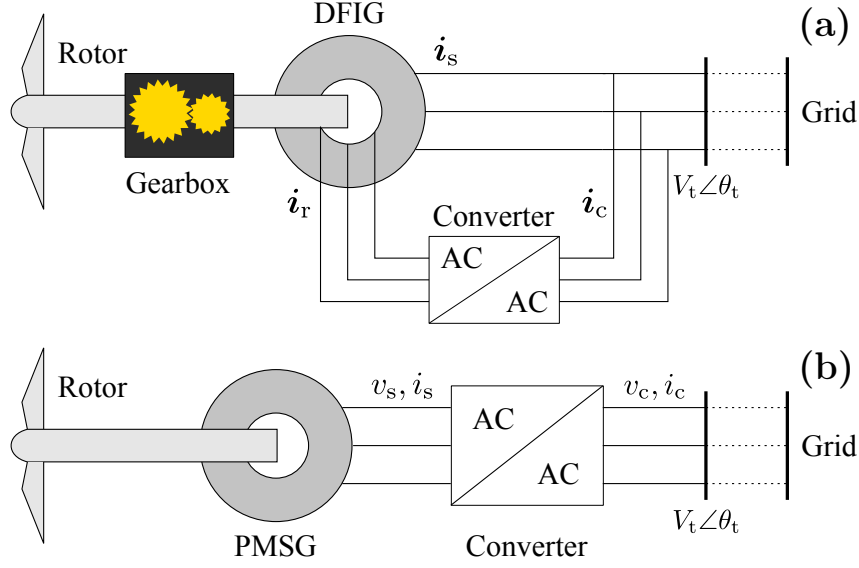


Figure 4.4: (a) Doubly-Fed Induction Generator (DFIG) and (b) direct drive (gearless) WT with a PMSG connected to a power grid.

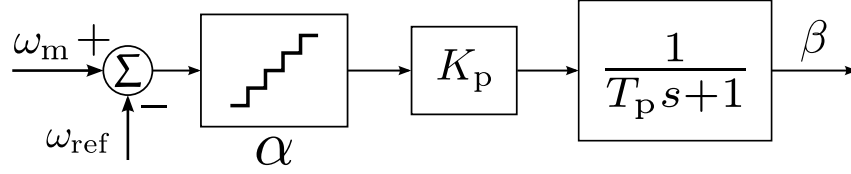


Figure 4.5: Illustration of the utilized pitch angle control, where s is the Laplace transform variable. ω_{ref} is the rotor speed reference.

where the subscript s denotes quantities on the synchronous generator side: see Fig. 4.4 (b). The electrical torque T_e is given as

$$T_e = i_{qs} i_{ds} (x_q - x_d) + i_{qs} \Psi_p, \quad (4.12)$$

where x_d and x_q are the synchronous reactances in p.u. For control strategy purposes we let $T_e := P_w^* / \omega_m$ which yields

$$P_w^* = \omega_m i_{qs} (i_{ds} (x_q - x_d) + \Psi_p), \quad (4.13)$$

where ω_m is given in p.u. The following four equations with six unknown variables (v_{ds} , v_{qs} , i_{ds} , i_{qs} , ω_m , P_s) describe the PMSG considered here [7, 97, 100]:

$$\begin{aligned} v_{ds} &= -R_s i_{ds} + \omega_m x_q i_{qs}, \\ v_{qs} &= -R_s i_{qs} - \omega_m (x_d i_{ds} - \Psi_p), \\ P_s &= v_{ds} i_{ds} + v_{qs} i_{qs}, \\ P_s &:= P_w^* = \omega_m i_{qs} \Psi_p, \end{aligned} \quad (4.14)$$

where R_s is the stator resistance and Ψ_p the permanent flux. Because the converter decouples reactive power exchange between the grid and the generator, we let Q_s be arbitrary and is thus not included in the above system of equations. The last equation is a simplification of (4.13) where the term $i_{ds}(x_q - x_d)$ is assumed to be small compared to Ψ_p and omitted for simplicity. The output power of the converter which is considered as ideal is given by

$$\begin{aligned} P_c &= v_{dc}i_{dc} + v_{qc}i_{qc}, \\ Q_c &= v_{qc}i_{dc} - v_{dc}i_{qc}, \end{aligned} \quad (4.15)$$

with $v_{dc} = V_t \sin(-\theta_t)$ and $v_{qc} = V_t \cos(\theta_t)$, where V_t and θ_t are the terminal voltage and voltage angle (see Fig. 4.4 (b)) and the subscript c denotes converter side quantities. The differential equations for each WT are similarly to (4.10) given as [92, 97]:

$$\begin{aligned} \frac{dv_{w,i}}{dt} &= \frac{v_{in,i} - v_{w,i}}{\tau}, \\ \frac{\omega_{m,i}}{dt} &= \frac{P_{w,i}/\omega_{m,i} - T_{e,i}}{2H_m}, \\ \frac{di_{qs,i}}{dt} &= \frac{i_{qsr,i} - i_{qs,i}}{T_s}, \\ \frac{di_{dc,i}}{dt} &= \frac{K_v(V_{ref,i} - V_{t,i}) - i_{dc,i}}{T_v}, \\ \frac{d\beta_i}{dt} &= (K_p\alpha(\omega_m - \omega_{ref}) - \beta) \frac{1}{T_p}, \end{aligned} \quad (4.16)$$

The current i_{qs} is controlled such that it follows the optimal power and rotor speed relation according to the last equation of (4.14), and i_{dc} is used for voltage control where V_{ref} denotes the reference voltage. T_p, T_v, T_s are time constants and K_p, K_v control gains. The “references” for generator currents i_{qs} and i_{dc} for each generator are

$$i_{qsr,i} = \frac{P_{w,i}^*}{\omega_{m,i}\Psi_p}, \quad i_{dcr,i} = \frac{i_{dc,i}V_{t,i}\sin(\theta_{t,i}) + P_{s,i}}{V_{t,i}\cos(\theta_{t,i})},$$

where $i_{qsr,i}$ is derived from the last equation of (4.14), and $i_{dcr,i}$ is derived from (4.15).

4.4 Initialization and simulation of a wind farm

Here, a WF with the nameplate capacity 60 MW will be simulated using detailed weather simulation data of Japan’s north-western coast (Aomori Prefecture), provided by the Cloud Resolving Storm Simulator (CReSS) model [63, 101]. At present, there are plans to construct multi-megawatt WFs near the west coast of northern Japan [102], which inspire this numerical case study. The WF simulated here comprises twenty 3 MW PMSGs, and is connected to the commercial grid represented by an Infinite Bus (IB). An IB is assumed to always maintain a constant voltage [7]. In the following, the procedure of initializing and simulating a WF is described, including a detailed description of the DAEs governing its dynamics.

4.4.1 Initialization of wind turbines

Wind speed data from CReSS at m locations are utilized, and thus it is desired to initialize simulations using initial wind speed measurements $\mathbf{v}_{w0} := [v_{w1,0}, \dots, v_{wm,0}]^\top$. Firstly, WTs are initialized with $\mathbf{v}_{w0}$ as input, and secondly load flow calculations and dynamic simulations are run. On the other hand, if wind speeds are generated as described in [103] with power outputs predefined or from the load flow solution, then (4.14) can be solved since P_s and ω_m (through the algebraic relationship between electric power and rotor speed) are known: see [104]. Here, simulation variables for each PMSG are initialized using an initial wind speed measurement $v_{w0} := v_{wi,0}$ as input as follows:

1. Mechanical power P_w from (4.2) with $v_w = v_{w0}$ is given by

$$P_{w0}(\lambda(\omega_m), \beta) = T_e \omega_m,$$

where the electrical torque T_e is given by (4.12).

2. Equation (4.14) is augmented with one new variable and three new equations. We let β be the new initialization variable and $P_s = P_w^*$. Then, using the Newton-Raphson (NR) method, (4.14) is solved together with

$$\begin{aligned} 0 &= P_{w0}(\lambda(\omega_m), \beta) - T_e \omega_m, \\ 0 &= P_w^*(\omega_m) - P_w^*, \\ 0 &= K_p(\omega_m - \omega_{\text{ref}}) - \beta, \end{aligned} \tag{4.17}$$

where the second equation describes the optimal power-speed relationship, and the last one is the pitch control equation in (4.16) with the time-derivative of the pitch angle set to zero.

3. Using v_{w0} , initial values can be derived for ω_m and P_w^* using data as shown in Fig. 4.3 (a). Further, the initial values of the NR iteration for remaining variables are set as $(V_{qs0}, i_{qs0}, V_{ds0}, i_{ds0}) = (\omega_{m0} \Psi_p, P_{w0}^*/V_{qs0}, \omega_{m0} x_q i_{qs0}, 0)$. The initial pitch angle β_0 is set with the last equation of (4.17) if $\omega_{m0} > 1$ or otherwise set to zero. Thus, a solution for the total seven variables $(v_{ds}, v_{qs}, i_{ds}, i_{qs}, \omega_m, P_w^*, \beta)$ can be obtained.
4. Solve the power flow equations (see the first two equations of (4.19)). Using (4.15) where Q_c is obtained from the power flow solution and $P_c = P_s = P_w^*$, initial values of the remaining differential and algebraic variables can be determined.

4.4.2 Wind farm modeling

Let us now simulate a WF comprising twenty identical WTs: see Fig. 4.6, and modeled as PMSGs. Transmission line parameters are same as in [100] but transferred to a 100 MVA system rating and the voltage levels stated in Fig. 4.6 are same as in [105]. The system is simulated with the state variables $\mathbf{x} = [\mathbf{v}_w^\top, \boldsymbol{\omega}_m^\top, \mathbf{i}_{qs}^\top, \mathbf{i}_{dc}^\top, \boldsymbol{\beta}^\top]^\top$ and algebraic variables $\mathbf{y} = [\boldsymbol{\theta}^\top, \mathbf{V}^\top, \mathbf{i}_{ds}^\top, \mathbf{i}_{qc}^\top, \mathbf{P}_w^{*\top}]^\top$. Note that e.g. $\boldsymbol{\omega}_m = [\omega_{m,1}, \dots, \omega_{m,20}]^\top$ and the other

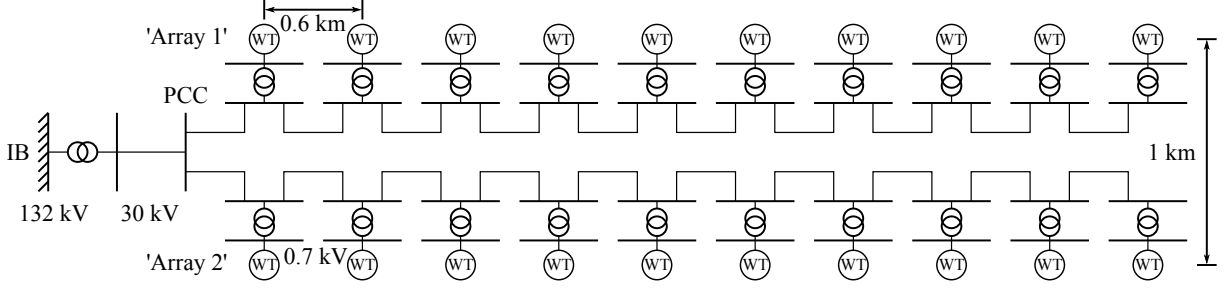


Figure 4.6: Simulated wind farm comprising twenty Wind Turbines (WTs) arranged in two arrays connected to an Infinite Bus (IB).

variables are defined analogously. $\mathbf{V}$ and $\boldsymbol{\theta}$ comprise the n voltage amplitudes and angles of buses (nodes) in the WF: see Fig. 4.6. The set of DAEs,

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{y}), \quad \mathbf{0} = \mathbf{g}(\mathbf{x}, \mathbf{y}), \quad (4.18)$$

is solved with a trapezoidal solver [97] implemented in MATLAB. The differential equations $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ are given by (4.16). In addition, the set of algebraic equations are the following:

$$\begin{aligned} \mathbf{0} &= \mathbf{P} + \mathbf{P}_{\text{GL}}, \\ \mathbf{0} &= \mathbf{Q} + \mathbf{Q}_{\text{GL}}, \\ \mathbf{0} &= \mathbf{P}_{\text{s}} - \mathbf{P}_{\text{w}}^*, \\ \mathbf{0} &= \mathbf{i}_{\text{qcr}} - \mathbf{i}_{\text{qc}}, \\ \mathbf{0} &= \mathbf{P}_{\text{w}}^*(\boldsymbol{\omega}_{\text{m}}) - \mathbf{P}_{\text{w}}^*. \end{aligned} \quad (4.19)$$

Here $\mathbf{P} = [P_1, \dots, P_n]^\top$ and $\mathbf{Q} = [Q_1, \dots, Q_n]^\top$ are the sums of power flows at buses given by $P_i = \sum_{j=1}^n V_i V_j Y_{ij} \cos(\phi_{ij} + \theta_j - \theta_i)$ and $Q_i = -\sum_{j=1}^n V_i V_j Y_{ij} \sin(\phi_{ij} + \theta_j - \theta_i)$, where $V_i \angle \theta_i$ is the voltage phasor at bus i and $|Y_{ij}| \angle \phi_{ij}$ is the admittance of the transmission line connecting bus i and j and $P_{\text{GL},i}$ (or $Q_{\text{GL},i}$) is $-P_{\text{G},i}$ (or $-Q_{\text{G},i}$) if i corresponds to a generator bus and $P_{\text{L},i}$ (or $Q_{\text{L},i}$) if load bus.

4.4.3 Simulation results

Now, weather prediction data from CReSS are utilized, and simulations are conducted. Twenty measurement points are selected and a 300 s window of data is used as input to the test system. WTs are spaced 0.6 km apart in each array and the distance between the two WT-arrays is 1 km as shown in Fig. 4.6. This configuration is similar to [105]. The wind speed data are taken from a simulated weather scenario where a cold weather storm hits the west coast of Japan, as illustrated in Fig. 4.7. Seen from left to right and starting from the first row, the first snapshot shows the wind profile at initial time and the last one at 2 hours later. The coloring of the figures vary from dark blue (lowest) to red (highest) and indicates the wind speed (see the color scales). It is clearly visible

Table 4.1: Parameters for the WT generator system used in wind farm simulation: similar to [92, 97].

Wind turbine parameter	Value
Rated power	3 MW
Nominal speed	15 rpm
Number of poles p	96
Rotor diameter $d = 2r$	96 m
Direct axis reactance x_d	1.21 p.u.
Quadratic axis reactance x_q	0.606 p.u.
Stator resistance R_s	0.02 p.u.
Inertia constant H_m	1.6 s
Permanent flux Ψ_p	1.4 p.u.
Pitch controller gain K_p	95
Voltage controller gain K_v	10
Pitch controller time constant T_p	4 s
Voltage controller time constant T_v	0.01 s
i_{qs} controller time constant T_s	0.01 s
Wind filter time constant τ	4 s

how the storm approaches the coast and finally gives rise to strong winds along the whole coastline. Input data to our simulation were for a 5 minute window that starts at the time of the second snapshot. Locations of the WTs in the WF are given by the black dots. Note that the locations do not exactly match any ongoing project around this area. The wind speeds have longitude and latitude components ($v_{\text{long}}, v_{\text{lat}}$) and here the absolute values $\tilde{v}_w = \sqrt{v_{\text{long}}^2 + v_{\text{lat}}^2}$ are used as input. Wind speed predictions have been acquired with a sampling period of 2 s. The time step in simulations is set to 10^{-2} s and thus interpolation between wind speed samples is performed. The wind speed sequences used as input to the simulations are shown in Fig. 4.8 (a) for the two arrays of turbines. The pitch angle control acts like a governor to limit extracted energy in the case of high wind speeds and is activated when ω_m exceeds 1 p.u. Thus, for initial wind speeds larger than ≈ 11.7 m/s, pitch control is activated ($\beta > 0$), the WT is operated at super synchronous speed ($\omega_m > 1$), and the output power P_w^* is kept at 1 p.u. In Fig. 4.8 (b), the rotor speeds ω_m are given. It is easy to see how the rotor speeds follow the behavior of the wind speeds and are kept close to 1 p.u. ($= \omega_{\text{ref}}$) for high wind speeds due to the pitch angle controller. However, there will be small deviations from the reference speed (1 p.u.) because only proportional control is used here. Small oscillatory transients are visible in the rotor speed responses shortly after increasing above 1 p.u. and are caused by an excited mode with low frequency that appears for activated pitch control. Fig. 4.8 (c) show the pitch angle

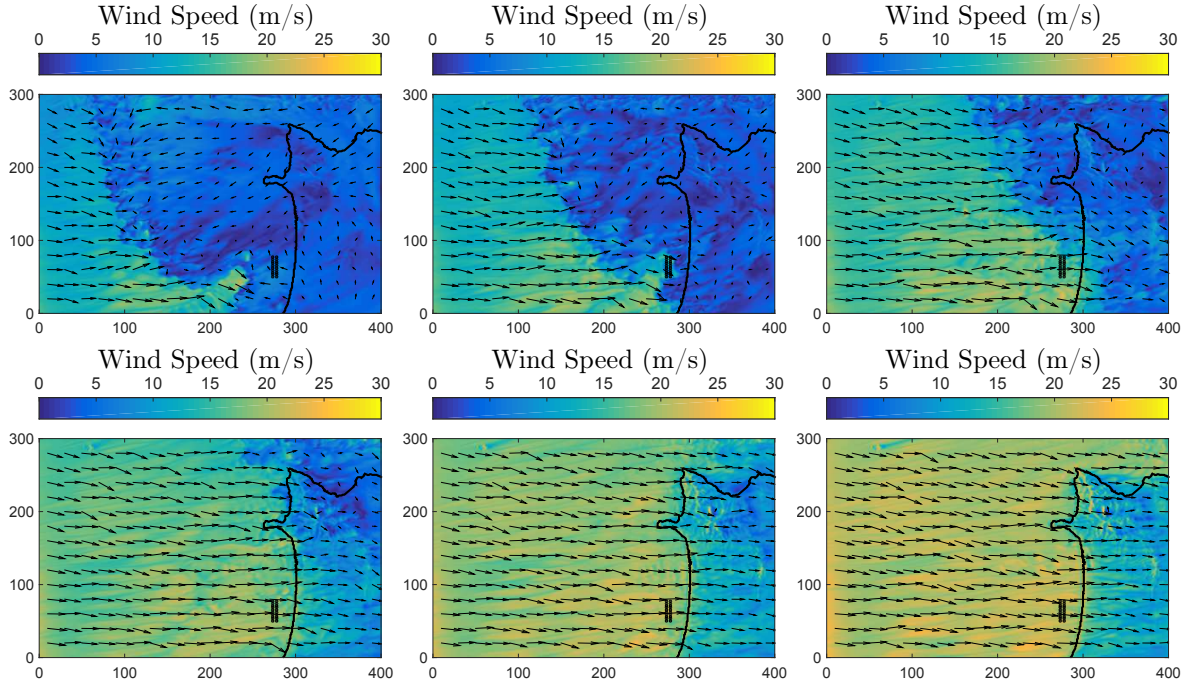


Figure 4.7: Wind profiles illustrating the weather scenario from which the wind speed measurements are acquired. The data are ordered in a 400×300 grid of measurements of an area outside the north-western coast of Japan. The coastline is depicted by the black solid curve, and the locations of turbines in the wind farm are given by the black dots. Arrows indicate the wind direction and magnitude.

responses which as expected are only larger than zero for super synchronous speeds. In Figs. 4.8 (d) and (e), active and reactive outputs of the converters are given. As expected and desired, active power outputs from WTs are limited and kept at 1 p.u. for super synchronous speeds.

4.5 Conclusions

This chapter described the fundamental modeling of wind turbines and wind farms, including basic controllers, and also outlined an initialization procedure starting from measurements of wind speeds. Simulations of a wind farm were explained, and run by incorporating wind speeds acquired from weather simulations. The contributions of this chapter is two-fold. First, a new initialization procedure of DAE-based simulation of a wind farm was proposed. Secondly, detailed weather predictions from CReSS were newly utilized in the wind power simulation and analysis. In the following chapters, we will refer to the modeling introduced and explained here.

Appendices

4.A The dq -transformation

The following explanation of the dq -transformation follows [7]. d stands for direct, and q for quadratic. The transformation is commonly used in the mathematical modeling of electrical machines since it eliminates time dependence of machine inductances obtained in the three-phase description. The description here is in terms of currents, but analogously applies to voltages as well. The so-called d - and q -axis currents from the three phase currents i_a , i_b , and i_c are given as follows:

$$\begin{aligned} i_d &= k_d \left[i_a \cos(\theta) + i_b \cos\left(\theta - \frac{2\pi}{3}\right) + i_c \cos\left(\theta + \frac{2\pi}{3}\right) \right], \\ i_q &= -k_q \left[i_a \sin(\theta) + i_b \sin\left(\theta - \frac{2\pi}{3}\right) + i_c \sin\left(\theta + \frac{2\pi}{3}\right) \right], \end{aligned} \quad (4.20)$$

where θ is a time-varying phase-angle, k_d and k_q two constants often set at $2/3$, and d and q represent two axes rotating 90 degrees apart. The three-phase currents, all with amplitude i_m , are defined as follows:

$$\begin{aligned} i_a &= i_m \sin(\omega_s t) \\ i_b &= i_m \sin\left(\omega_s t - \frac{2\pi}{3}\right), \\ i_c &= i_m \sin\left(\omega_s t + \frac{2\pi}{3}\right), \end{aligned} \quad (4.21)$$

i.e. oscillating with an angular frequency ω_s phase-shifted 120-degrees from each other, and t denotes the time. Combining (4.20) and (4.21) yields after some manipulation:

$$\begin{aligned} i_d &= k_d \frac{3}{2} i_m \sin(\omega_s t - \theta), \\ i_q &= -k_q \frac{3}{2} i_m \cos(\omega_s t - \theta). \end{aligned} \quad (4.22)$$

This transformation can conveniently be computed by a transformation matrix as follows:

$$\begin{bmatrix} i_d \\ i_q \\ i_0 \end{bmatrix} = \frac{2}{3} \begin{bmatrix} \cos(\theta) & \cos\left(\theta - \frac{2\pi}{3}\right) & \cos\left(\theta + \frac{2\pi}{3}\right) \\ -\sin(\theta) & \sin\left(\theta - \frac{2\pi}{3}\right) & \sin\left(\theta + \frac{2\pi}{3}\right) \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix}, \quad (4.23)$$

where i_0 is the zero-sequence current given by $i_0 = (i_a + i_b + i_c)/3$ and is equal to zero under so-called balanced conditions.

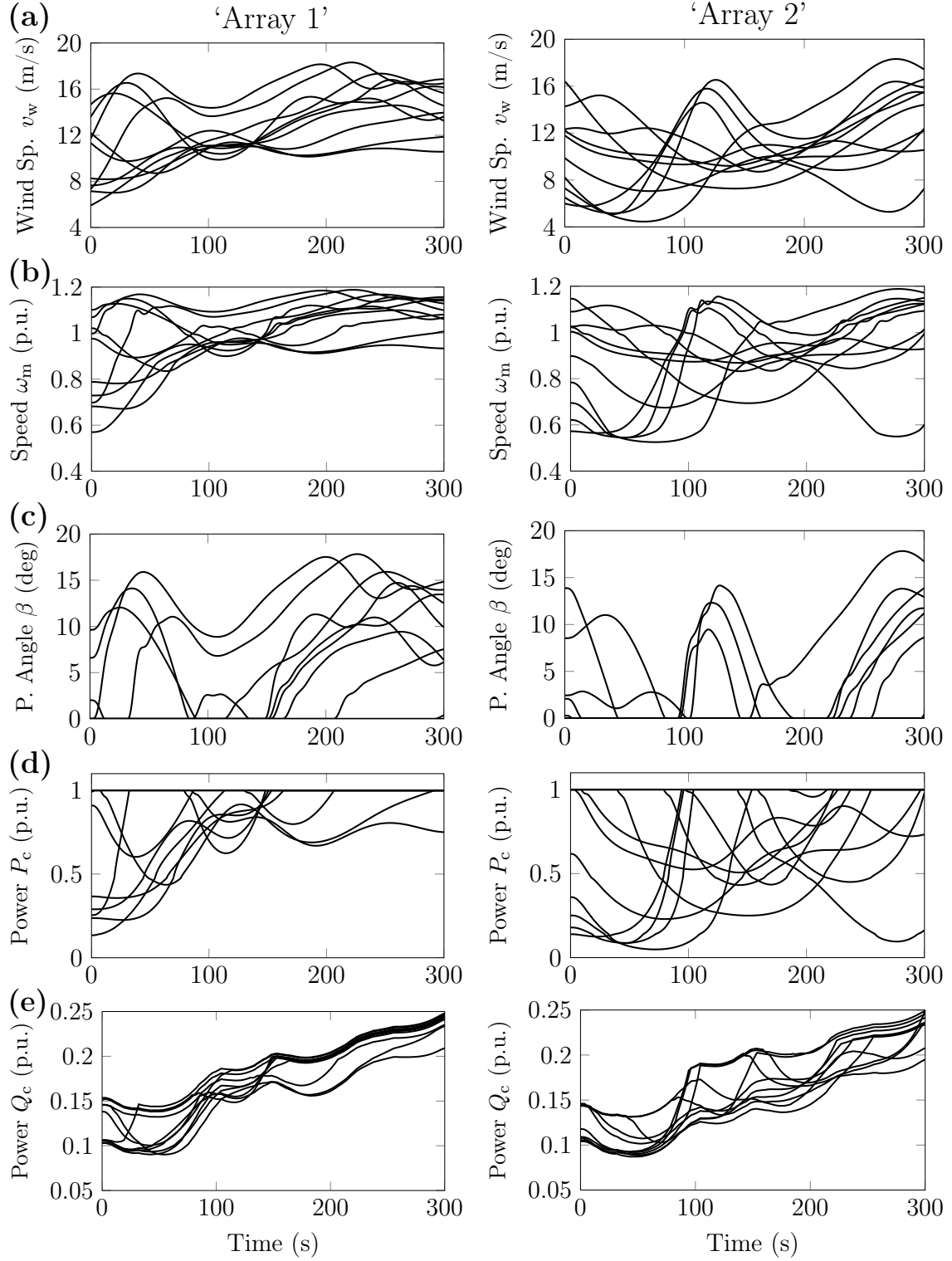


Figure 4.8: Simulation results for 'Array 1' (left side) and 'Array 2' (right side) of wind turbines. Fig. (a) show wind speeds for all measurement points, (b) give the turbines' angular speeds, and (c) show the pitch angle responses. The active power output of the turbines are given in (d) and reactive power in (e).

Chapter 5

Assessment of short-term offshore wind power fluctuations in Japan

5.1 Introduction

This chapter investigates short-term wind speed and wind power fluctuations with measurements from an offshore Wind Farm (WF) in Ibaraki Prefecture, Japan. Rapid wind power fluctuations become a challenge to power system operation and control as the penetration level increases [6, 40]. Currently, Japan does not have a significant amount of installed wind power; with 3 GW in 2015 it covered only about 0.5% of the total electricity demand [106]. However, wind energy resources around Japan are excellent with a potential annual energy production of 2700 TWh [25], where offshore wind accounts for as much as 82%. Japan seems to share this view by regarding floating turbines for offshore wind as a promising technology for the future [106]. Due to new energy policies following the Fukushima accident, we can expect an increase in installed wind power capacity in Japan in the forthcoming years [43]. On the other hand, Japan has recently achieved a significant penetration of solar power [39], currently providing about 10% of the electricity demand. In fact, Japan is third in the world in terms of installed solar power capacity, and has doubled its renewable energy capacity in the last three years [39]. It has been shown that both solar and wind power are highly intermittent and exhibit non-Gaussian characteristics [107]. In [107, 108] it was shown that these non-Gaussian characteristics remain in the aggregated power, even for vastly distributed generation. This is caused by long-range correlations, which were, e.g., demonstrated for wind speeds in [26]. These characteristics have potential implications for power system stability [6, 40, 107]. In order to cope with large-scale renewable generation, storage capability is indispensable [109]. However, it was argued in [109] that Germany can manage as much as a 50% penetration of wind and solar power without additional storage, and that additional storages would be needed when the combined solar and wind power generation exceeds about 80% of the demand.

5.1.1 Purpose and contributions

The short-term variability—here referring to the time-scale of seconds—of wind power is important to evaluate since wind speed fluctuations have been shown to be amplified by the nonlinear conversion to wind power and fed into the grid, despite the aggregation of multiple Wind Turbines (WT) [64, 107, 108]. Large short-term wind power fluctuations impact the operation and stability of power systems and are thus of vital importance to characterize [6, 40, 107]. To capture the realistic wind power fluctuations on the time-scale of seconds, a physical model including mechanical systems and controls is needed [64]. For example, such as the ones previously outlined in Chapter 4, described by Differential-Algebraic Equations (DAE), and with a scalar time-series of wind speeds as an input. On the other hand, the authors of [64, 110] proposed a model based on the so-called Langevin process, consisting of a deterministic drift towards a standard power curve, driven by Gaussian noise, formulated as a Stochastic Differential Equation (SDE). The stochastic part is motivated by the fact that the wind field is commonly reduced to scalar wind speeds v_w in models, thus ignoring the non-homogeneous wind field’s uneven impact on the blades [110].

The purpose of the work presented in this chapter is two-fold. First, this chapter aims to clarify short-term output characteristics of WTs and a WF (the aggregate output of seven WTs), for a WF located offshore in Japan. Their characteristics become important as the amount of wind power increases—in order to properly assess potential impacts on voltage and power quality [6]. Second, two types of WT models are constructed by utilizing measurements: a static power curve derived from an estimated curve of the turbine’s aerodynamic efficiency, and a more detailed dynamic model incorporating the data-derived WT-characteristics and a standard model of the electro-mechanical system, as explained in Chapter 4. With the detailed dynamic model, it is investigated to what extent observed WT-output fluctuations are captured by the model. If the fluctuations are of deterministic electromechanical origin, then controls to reduce fluctuations are possible.

The contributions of this study are as follows. First, clear non-Gaussian characteristics of WT and WF outputs are demonstrated. Then, the extent to which aggregated WT fluctuations are mitigated on a WF-scale it is demonstrated. It is shown that almost MW-scale variations are not uncommon in the WF output on the time-scale of seconds, despite the maximum distance between two WTs being 1.5 km. E.g., output changes of up to five standard deviations are not at all mitigated. This stands in sharp contrast to the notion of short-term wind speed and wind power fluctuations being uncorrelated and random, and stresses the importance to thoroughly assess the potential impact of short-term fluctuations of renewable generation in Japan. Second, by using a data-derived dynamic model, we quantify the amount of observed fluctuations captured by the model, using only a scalar time-series as input. In particular, a so-called R^2 -value [111] of about 0.9 is achieved, indicating an almost linear correlation between modeled and simulated output.

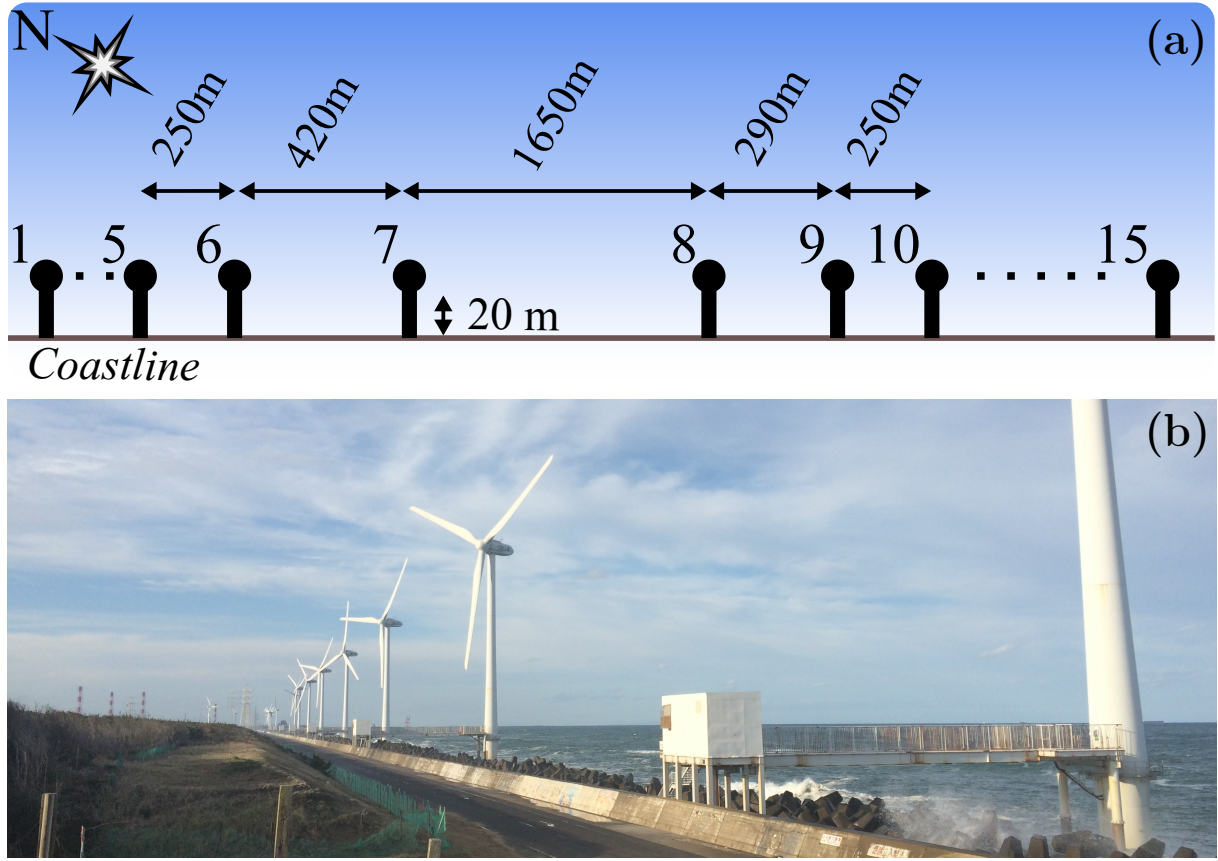


Figure 5.1: (a) Illustration of the placement of fifteen WTs along the coast of Ibaraki prefecture in Japan, as seen from above. The distances between WTs are generally 250 m except for three cases which are shown in the figure. Turbines are depicted by black dots and piers connecting them to land by thick vertical lines. Capital N in the upper left corner indicates the north. (b) Photo showing some of the WTs along the coast.

5.1.2 Outline

The remainder of this chapter is organized as follows. Section 5.2 describes the configuration of the target WF. Data from the same WF are also used in Chapters 6 and 8. Section 5.3 provides the mathematical methods used to characterize the short-term wind speed and wind power fluctuations. The results are presented and discussed in Section 5.4. Finally, conclusions are given in Section 5.5.

5.2 Description of target wind farm

The WF comprises fifteen 2-MW WTs close to shore and is depicted in Fig. 5.1 (a), as seen from above, where WTs are indicated by large black dots and the piers supporting them by thick vertical lines connecting the WTs to the coastline. The distance from

land to the WTs is about 20 m, i.e. they are in immediate proximity to shore. Figure 5.1 (b) clearly shows this. This configuration naturally results in significantly different turbulence-characteristics of the incoming wind depending on the wind direction, due to drastically different terrain characteristics. On land, buildings are sparsely scattered, including one of comparable height to the WTs. Wake-effects are expected to be most prevalent for approximately north-western and south-eastern wind directions. For this configuration, small changes in the wind direction can greatly affect the WF output, which makes it challenging to accurately predict the WF output without detailed knowledge about the site and accurate, highly-resolved weather predictions, and this is investigated in Chapter 6. This together with the fact that the WF is one of only a few offshore WFs in Japan makes it a highly interesting case study. All WTs have the same specification, and are so-called downwind turbines [66], for which the wind passes the nacelle before the blades, in contrast to the more common, so-called upwind turbines [66], where the situation is opposite. Thus, for wind blowing from sea towards land, the WTs' rotors are facing the shore.

5.3 Methods for statistical evaluation

In the statistical evaluation of short-term wind speed and power fluctuations, wind speeds or output powers $u(t)$ are decomposed as follows [64, 110, 112]:

$$u(t) = \bar{u}_T + u'(t), \quad (5.1)$$

where $\bar{u}_T = \langle u_T \rangle$ is the mean ($\langle \cdot \rangle$ denotes the mean operation) over some period T , and $u'(t)$ the fluctuations (deviations) from $\bar{u}$. According to the IEC standard [113], wind speed fluctuations are defined as deviations from the $T = 10$ -min averages $\bar{u}_T$, and the same definition for output power is used for the sake of consistency in the comparison. The variance is calculated as $\sigma^2 = \langle u'^2 \rangle$, and standard deviation as $\sigma = \sqrt{\langle u'^2 \rangle}$. The turbulence intensity [66] of wind is defined as

$$TI := \frac{\sigma}{\bar{u}_T}, \quad (5.2)$$

and it changes every 10 min since both the mean $\bar{u}_T$ and σ vary with time (slowly in comparison to $u'(t)$). With the above definitions, we can rewrite (5.1) as

$$u(t) = \bar{u}_T (1 + TI \cdot z'(t)), \quad (5.3)$$

where $z'(t) = u'(t)/\sigma$ are the fluctuations normalized by σ . This chapter will look at Probability Density Functions (PDF) of fluctuations $p(z'(t))$, where p denotes the PDF, to qualitatively compare statistics of wind speed, and measured and modeled power outputs. Generally, the PDF of a random variable X is denoted by p_X , and the associated so-called Cumulative Density Functions (CDF) is defined by $F_X(x) = \int_{-\infty}^x p_X(t)dt$. The probability $P(X \leq x)$, i.e. the probability that X is less than a certain value x , corresponds to

$F_X(x)$. PDFs p_X are estimated with the function `ksdensity` in MATLAB, and the CDFs are calculated by integrating p_X with the trapezoidal rule.

We also consider so-called increments of $u(t)$ [64] which are simply defined as

$$du(t, \tau) = u(t + \tau) - u(t), \quad (5.4)$$

where τ is a time-shift in seconds, which lets us look at wind speed and wind power variability at different time-scales determined by τ . Wind speed increments are related to wind gustiness which affects load on WT structures [114]. Wind power increments are closely related to *wind power ramps* which are often analyzed on hourly time-scales, and important to predict to ensure safe power system operation [115]. Such an analysis is conducted in Chapter 6.

As measures for comparison and *normality* of fluctuations $u'(t)$, we look at skewness Sk and flatness F (or kurtosis), here estimated for the zero-mean fluctuations $u'(t)$ as

$$Sk(u') = \frac{\langle u'^3 \rangle}{\sigma^3}, \quad F(u') = \frac{\langle u'^4 \rangle}{\sigma^4}.$$

These are the third and fourth order normalized moments $\mu_n = \langle u'^n \rangle / \sigma^n$ (n determining the order). Their theoretical values for normal distributions are: $Sk = 0$, and $F = 3$. The skewness quantifies asymmetry in the distribution: hence a symmetric normal distribution has a skewness of zero. The flatness quantifies the “thickness” of the tails of the PDF: a higher value indicates higher probability of more extreme values [116]. Now, assuming stationarity within 10-min intervals, (5.4) can be approximated as

$$du(t, \tau) = u(t + \tau) - u(t) \approx u'(t + \tau) - u'(t), \quad (5.5)$$

which is a very accurate approximation for our wind speed data. This is strictly true if $\bar{u}_T$ is the same for consecutive 10-min intervals. It is reasonable to assume that $\bar{u}_T$ generally does not change rapidly on such time-scales. Now, assuming $\langle u'(t)^2 \rangle \approx \langle u'(t + \tau)^2 \rangle$, which holds for long time-series, following [112], the second moment of increments is given by

$$\mu_2(du(t, \tau)) = \langle (u'(t + \tau) - u'(t))^2 \rangle = \langle 2u'(t)^2 - 2u'(t + \tau)u'(t) \rangle. \quad (5.6)$$

Note that the sample autocorrelation of $u'(t)$ is computed by

$$R_{u', u'}(\tau) = \frac{1}{\sigma^2} \langle u'(t + \tau)u'(t) \rangle = \langle z'(t + \tau)z'(t) \rangle. \quad (5.7)$$

Thus, with (5.7), we can rewrite (5.6) as follows:

$$\mu_2(du(t, \tau)) = 2\sigma^2(1 - R_{u', u'}(\tau)). \quad (5.8)$$

Hence, we see that the second moment of increments depends on the autocorrelation. Similarly, higher order moments $\langle u'^n \rangle$ depend on higher order correlations $R_{u'^n, u'^m}(\tau)$ [112]. These higher order correlations are obviously manifested in the increment PDF since they

affect its shape. To quantify deviation from Gaussianity, [112] derived an explicit formula for $p(du(\tau))$, where

$$\lambda^2(\tau) = \frac{\ln(F(du(\tau))/3)}{4}, \quad (5.9)$$

is identified as a parameter determining the shape of the distribution. We see from the definition that the shape factor λ^2 for normally distributed (increment) data would result in a value of zero. In this chapter $\lambda^2(\tau)$ is utilized to quantitatively compare shapes of PDFs of power increments from different WT models.

To compare how close the outputs of different WT model are to the measured output P'_w , the so-called R^2 measure [111] is used:

$$R^2 = 1 - \frac{\sum_{i=1}^N (P'_{w,i} - P_{w,\text{mod},i})^2}{\sum_{i=1}^N (P'_{w,i} - \langle P'_w \rangle)^2}, \quad (5.10)$$

where $P_{w,\text{mod}}$ is the output of a WT-model. R^2 becomes 1 for perfect agreement between measurements and model outputs, and $R^2 = 0$ for $P_{w,\text{mod}} = \langle P'_w \rangle$. Thus, a model for which R^2 is larger than zero is better than the “mean-model” $P_{w,\text{mod}} = \langle P'_w \rangle$.

5.4 Statistics of short-term wind speed and power

Here, we first derive WT models and then look at the statistics of short-term wind speed and output power fluctuations at the WF depicted in Fig. 5.1.

5.4.1 Deriving models for comparison

A model of the WT can be estimated from measurements. To do so, from (4.3), we have

$$c_p(\lambda) = c_1 \left(c_2 \left(\frac{1}{\lambda} - c_9 \right) - c_6 \right) \exp \left(-c_7 \left(\frac{1}{\lambda} - c_9 \right) \right), \quad (5.11)$$

where c_i are constants to be determined from the measured data, and β in (4.3) has been disregarded here since it can be assumed zero below rated power for which situation c_p is modeled here. c_p can then be estimated from data as follows:

$$c_p = \frac{P'_w}{\frac{1}{2} \rho A_r v_w^3}. \quad (5.12)$$

The c_i parameters of c_p are estimated with the `lsqcurvefit` function (nonlinear curve-fitting optimal in a least-square sense) in MATLAB by solving $c_p = c_p(\lambda)$ (using (5.11) and (5.12)). Calculated values of c_p plotted against λ are shown in Fig. 5.2(a) together with the estimated $c_p(\lambda)$ ‘Mod. Curve’ and a curve of the averaged data ‘Avg. Curve’.

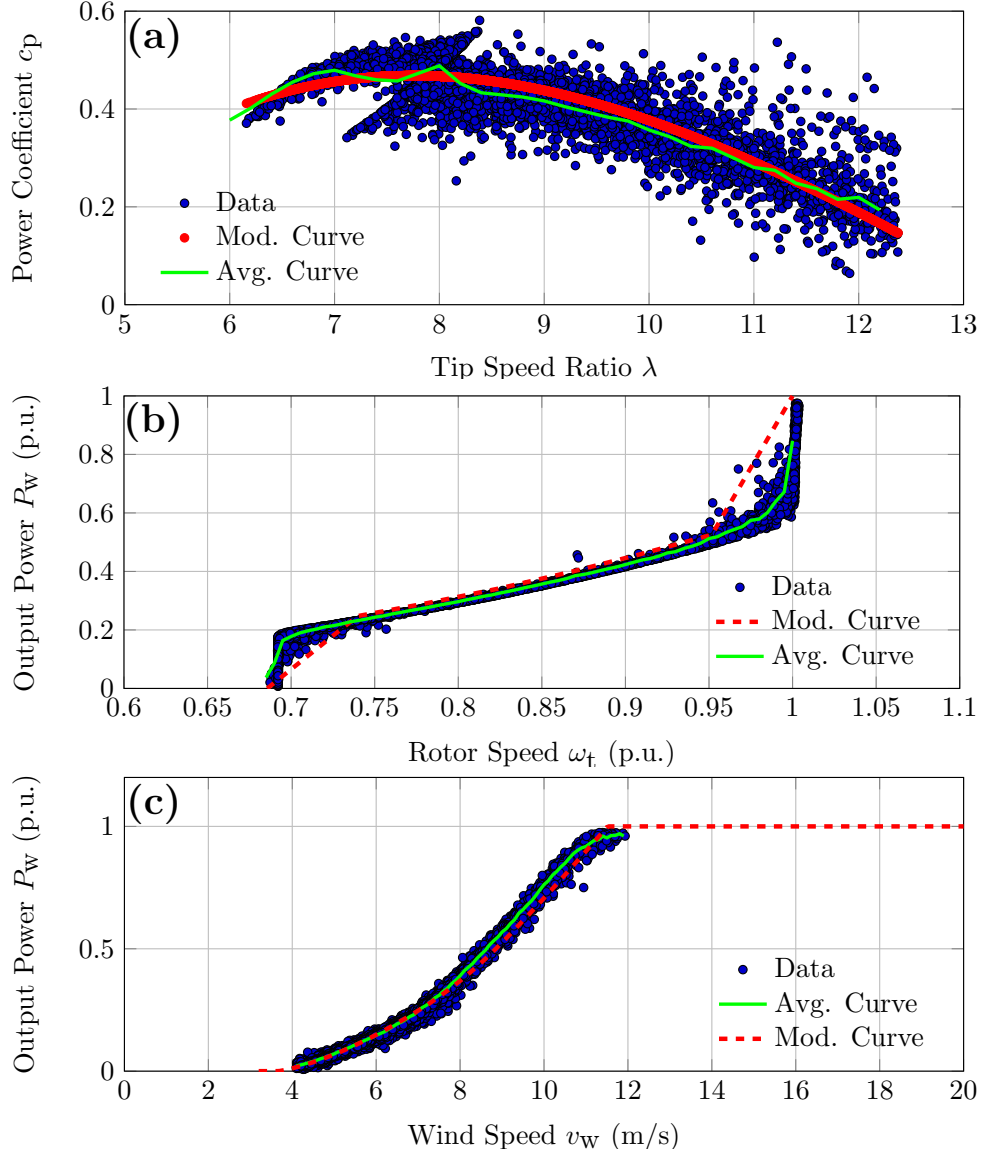


Figure 5.2: (a) Power coefficient c_p . The red curve depicts the derived c_p from a curve-fit using (5.11) and (5.12). The green ‘Avg. Curve’ shows data averaged for small intervals of tip speed ratio λ . (b) Output power vs. rotor speed for the measurements, and optimal output power vs. rotor speed for the model. (c) Measured data, and averaged and modeled power curves: ‘Avg. Curve’ and ‘Mod. Curve’. The modeled curve is derived using (4.5).

With the estimated $c_p(\lambda)$, an optimal relation between P_w and ω_t , $P_w^*(\omega_t)$ is derived by tracking $\max(P_w)$ for different v_w and ω_t (Maximum Power Point Tracking (MPPT)). $P_w^*(\omega_t)$ is used to control the WT at optimal power, as described in Chapter 4. Measured, averaged, and modeled P_w vs. ω_t are shown in Fig. 5.2 (b) in p.u. Note that we let $P_w^*(\omega_t)$ deviate from MPPT for low and high ω_t to keep an unique relation between P_w and ω_t (otherwise, a vertical slope is obtained for $P_w \gtrsim 0.6$ p.u. as observed in Fig. 5.2 (b)).

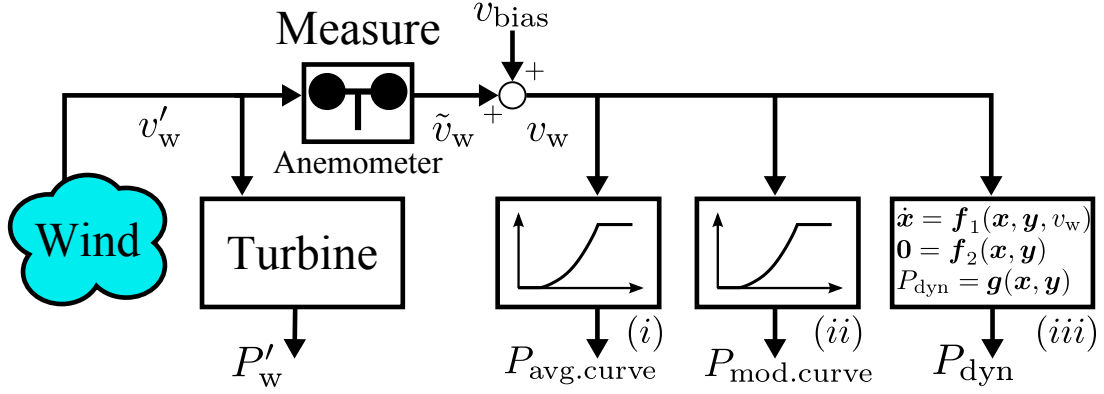


Figure 5.3: Outline of measurements and models. The measurement data used in this chapter are wind speeds $\tilde{v}_w$ measured with anemometers at the nacelles of WTs and output powers P'_w . The wind speeds $v_w(t) = \tilde{v}_w(t) + v_{bias}$ are used as input to three WT models.

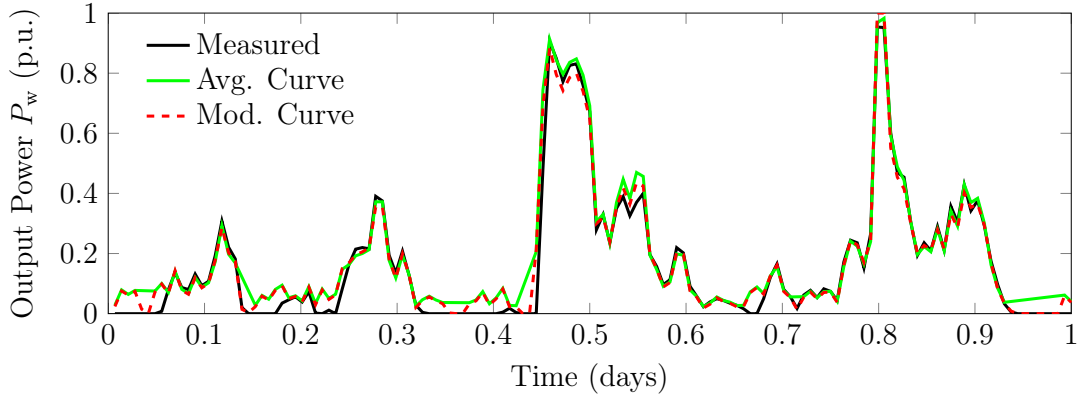


Figure 5.4: Comparison between measured WT outputs and outputs using the averaged and modeled power curves shown in Fig. 5.2 (c).

It is now possible to solve (4.5) for a number of wind speeds to derive a power curve ‘Mod. Curve’: see Fig. 5.2 (c). The outputs with the average power curve ‘Avg. Curve’, and ‘Mod. Curve’ will be evaluated together with a DAE model incorporating the derived c_p . For the DAE model, the modeling except for c_p , and known physical dimensions/parameters of the real WT, is according to the DFIG (type used in the WF), outlined in Chapter 4, with standard parameters from [92, 97, 98]. In summary, three models are derived to model the power conversion of the WT:

- (i) Averaged power curve: green ‘Avg. Curve’ in Fig. 5.2.
- (ii) Power curve derived using estimated c_p : dashed red ‘Mod. Curve’ in Fig. 5.2.
- (iii) Dynamic model using (4.10) with the two-mass mechanical model (4.6) instead of the one-mass model, and with the mechanical power P_w (see (4.2)) incorporating the modeled c_p .

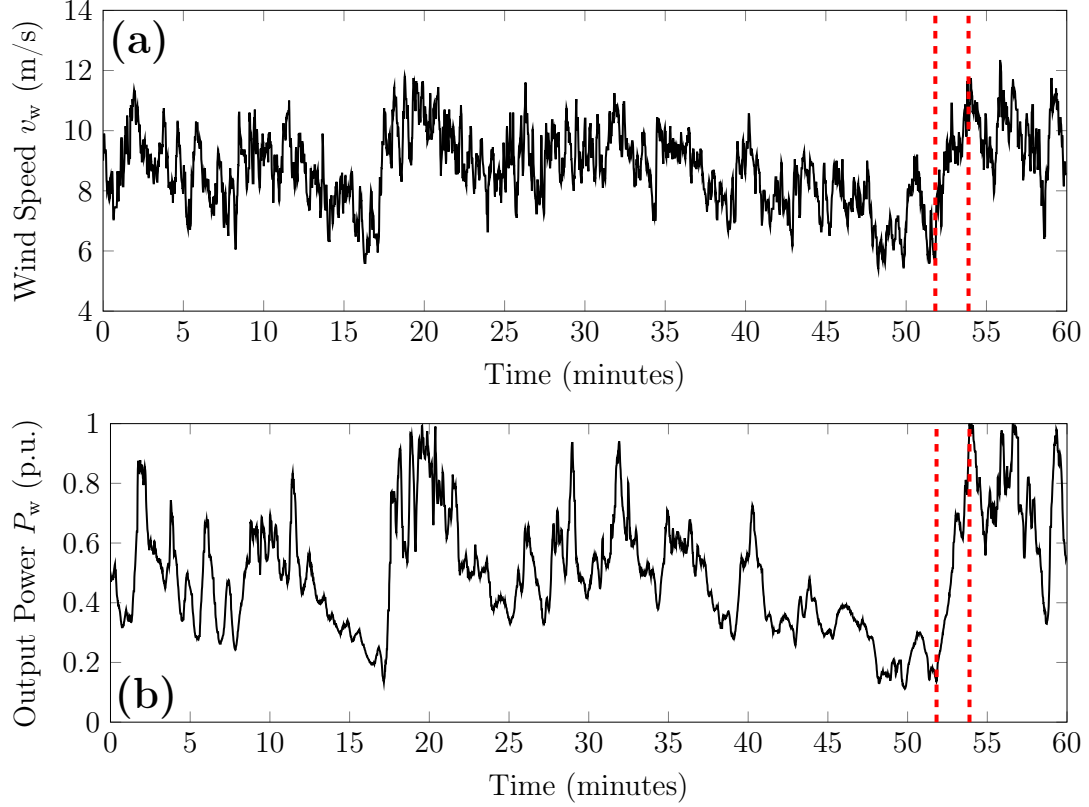


Figure 5.5: One hour example time-series of measured (a) wind speeds and (b) output powers. The interval marked in red is used in Fig. 5.6.

See Fig. 5.3 for an illustration of available measurements, approximations, and models. Note that a wind speed bias v_{bias} is added to the measurements before using them as input to the models. This bias represents an observed influence of the WT rotor on the measured wind speeds. For simplicity, we let $v_{\text{bias}} = 1.1$ m/s, although it is obviously an oversimplification. In Chapter 6, a linear function $\tilde{v}_w = f(v'_w)$ is used instead of v_{bias} .

Figure 5.4 shows the 10 min-averaged WT outputs for a duration of one day, together with outputs from the modeled and averaged power curves. It clearly shows that a power curve $P_w(v_w)$ estimates the measured P'_w well in the long term, although differences are seen particularly for small and large outputs. In the following section, the insufficiency of using a simple power curve to estimate the short-term output is demonstrated, and a statistical evaluation of WT outputs is conducted.

5.4.2 Statistical evaluation and comparison

Here, we will first investigate and compare statistics of short-term wind speed and wind power fluctuations. After that, WT and WF power variability in terms of increments are compared. Finally, we quantify the effectiveness of the data-derived dynamic WT model

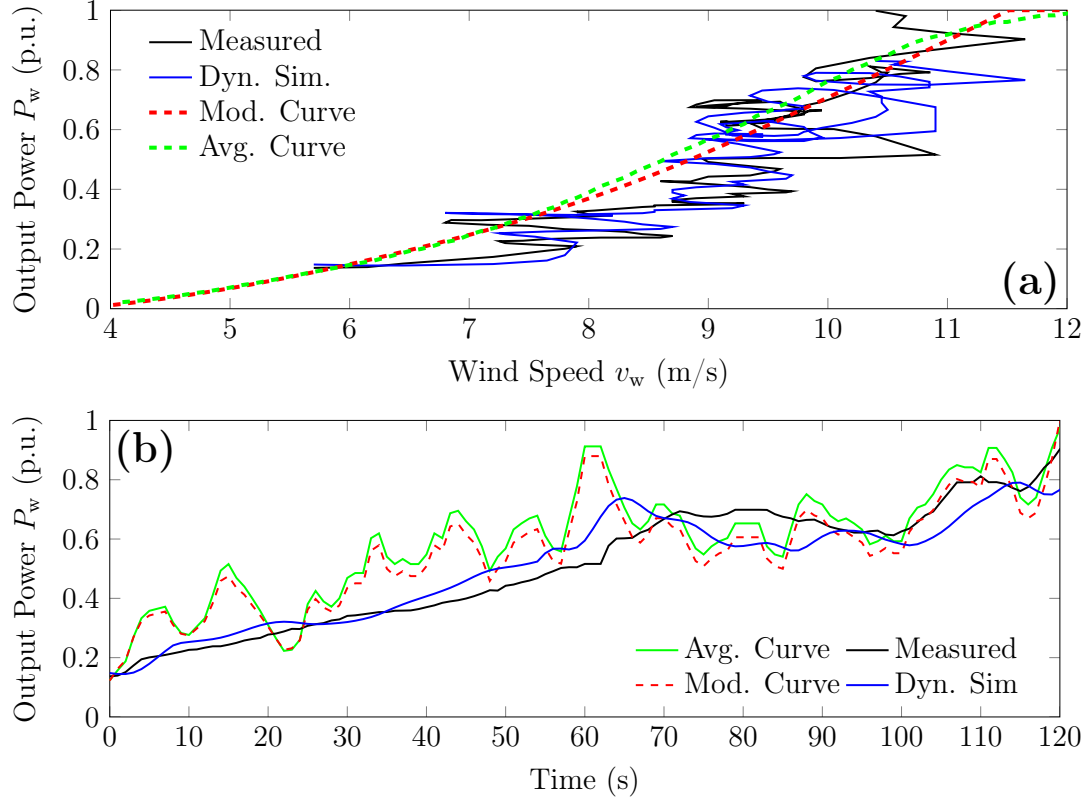


Figure 5.6: Measured and simulated power output plotted against (a) wind speed and (b) time, for the interval indicated in red in Fig. 5.5.

in capturing realistic output fluctuations.

Examples of 1-Hz sampled wind speeds $\tilde{v}_w$ and P'_w are shown in Fig. 5.5. The power output looks more “smooth” than $\tilde{v}_w$, which is due to filtering in the electromechanical conversion process, and because the WT-blades act as a low-pass filter for the wind. In Fig. 5.6 (a), about 2-min long trajectories of measured and simulated (using the dynamic model of a DFIG (4.10)) P_w plotted against v_w are shown. It is observed that $P_w(v_w)$ approximates the long-term, averaged output P_w well, which was demonstrated for the 10-min data. However, it cannot capture the behavior well in the short-term, since both measurements and data from dynamic simulation ‘Dyn. Sim.’ exhibit large deviations from the power curve. This is due to controls and inertia of the WT [64, 110], which the proposed dynamic model includes. For example, examining the time-series in Fig. 5.6 (b), we see that the measured WT output and the output from the dynamic simulation fluctuate in a similar manner, qualitatively different from the outputs estimated with the power curves ‘Mod./Avg. Curve’. Therefore, we understand that the dynamic model is necessary for evaluation of short-term power fluctuations. This behavior has also been described as a Langevin process in [64, 110] consisting of a deterministic drift towards the power curve plus random fluctuations.

Table 5.1: Flatness and skewness of normalized fluctuations z' and increments $du(t, \tau)/\sigma$ (with $\tau = 4$ s), R^2 -values between measured and modeled outputs, and RMSE. Note that v_w : measured wind speed; $p_{w,\text{meas}}$: measured WT output; $p_{w,\text{avg}}$: WT output with averaged power curve; $p_{w,\text{mod}}$: WT output with modeled power curve; $p_{w,\text{dyn}}$: WT output with dynamic simulation utilizing a DAE-model.

	v'_w	$p'_{w,\text{meas}}$	$p'_{w,\text{avg}}$	$p'_{w,\text{mod}}$	$p'_{w,\text{dyn}}$
F	3.54	3.64	3.84	3.75	3.13
Sk	0.01	0.27	0.59	0.58	0.37
	dv_w	$dp_{w,\text{meas}}$	$dp_{w,\text{avg}}$	$dp_{w,\text{mod}}$	$dp_{w,\text{dyn}}$
F	8.70	40.7	16.4	17.1	23
Sk	0.45	0.86	0.42	0.43	0.73
	$P_{\text{avg.curve}}$		$P_{\text{mod.curve}}$		P_{dyn}
R^2	0.85		0.86		0.90
RMSE	0.071		0.065		0.054

To quantitatively compare the measurements with simulations, we will now look at statistical properties of wind speed and wind power fluctuations: v'_w and p'_w . The data consist of WT measurements acquired over a period of 3.6 days. For the same period, we also look at the aggregated and normalized WF output of WTs 9–15: see Fig. 5.1 (a). Normalized PDFs of v'_w ‘Wind Speed’ and p'_w ‘WT Output’ are shown in Fig. 5.7 (a). Keep in mind that fluctuations are defined as deviations from the 10-min average. It is noted that the wind speed, WT output, simulated outputs, and the WF output ‘WF Output’ all become close to the Gaussian PDF for fluctuations smaller than about four standard deviations (4σ) in magnitude. The simulated outputs consist of outputs from dynamic simulation of the DAE model: ‘Dyn. Sim.’, and the two power curves shown in Fig. 5.2 (c). These results indicate that the model (5.1) with $T = 10$ min corresponds well with u' represented by Gaussian noise, and that the assumption of stationarity to some extent holds in each 10-min period. The flatness F for all cases is given in Table 5.1, and the values are overall fairly close to Gaussian. Values on Sk given in the table also become close to Gaussian. In particular, Sk for wind speeds is very close to zero and indicates a symmetric distribution.

Now, PDFs of increments $du(t, \tau)$ are considered, and given in Fig. 5.7 (b) in arbitrary unit (a.u.) for a time-shift $\tau = 4$ s. The PDFs are arbitrarily shifted in the y -direction for the sake of easy visual comparison. Compared to Fig. 5.7 (a), all PDFs show long tails, indicating non-Gaussian characteristics. In particular, we see that the PDF of WT output increments exhibits the longest tails followed by dynamic simulation and total WF output. The non-Gaussian PDF characteristics are influenced by non-zero autocorrelations of various orders. Evidently, these characteristics are most prevalent in the power outputs due to the nonlinear electro-mechanical conversion of wind to electric power. Fast wind speed fluctuations of a few m/s can cause large variations in the WT output, amplifying the intermittency of the wind: see e.g. Fig. 5.5. We can also understand that low-

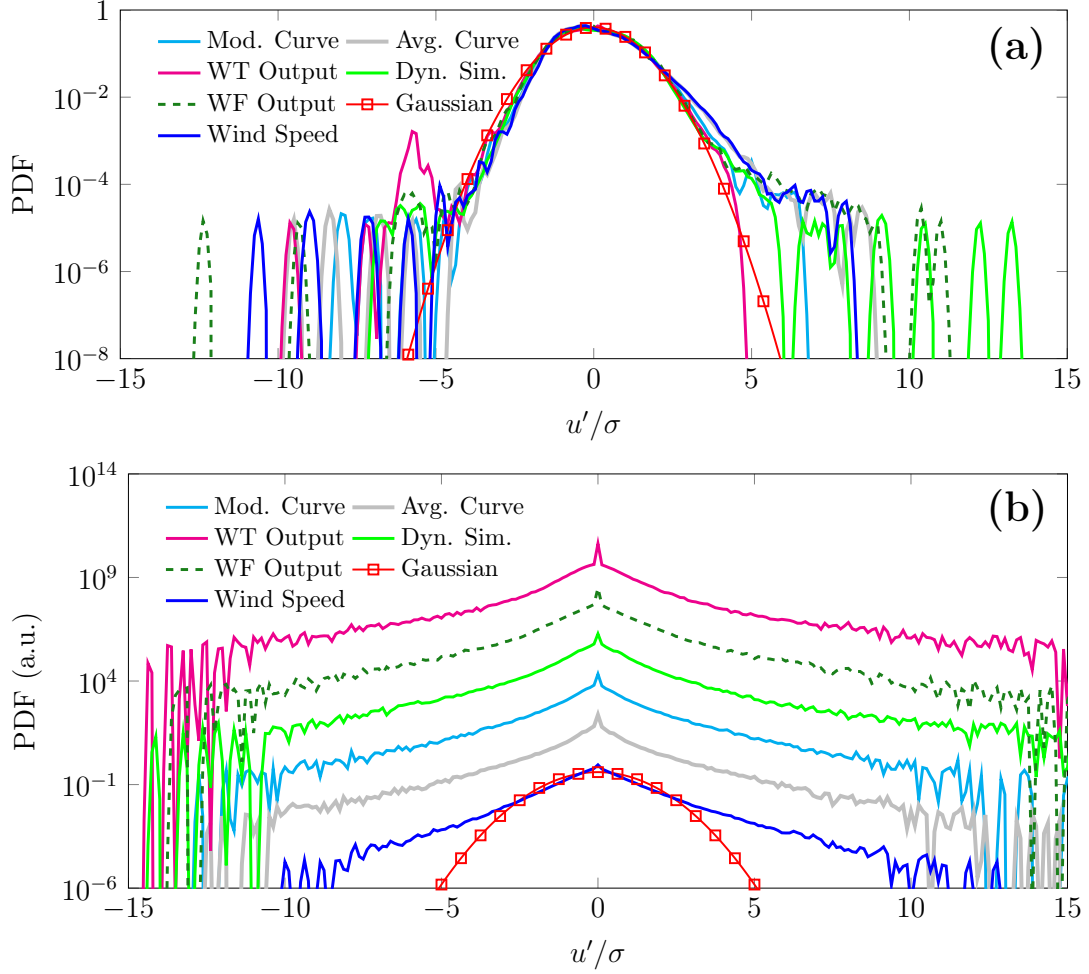


Figure 5.7: Probability Density Functions (PDF) of (a) normalized fluctuations u'/σ ; (b) increments $du(t, \tau)/\sigma$ ($\tau = 4$ s). The Gaussian PDFs have standard deviation $\sigma = 1$ (same as the standard deviations of u'/σ and $du(t, \tau)/\sigma$). The PDFs are arbitrarily shifted in the y -direction in (b) for the sake of easy comparison.

pass-filtering-effects in the power conversion due to WT inertia and blades increase the autocorrelation, consequently affecting the increment PDFs. The values on flatness F of the increments are given in Table 5.1. It is observed that F of power increments dp clearly differ from fluctuations p' , and become particularly non-Gaussian for the measured power $dp_{w, \text{meas}}$, and $dp_{w, \text{dyn}}$ from dynamic simulation. F also becomes larger for $dp_{w, \text{avg}}$, $dp_{w, \text{mod}}$, and dv_w compared with p' and v'_w , but not as large. Looking at Sk , it can be seen that $dp_{w, \text{meas}}$ and $dp_{w, \text{dyn}}$ seem to exhibit the most asymmetrical distributions.

The shape parameters λ^2 of $du(t, \tau)$ are plotted in Fig. 5.8 for the same cases as in the previous paragraphs. It is observed that the results for the power curves become almost identical, and λ^2 for the measured WT output becomes largest (except for $\tau = 1$ s). The result for the dynamic simulation is closer to the measured power than the power

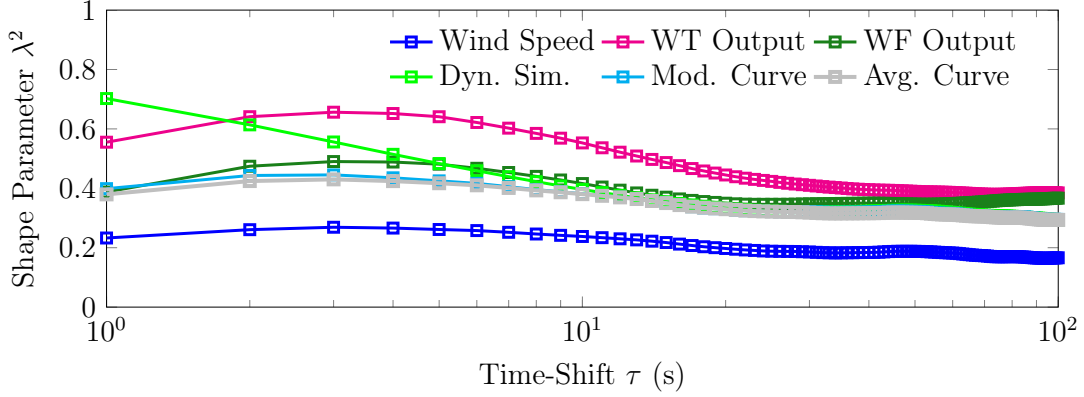


Figure 5.8: Shape parameter λ^2 vs. time-shift τ .

curves, in particular for $\tau = 2$ s, where the results are almost identical. Eventually, λ^2 of the increments of the dynamic simulation output converge to λ^2 of the power curve outputs at $\tau \approx 20$ s. This is an expected result since the modeled power curve and the dynamical model are based on the same data-derived power coefficient c_p , and the dynamic power conversion mainly affects the output characteristics on time-scales of tens of seconds [64, 117]. This is also demonstrated in Fig. 6.3(b) in the next chapter by a spectral analysis. Another interesting takeaway from Fig. 5.8 is that ‘Dyn. Sim.’ and ‘WF Output’ become close at around $\tau = 4$ s, which indicates that their output characteristics on the time-scale of several seconds are close. This is discussed further in the next paragraph, which looks at the probability of output increments.

Now, WT and WF increments are compared by looking at the probability of a certain increment x given by $P(du/\sigma > x) = 1 - F(x)$, where $F(x)$ denotes the CDF of increments. Figure 5.9 shows $P(du/\sigma > x)$ for the measured and simulated WT outputs, and for WF outputs for $\tau = 1, 4$ and 8 s. These results clearly show that the probability of sudden changes in the outputs on the time-scale of seconds are very close for the WF and WT up to about 5–6 standard deviations. As an example, here, an increment of 7 standard deviations corresponds to a change in WF output of about 0.5 MW: quite a large variation. Such variations are potentially challenging to power system operation, as was discussed earlier. Interestingly, the results for ‘Dyn Sim.’ appear to be in between ‘WT Output’ and ‘WF Output’ for $\tau = 1$ and 4 s, but for $\tau = 8$ we see an extremely close resemblance between the output increment probabilities for ‘Dyn Sim.’ and ‘WF Output’. This result suggests that a part of what is not captured by the dynamic simulation may be caused by either wind fluctuations on the time-scale of seconds not captured by the scalar input to the model, or by the WT model itself. Further, that these fluctuations are mitigated in the aggregated power, and consequently close results between the dynamic simulation and WF output are obtained.

Now, to further evaluate the different WT-models, we compute R^2 values according to (5.10), and the results are given in Table 5.1. The results strengthen the conclusion that the dynamic model outperforms the power curves (which represent the “long-term”

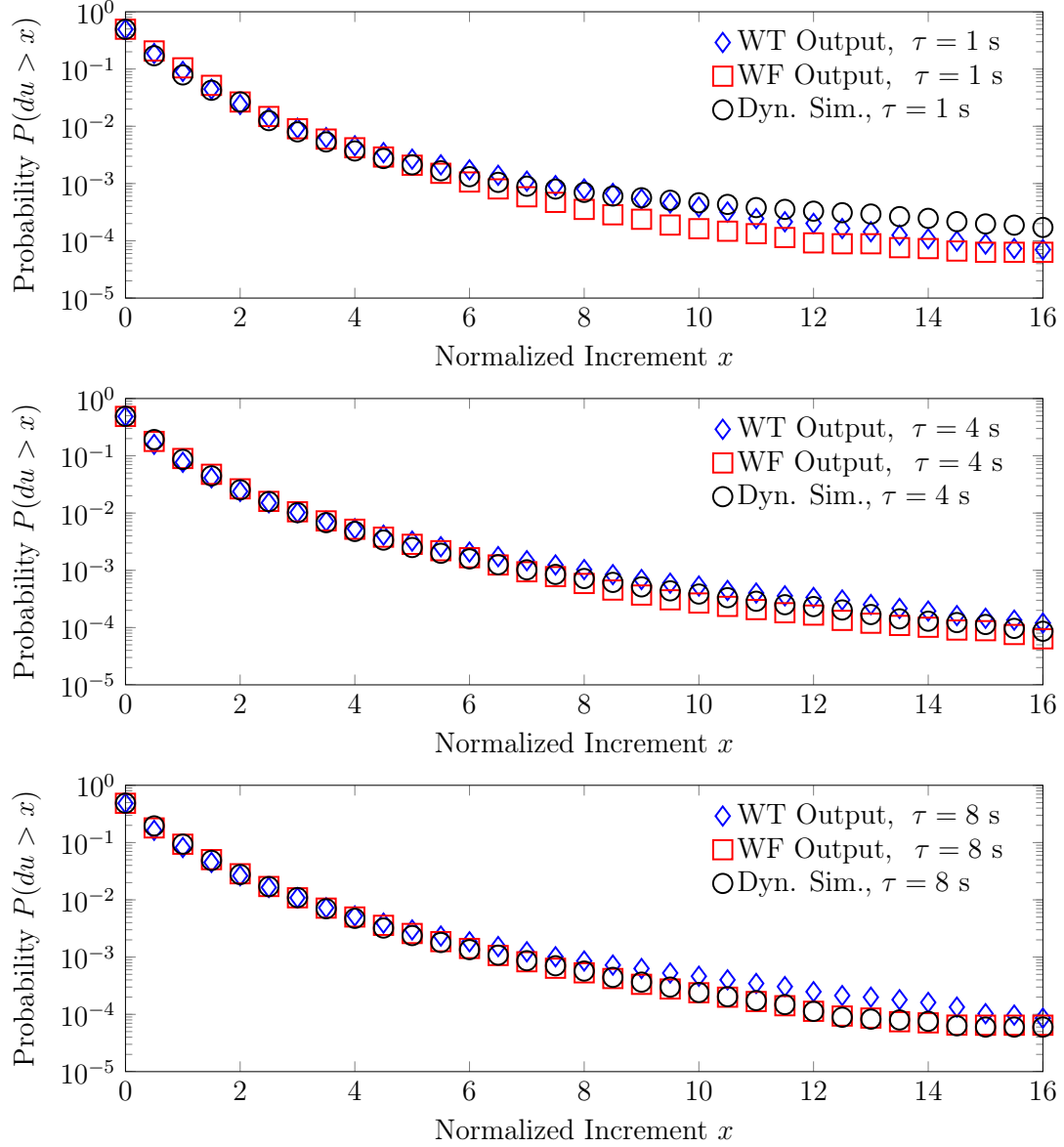


Figure 5.9: Probabilities $P(du > x)$ of WT and WF normalized output increments $du(t, \tau)/\sigma$ larger than x .

average). Additionally, Root Mean Square of Errors (RMSE) between measured and modeled outputs for the full 3.6-day period are given in Table 5.1. RMSE becomes smallest for P_{dyn} . This is due to the fact that the dynamic model captures more realistic fluctuations because of its dynamic mechanical and control system, as was exemplified e.g. in Fig. 5.6, and which is evident from Fig. 5.8.

5.5 Conclusions

Short-term wind speed and wind power characteristics were statistically evaluated at an offshore wind farm located in Ibaraki Prefecture, Japan. Three models of a wind turbine's power conversion were derived from wind turbine measurements. Time-domain simulation and statistical analysis were conducted, and indicated that the power fluctuations were better represented by the more detailed differential algebraic model than with standard power curves. The results indicated that a large part of the fluctuations around the power curve are due to the deterministic electromechanical conversion or control system and are not solely described stochastically. Thus, we suggest here that developing a representative short-term model of the WT is essential for understanding and controlling fluctuations in the short-term power conversion. Furthermore, we showed that aggregated WT output fluctuations were not mitigated on a second scale for variations less than about 5–6 standard deviations. Since the balance of active power is closely linked to the system frequency: see e.g. (1.3), this implies that large WF output variations are fed into the power system which could possibly cause both frequency and voltage fluctuations. This relates to so-called wind power smoothing, where the reduction in variability in aggregated power is commonly known as the smoothing effect. Related to this, in Chapter 7, we investigate smoothing effects of large-scale wind power in Japan.

Chapter 6

Assessment of offshore wind power characteristics in Japan

6.1 Introduction

In this chapter, wind conditions and output power characteristics of a Wind Farm (WF) in Japan are evaluated with highly-resolved weather predictions from the so-called Cloud Resolving Storm Simulator (CReSS) [63]. In difference to Chapter 5, which focused on statistics of short-term wind speed and Wind Power (WP) fluctuations, this chapter focuses on evaluating wind and WP predictions from CReSS. Later on, in Chapter 7, the data are used to study smoothing effects of aggregated wind power in Japan. The CReSS data used are mainly hourly predictions, but also data from extremely detailed 200-m resolution simulations are considered. The WF is the same one described in Chapter 5 (Section 5.2). Predicted wind speeds, directions, and powers are compared with measurements from turbines in the WF. This investigation is motivated by the rapid growth of WP around the world, which makes it the fastest growing source of renewable energy [37, 38], and the potential and expected growth of offshore WP in Japan [25, 43].

The wind conditions around a potential WF site need to be carefully assessed before its construction stage and are in practice evaluated with at least a year’s worth of data from measurement towers (called met masts). Alternatively, so-called Doppler lidars have emerged as a promising technology for replacing traditional met masts [118, 119]. A suitable WF-configuration can then be decided with the aid of commercially available softwares that enable analysis of the wind flow fields in the WF under different conditions [66]. The potential usages of CReSS are manifold in this context: first, CReSS provides users with wind speeds at desired heights, which becomes even more important as the WTs grow larger and the cost of the met masts increases. Second, instead of a single measurement, a large number of measurement points are available and can be decided by adjusting the spatial resolution. Thus, coverage of a large area can be achieved. In other words, CReSS could be used as a tool for preliminary evaluation of a potential site before deciding to perform in-situ measurements—thus potentially saving resources. Assessment

of WF sites becomes particularly expensive for offshore locations [120]. Further, CReSS can provide initial conditions to CFD models that simulate detailed wind conditions and WF wakes. Such coupling between mesoscale and microscale simulation is investigated in e.g. [121–123]. Finally, CReSS can be exploited purely for Numerical Weather Predictions (NWP) to provide power system operators with detailed wind and WP predictions.

6.1.1 Literature review

A great deal of work has been done in the field of wind and WP prediction, including WP forecasting using NWPs (physical weather models). Many such studies are discussed in [38, 124, 125]. A particularly challenging task for NWP-models is prediction of near-surface wind due the complex flows in the lower atmosphere and terrain effects [126]. In [126], NWPs were conducted with the Weather Research and Forecasting (WRF) model with boundary and initial conditions from different reanalyses (results of data assimilation schemes incorporating weather models and observations). It was shown that the choice of boundary and initial conditions had large influences on the prediction error. The authors of [127] also used the WRF model, and their results showed differences between different parameterizations, particularly for high wind speeds. The paper [128] also discussed the importance of deliberately selecting appropriate physical parameterization schemes in the WRF model. In [129], averaged wind speeds from nacelles of WTs were evaluated against met mast data, and it was concluded that averaged nacelle-measurements of WTs better represented the actual WF conditions. It was also found that averaged nacelle wind speeds were in better agreement with NWPs. In [130], a hybrid wind power forecasting system incorporating the WRF model and Kalman filtering was proposed and evaluated against wind speeds measured at the nacelles of WTs on Awaji Island, Japan. A substantial reduction in prediction error was demonstrated by using Kalman filtering in comparison to raw predictions. In conclusion, we can understand that the performance of an NWP with a particular set of physical schemes depends on many factors including the meteorological conditions [128, 131], and in particular the atmospheric stability [127], and hence that a quantitative comparison between NWP models can only be made for the same location with identical boundary and initial conditions.

NWPs have also been used to forecast WP ramp events [132]—i.e. sudden variation in WP over a short period of time—and it has been argued that NWPs are not suitable for ramp forecasting in the less than 6-hour scale [132, 133]. It was found that the ramp timing might be inaccurate with NWPs [134], but that a combination of different NWPs would improve the accuracy. The authors of [115] demonstrated how an ensemble of ramp forecasts could be utilized to accurately predict the WP ramps.

6.1.2 Purpose and contributions

The purpose of this investigation is two-fold. First, we experimentally investigate CReSS as a tool for detailed analysis and evaluation of WP and potential WF sites. This is

Table 6.1: Summary of data sets and their usages in the current comparison.

#	Type	Time-resolution	Spatial resolution	Acquisition period	Usage
1	WT-data	10 min	15 WTs	July 2014–July 2015	Wind/power comparison. Power curve.
2	WT-data	1 s	15 WTs	Autumn 2015	Wind comparison.
3	CReSS	1 h	2 km	July 2014–July 2015	Wind/power evaluation/comparison.
4	CReSS	1 s	200 m	Autumn 2015	Wind comparison.

done by evaluating predictions from CReSS against WT-measurements, which has never been reported before. The measurements are from the offshore WF in Japan described in Section 5.2. Second, CReSS is evaluated as a tool for wind and WP prediction, and for assessing WP variability. Two different data sets from CReSS are used: see Table 6.1. The first one (#3) is comprised of 30-hour ahead predictions of 2-km spatial resolution and 1-h temporal resolution. As the second data (#4), extremely detailed weather simulations of a domain encompassing the real WF were conducted with CReSS, down to a 200-m spatial resolution and 1-s time resolution. In those simulations, weather conditions are predicted about 5–7 h in the future at three different dates where overlapping 1-Hz measurements at the WF are available.

Predictions from CReSS are evaluated against measured wind speeds, output powers, wind directions, and spectral coherence between wind speeds at different WT locations. In the evaluation of WF outputs, we look at both the WP variability and the WF’s annual energy production. It is then shown that both WP variability on an hourly scale and the WF’s energy production are captured by CReSS—thus confirming the effectiveness of CReSS for large-scale WP analysis. Also, measured and predicted spectra of wind speeds are evaluated and found particularly close for time-scales on the order of hours. The spectral coherence [135, 136] determines the smoothing characteristics of the WF’s output, and is thus an important measure when assessing the WF characteristics. For the detailed simulations, it is shown that the wind direction and spectral coherence between different locations become close to measurements—indicating the potential use of CReSS to analyze prospective WF locations at a detailed scale.

6.1.3 Outline

The remainder of this chapter is organized as follows. Section 6.2 describes the measurement data and their processing before evaluation. Section 6.3 introduces the CReSS weather model, explains the simulation setting, and describes the utilized outputs from the model. The mathematical methods used in the comparison are provided in Section 6.4. In Section 6.5 predicted wind and WP characteristics are evaluated, and the prediction errors are analyzed and discussed. Conclusions are given in Section 6.6.

6.2 On measurements and data processing

This section describes the obtained WF measurements and how they are processed before evaluation against predictions from CReSS. Two types of data from the WF are considered in this chapter (see Table 6.1):

1. One year 10-min averaged data for every WT, which were acquired between July 2014 and July 2015.
2. 1-Hz sampled data were acquired around three dates in August/September of 2015, where each set of WT-data consists of 3600 samples, resulting in a temporal length of 1 h. There are in total 44 data sets of such 1-Hz data, where each WT is monitored during three 1-h periods expect for one WT which is only monitored twice. Data around a specific date are not synchronously measured at the WTs, but pairwise overlaps between two WTs occur.

Both types of data include the following measurements: wind speeds $\tilde{v}_w$, electric powers $\tilde{P}_w$, rotor speeds $\tilde{\omega}_r$, and wind/nacelle directions $\tilde{\phi}/\tilde{\phi}_n$. The power extracted by a single WT is given by (4.2). By assuming that c_p is constant on average for each 10-min period by controlling the tip speed ratio λ around its optimal value (called Maximum-Power-Point-Tracking (MPPT)) [65], 10-min mean wind speeds are estimated by

$$\bar{v}_w = \left(\frac{2\tilde{P}_w}{\rho c_p A_r} \right)^{1/3}, \quad (6.1)$$

from measurements of electric power outputs, for wind speeds above cut-in v_{c-in} and below nominal (where the WT starts to produce rated power). Wind speeds and directions are measured by anemometers mounted at the nacelle of the WTs. Such measurements are inevitably influenced by the rotor and the rest of the WT structure [129]. According to our observations, wind speeds are generally measured lower at WTs in operation than at non-operating ones. Measurements at the nacelle and from met masts have a close to linear relation [137, 138], and nacelle measurements can provide accurate estimations of the true wind speeds [139]. However, [140] found a nonlinear relation between nacelle and mast measurements. Here, considering $\bar{v}_w$ as a good approximation of the “effective” wind speed experienced by the WT, it is suggested based on [137–139] to adjust the wind speed measurements at each WT according to the following linear model:

$$\bar{v}_w := c_1 \tilde{v}_w + c_2, \quad (6.2)$$

where optimal c_1 and c_2 can be determined in a least-squares sense by a linear regression.

When conducting the comparison between 1-h averaged measurement data and prediction data, if not stated otherwise, the rules listed below are followed to ensure more reliable and similar conditions for the comparison:

1. Two different cases are considered separately: WTs in operation (on) and not in operation (off). A WT in operation is defined as when its blades are rotating (above or equal to the designed rotor speed at v_{c-in}) and power is being produced. A WT not in operation has its rotor standing still and does not produce power.
2. The absolute difference between nacelle and wind directions is less than 30° for every 10-min measurement.
3. When calculating the averaged WF speed/power/direction, we require that at least 12 of the 15 WTs are in operation.

Note that in most cases when a WT is not operating, it is due to low wind speed, hence the case of non-operating WTs is almost exclusively restricted to low wind speeds. The wind speed at the τ -th hour is the mean of the speed at that time and the five preceding 10-min measurements: $\tilde{v}_\tau := \langle [v_\tau, v_{\tau-10\text{min}}, \dots, v_{\tau-50\text{min}}] \rangle$.

Wind directions are averaged in time, or spatially over WTs. For example, we average 10-min directions to 1-h averages, and then average that data over all WTs (spatially). The averaging procedure for wind directions is as follows. Denote the measured or predicted wind direction by ϕ_i (rad), where i either represents a time-instance or the index of a WT. Then, according to [141], the mean angle (here, the mean incoming wind direction ϕ_{in}) of N measurements ϕ_i ($i = 1, \dots, N$) is calculated by

$$\phi_{in} = \text{atan2} \left(\frac{\sum_{i=1}^N \sin(\phi_i)}{\sum_{i=1}^N \cos(\phi_i)} \right) = \text{atan2} \left(\frac{\phi_{num}}{\phi_{den}} \right) := \text{atan2}(\phi_{num}, \phi_{den}), \quad (6.3)$$

where **atan2** is an arctangent function with two input arguments—numerator ϕ_{num} and denominator ϕ_{den} separately—to keep track of the quadrants in order to output a result in the correct one. More specifically, **atan2** depends on the signs of ϕ_{num} and ϕ_{den} : e.g. if $\phi_{num} > 0$, then $\text{atan2}(\phi_{num}, \phi_{den}) = \arctan(\phi_{num}, \phi_{den})$. It is used in various programming languages, and is computed with the function **atan2** in MATLAB.

The occurrence of mean wind speed directions, computed by averaging over all WTs using (6.3), for the 10-min data is shown in Fig. 6.1. The dominant direction (where the wind comes from) is wind from $[30^\circ, 60^\circ]$ clockwise from north, i.e. wind from ocean, and from $[180^\circ, 210^\circ]$, i.e. wind from land.

6.3 CReSS model and outputs

CReSS is based on a non-hydrostatic and compressible weather model and was originally intended for simulation of cloud and precipitation systems including extreme weather phenomena such as typhoons. The basic equations of CReSS comprise momentum equations including the earth's rotation, thermodynamic equations, the continuity equation,

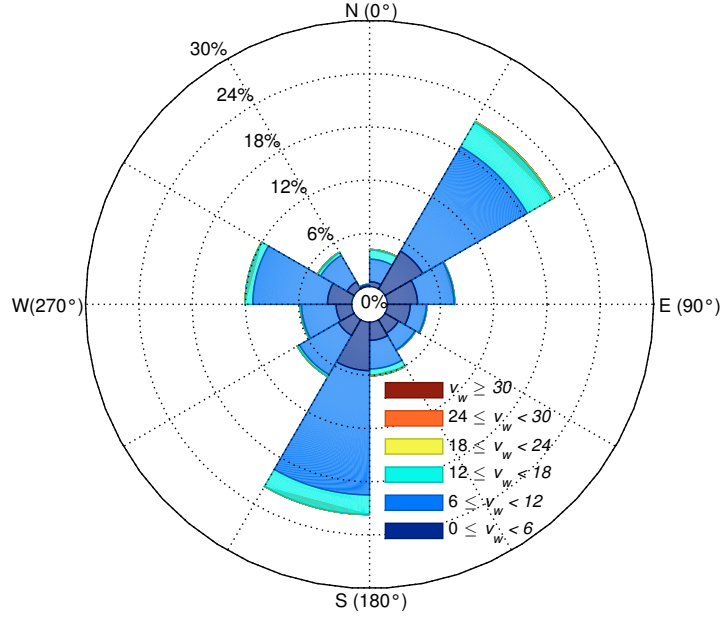


Figure 6.1: Occurrence (in percent) of mean wind directions at the WF. The mean direction is the mean of 10-min averaged directions over all WTs, calculated with (6.3). The diagonal axis indicates the occurrence frequency in percent. The dominant directions (where wind comes from) are clearly wind from $[30^\circ, 60^\circ]$ clockwise from North, i.e. wind from ocean, and $[180^\circ, 210^\circ]$, i.e. wind from land.

and mixing ratios of water vapor and hydrometers. The prognostic variables are three-dimensional components of wind velocity, Turbulent Kinetic Energy (TKE), perturbations (deviations from basic states) of atmospheric pressure and temperature, water vapor mixing ratio, and cloud physical variables [142]. Variables are defined on the so-called Arakawa-C grid [143] (same as the WRF model [144]) in the horizontal domain, and on the Lorentz grid in the vertical domain. A time-splitting method [145] is used in the time-integration, where terms related to acoustic waves are separately computed with a smaller time-step. Cloud physics are based on a bulk method of cold rain and is elaborately described in [142], and consider water vapor, rain, cloud, ice, snow, and graupel. The 1.5 order TKE closure scheme is adopted in CReSS, and there is no parameterization of the planetary boundary layer. The terrain is modeled with a 1-km resolution from the Global 30 Arc-Second Elevation (GTOPO30) from the U.S. Geological Survey (USGS). GTOPO30 was also used in e.g. [126] together with the WRF model. CReSS utilizes terrain-following coordinates to account for the terrain.

Data from two types of CReSS simulations are utilized. The first type is daily predictions conducted with a spin-up period of 3–6 h, followed by a 30–33 h prediction period, resulting in a total run-time of 36 h. Here, the spin-up period is assumed to be 6 h for every run to be on the safe side, with up to 30-h predictions that follow. The next 36-h run starts 24 h after the former run, and consequently a continuum of daily predictions is achieved. The domain $\mathcal{D}_1$ (see Fig. 6.2) of these simulations encompasses a $2300 \text{ km} \times$

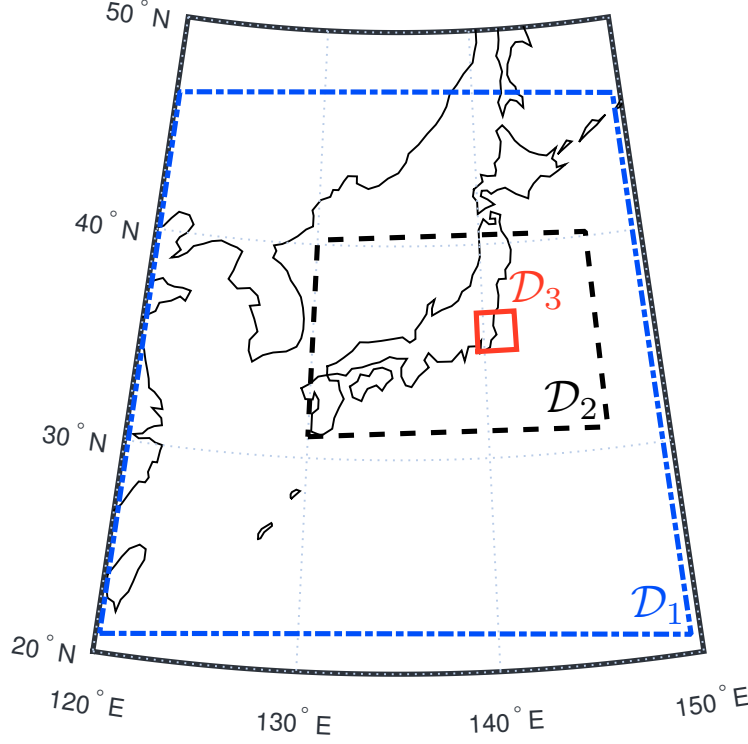


Figure 6.2: Simulation domains $\mathcal{D}$ of the CReSS simulations.

2800 km area with min/max longitudes of $120.20^\circ/149.50^\circ$ and latitudes of $20.88^\circ/46.48^\circ$. The horizontal resolution is 2 km and the output interval is 1 h. The boundary and initial conditions are obtained from the Global Spectral Model (GSM) provided by the Japanese Meteorological Agency (JMA).

The second type is detailed simulations of 200-m horizontal resolution and 1-s output interval. The simulations are nested within 1-km resolution simulations of western Japan of a domain $\mathcal{D}_2$ with a size of about $1500 \text{ km} \times 1000 \text{ km}$ with min/max longitudes of $130^\circ/146.639^\circ$ and min/max latitudes $31^\circ/40.223^\circ$. The domain $\mathcal{D}_3$ of the nested 200-m simulations is about $200 \text{ km} \times 200 \text{ km}$ with min/max latitudes of $34.96^\circ/36.80^\circ$ and longitudes of $139.62^\circ/141.86^\circ$. The boundary and initial conditions of the 1-km resolution simulations are provided by the regional reanalysis (MANAL) conducted by JMA and have 5-km horizontal resolution. The results are used for initial and boundary conditions of the 200-m resolution simulations.

Only wind speeds from CReSS are considered here. Examples of wind conditions around the WF obtained from CReSS are shown in Section 6.5, where the WT-positions are indicated by red dots, and the coloring indicates the horizontal wind speed magnitude. The wind speed magnitude is computed by $\hat{v}_w(t) = \sqrt{v_{x(t)}^2 + v_{y(t)}^2}$, where the subscripts denote components of the three-dimensional wind speed vector. Moreover, the wind speed predicted at a height z_1 , denoted by v_{z_1} , is scaled to the height of a turbine $z_2 \approx 60 \text{ m}$ by

the following formula [120, 146, 147]:

$$v_{z_2} = v_{z_1} \frac{\ln(z_2/z_0)}{\ln(z_1/z_0)}, \quad (6.4)$$

where $z_0 = 2 \times 10^{-4}$ m is the roughness length for open sea. For the 2-km simulations, wind speeds predicted at a height of 50 m are utilized, whereas the wind speed is predicted at WT-height in the detailed 200-m simulations.

6.4 Methods

This section first presents the standard statistical metrics which are used to evaluate predictions in Section 6.5. Following that, Power Spectral Density (PSD) of wind speed and wind power fluctuations are discussed, and some results on the estimated PSDs of measured and predicted wind speeds and WPs are presented.

6.4.1 Statistics for comparison

When comparing predictions against measurement, the prediction errors can be divided into two types: amplitude and phase errors [146]. Amplitude errors occur when the feature—for example a sudden increase or drop in wind speed—is captured, but the magnitude differs. Phase errors refer to the prediction errors occurring when the feature is captured but time-shifted. Phase errors and their contribution to the prediction errors can be quantified by considering auto-correlations [146]. The prediction error is given by

$$\Delta\hat{y} := \hat{y} - \tilde{y}, \quad (6.5)$$

where $\hat{y}$ and $\tilde{y}$ are the predicted quantity by CReSS and the measured quantity at the WF, respectively, and y is thus either wind speed v_w or power P_w . Prediction errors are evaluated by the Root Mean Square of Errors (RMSE) defined as

$$r_\Delta := \sqrt{\langle (\Delta\hat{y})^2 \rangle}, \quad (6.6)$$

where r_Δ is given in [146] as follows:

$$r_\Delta^2 = \Sigma_{\text{bias}}^2 + \sigma_\Delta^2 = \Sigma_{\text{bias}}^2 + (\Delta\sigma)^2 + \xi^2, \quad (6.7)$$

where the so-called error bias is defined as $\Sigma_{\text{bias}} := \langle \Delta\hat{y} \rangle$, and $\sigma_\Delta := \sigma(\Delta\hat{y})$ is the Standard Deviation of Errors (STDE), where $\sigma(\cdot)$ denotes the standard deviation operation. Moreover, the standard deviation error $\Delta\sigma$ and the so-called dispersion ξ are in [146] defined as follows :

$$\Delta\sigma := \sigma(\hat{y}) - \sigma(\tilde{y}), \quad \xi := \sqrt{2\sigma(\hat{y})\sigma(\tilde{y})(1 - r_c)}, \quad (6.8)$$

where r_c is the cross-correlation coefficient between $\hat{y}$ and $\tilde{y}$. The amplitude errors are quantified via the error bias and the standard deviation error $\Delta\sigma$. The dispersion ξ indicates the influence of phase errors to the RMSE.

In the case of wind direction, we denote the predicted and measured directions in radians by $\hat{\phi}$ and $\tilde{\phi}$, respectively, and calculate the shortest angular difference between measured and predicted direction at a time instance by

$$\Delta\hat{\phi} := \text{atan2}\left(\sin(\hat{\phi} - \tilde{\phi}), \cos(\hat{\phi} - \tilde{\phi})\right), \quad (6.9)$$

where it follows the notation in (6.3). Here, a positive value of $\Delta\hat{\phi}$ means that the predicted direction (angle) is larger than the measured direction in a clockwise sense. Since the direction is a circular variable, the standard deviation of wind direction prediction errors is computed in accordance with [141] as follows. Define a vector $\mathbf{r} := 1/N \cdot [\phi_{\text{num}}, \phi_{\text{den}}]^\top$, with notation from (6.3), and its norm is given by $R := \|\mathbf{r}\|$, where $\|\cdot\|$ denotes the euclidean norm. Then, the angular deviation is calculated by $\sigma_\phi := \sqrt{2(1 - R)}$.

To compare and evaluate variability of predicted and measured WF outputs ($\hat{P}_w$ and $\tilde{P}_w$), we compute WP increments $dP_w(t, \tau)$ by (5.4), defined as

$$dP_w(t, \tau) := P_w(t + \tau) - P_w(t), \quad (6.10)$$

where τ is now the time-shift in hours. Increments are called *ramps* if they exceed a certain threshold [115, 148]. Hourly variability of WP on a large scale was investigated with increments in [27], where the standard deviation of increments was found to be the best indication of reduced variability of aggregated WP (WP smoothing). A good agreement between predictions and measurements would justify the usage of CReSS for studying the smoothing of large-scale aggregated WP. Such a study is conducted in Chapter 7.

6.4.2 Power spectral density and spectral coherence

To quantify correlation in the frequency domain between two different time-series of wind speeds or powers, the so-called spectral coherence via PSD has been used [66, 135, 149]. Let $S(f)$ denote the PSD of a signal with variance σ^2 . The integration of $S(f)$ over frequency f is then by definition equivalent to the variance of the signal:

$$\int_0^\infty S(f) df = \sigma^2. \quad (6.11)$$

The PSD of WP has been shown to follow the so-called Kolmogorov spectrum [150, 151] according to $S(f) \sim f^{-5/3}$, and the spectrum of wind speed is similar [64]. The so-called Kaimal spectrum [152, 153]—described in e.g. [92] for generation of synthetic wind speeds—falls off with a similar rate as Kolmogorov's. The PSDs are estimated with the function `pwelch` in MATLAB.

Figure 6.3 (a) shows PSDs of measured 10-min averaged power output (one WT), 1-h wind speeds averaged over all WTs, and predicted (simulated) wind speeds from CReSS.

Figure 6.3 (b) shows PSDs of 1-Hz measured wind speed and power and 1-Hz wind speed from CReSS. It is evident from these examples that the PSDs of wind speed and power scale according to Kolmogorov’s spectrum. The largest deviations from the scaling-theory are observed for 1-Hz measured electric power. These results show that the conversion from wind to electric power primarily affects the spectrum in the short-term regime. It is also evident that the spectra of the 1-h wind speed predictions and 1-h averaged wind speeds are in excellent agreement: see Fig. 6.3 (a). For the 1-Hz data on wind speeds, the CReSS spectrum scales similarly but its magnitude differs. The cause of this is the absence of turbulent fluctuations observed in the real measured data, which is a result of inherent spatial and temporal smoothing effects in the CReSS model. A new method for generating synthetic wind speeds is proposed in [153], which also reviews commonly used methods, and could be used to add fluctuations to the CReSS data to obtain a more realistic spectrum.

Here, spectral coherence—i.e. the magnitude of spectral overlap between wind at different locations—is employed to compare the coherence between wind speeds at two WTs with real and simulated data. It is defined in [135, 136] as follows:

$$\Gamma_{ij}(f) = \frac{|S_{ij}(f)|}{\sqrt{S_{ii}(f)S_{jj}(f)}}, \quad (6.12)$$

where S_{ii} is the PSD of wind speed at a location i , and S_{ij} the cross spectral density between wind speeds at locations i and j . In [135], an exponential approximation of (6.12) is introduced as follows:

$$\Gamma_{ij}(f) := \exp(-Cd_{ij}^k f), \quad (6.13)$$

where d_{ij} is the distance between i and j , C is a decay coefficient, and k a constant. This approximation showed good agreement with real measurements in Japan [135].

6.5 Evaluation of predicted wind and wind power characteristics

This section first compares measured and predicted wind characteristics at the WF in terms wind speed and direction. Then, WP prediction errors, power increments and errors of power increments, and the WF’s annual energy production, are all investigated. Finally, spectral coherence of wind speeds between different WT-locations in the WF with the extremely detailed CReSS simulations (1 s, 200 m) are compared with 1-Hz measurements. Figure 6.4 shows example time-series of measured (1-h averaged, and averaged over at least 12 WTs) and predicted wind speeds, power outputs, and directions.

6.5.1 Wind speed

For this comparison, the data sets 1 and 3 in Table 6.1 are used. First, the two constants in the formula (6.2) for the adjustment of measured wind speed $\tilde{v}_w$ are determined. To

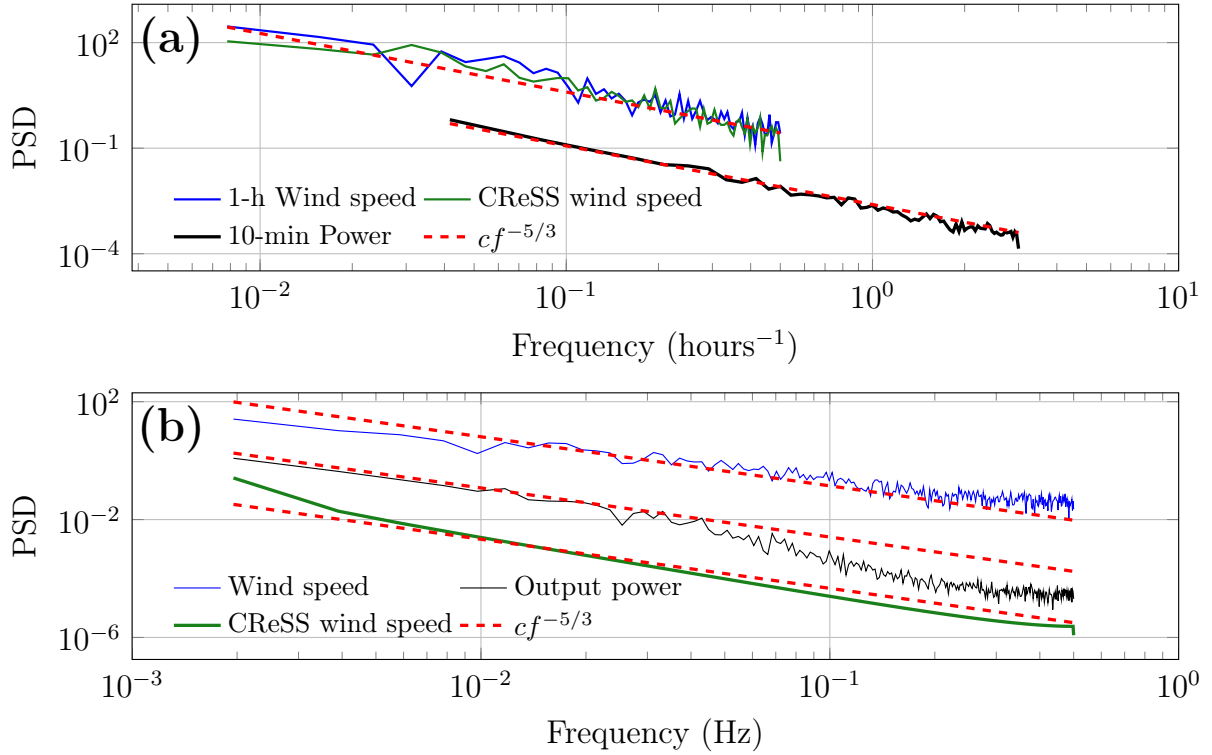


Figure 6.3: Power spectral densities (PSD) of (a) 1-h averaged wind speed measurements (over all WTs) and predictions from CReSS, plus 10-min averaged output power; (b) 1-Hz sampled data including both real measured data and predictions from CReSS.

this end, a reasonable assumption on c_p is needed—typically 0.46–0.48 with respect to electric power [87], or about 0.40 [154]—and here c_p is fixed at 0.42. Small variations of this constant would not considerably affect the results presented here, which is briefly discussed in the next paragraph. Then, the data are sorted such that only measurements from WTs in operation are included, the maximal difference between nacelle and wind directions is kept within 2° (the results will not vary much if this tolerance is relaxed but will result in a huge amount of data not providing much additional information), and including only wind coming from the ocean with $\tilde{v}_w$ satisfying $4 \text{ m/s} \leq \tilde{v}_w \leq 11 \text{ m/s}$. These conditions are enforced to ensure similar operating conditions and to avoid data from WTs operating in the flat region of the power curve, which would make (6.1) an invalid method of estimating the “effective” 10-min wind speed. In fact, about 85% of the measured and utilized wind speeds are less than or equal to 11 m/s. Adjusted wind speeds according to (6.2) are plotted against calculated wind speeds $\bar{v}_w$ in Fig. 6.5, and their relation is well approximated by a linear function. The constants in the model (6.2) become $c_1 = 1.12$ and $c_2 = 0.48$.

The PDFs of measured and predicted wind speeds are shown in Fig. 6.6 (a) for WTs in operation ‘ON’ with $c_p = 0.42$ and not in operation ‘OFF’; and in (b) for different values of c_p . The distribution of wind speeds is commonly modeled by the so-called Weibull

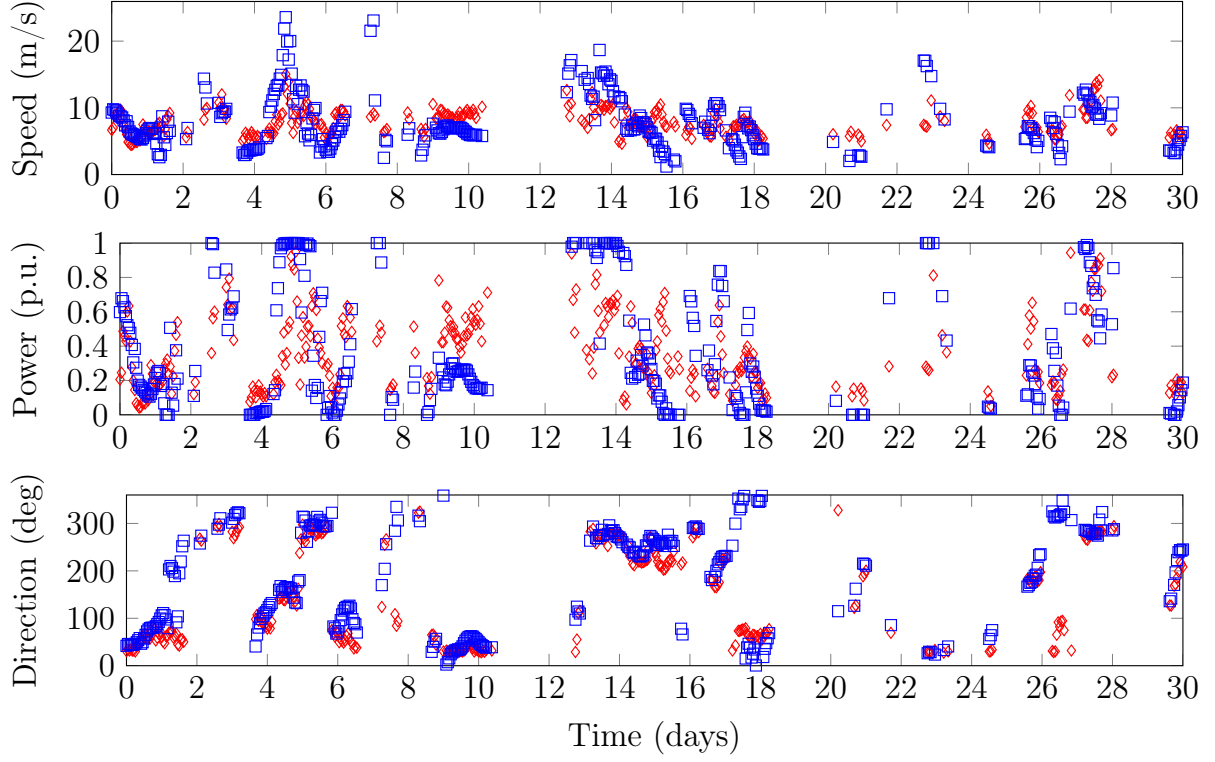


Figure 6.4: Example time-series of measured and predicted wind speeds, powers, and directions. Measurements are depicted by red diamond marks and predictions by blue squares.

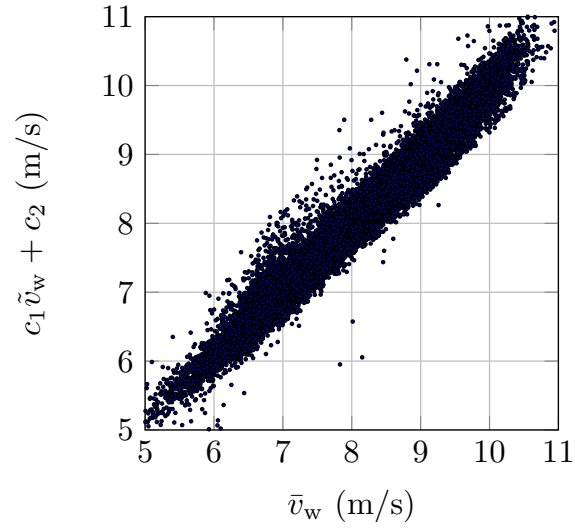


Figure 6.5: Measured wind speeds $\tilde{v}_w$ adjusted according to the linear model (6.2) and plotted against wind speeds calculated with (6.1).

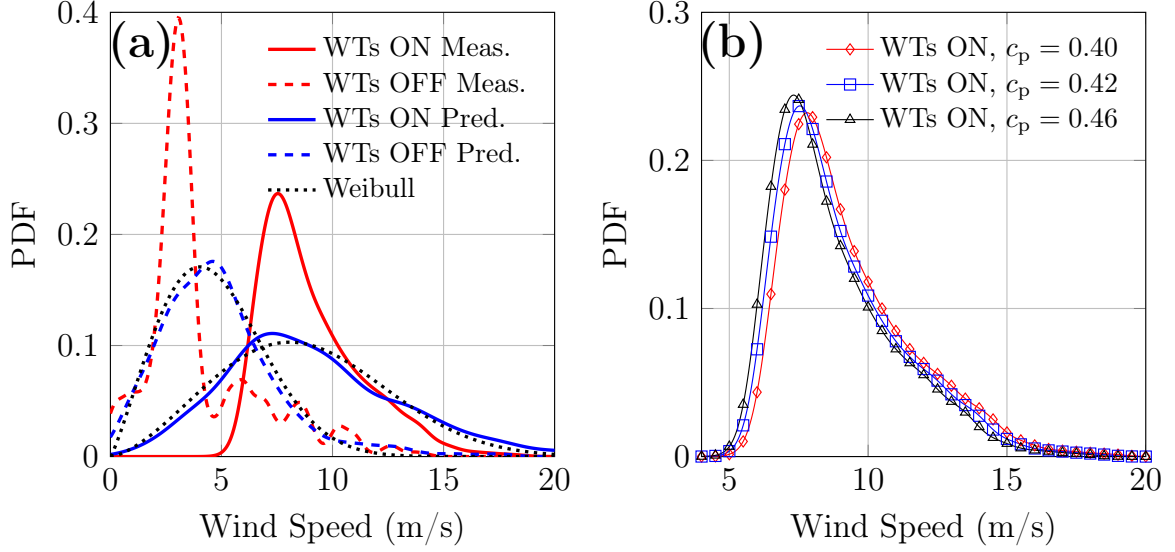


Figure 6.6: (a) PDFs of predicted and measured wind speeds and comparison with Weibull distributions. (b) PDFs of adjusted wind speed measurements according to (6.2) for different values of the power coefficient c_p .

distribution [155]. PDFs of wind speeds from CReSS are close to Weibull PDFs, while the measurements are not. The cause of this can be attributed to the two “data categories” that are adopted: WTs on and WTs off. In the case of WTs on, we measure wind speeds of at least v_{c-in} since this is required for power production. In the case of WTs off, mostly low speeds are measured because the most common cause of WTs off is due to low wind speed. Hence, a complete description of the wind speed distribution at the WF cannot be obtained in neither of these cases. In particular, wind speeds measured when the WTs are off are overestimated by CReSS. For cases when the WTs are on which are for “higher” wind speeds, CReSS predicts low wind speeds at many instances, and the peaks of the PDFs are close. Fig. 6.6 (b) shows that the PSD of wind speeds changes marginally when c_p is varied.

The three conditions outlined in Section 6.2 are enforced when sorting the data. All the measured 10-min wind speeds are now adjusted according to (6.2) and averaged to hourly wind speeds. Then, the averaged 1-h wind speeds are averaged over the WTs (at least 12) in the WF. The PDFs and CDFs of wind prediction errors $\Delta\hat{v}_w$ are shown in Figs. 6.7 (a)–(b) for the case of operating WTs, and for different averaging periods. Wind speeds averaged over several hours are computed by the moving average as follows:

$$u_n(d) := \langle [v_{n-d}, \dots, v_n, \dots, v_{n+d}] \rangle, \quad (6.14)$$

where d (hours) is a time-shift, and e.g. $d = 1$ h and $d = 3$ h give the 3- and 7-h averages, respectively, as they are called here. The moving average will smooth the wind speed and is expected to decrease the contribution of phase-errors to the RMSE. In Fig. 6.7 (a), the peaks of all PDFs are slightly and negatively shifted from zero, and the

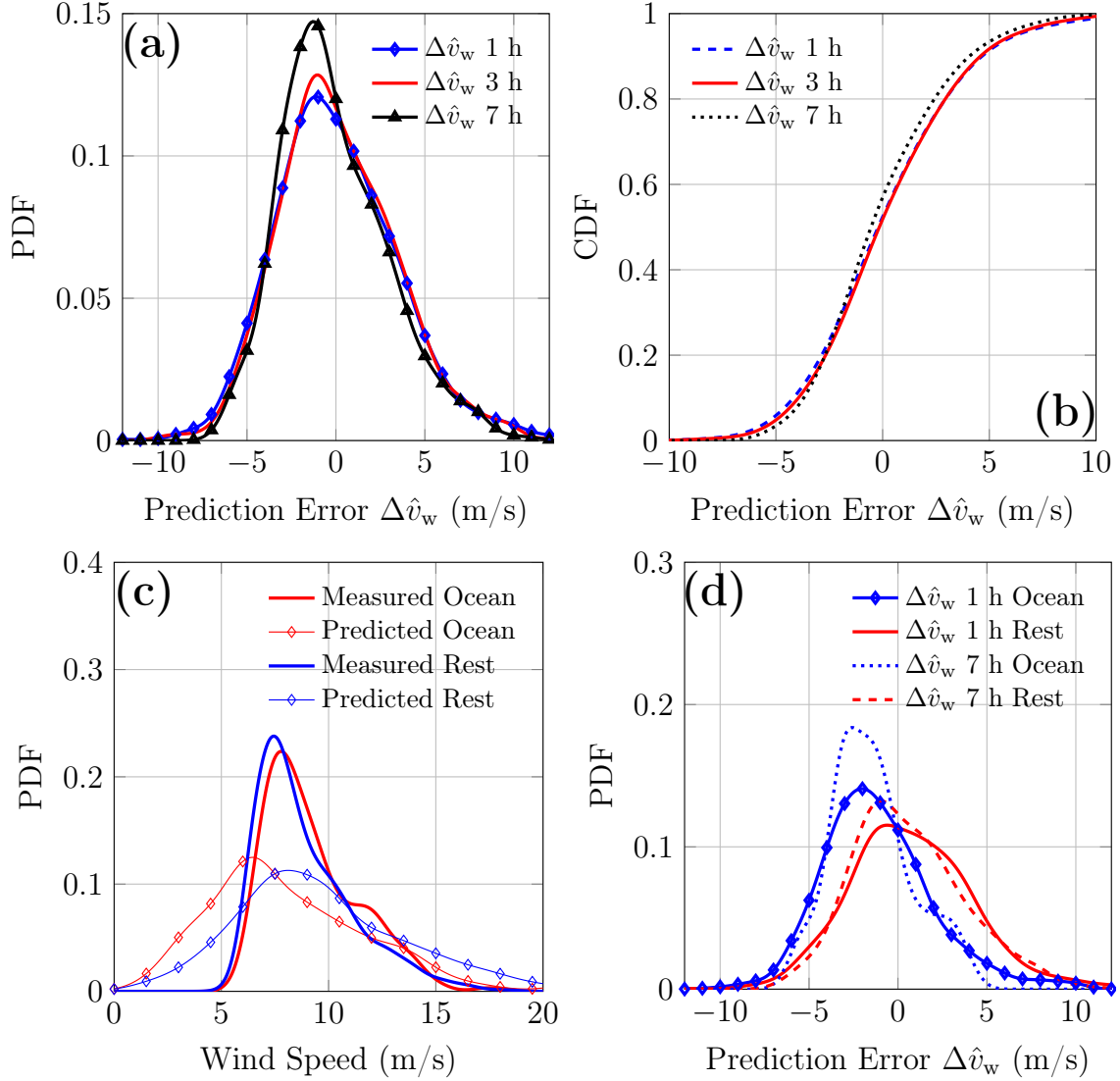


Figure 6.7: (a)–(b) PDFs and CDFs of wind prediction errors $\Delta \hat{v}_w$ for WTs in operation. (c) PDFs of wind speeds for wind from ocean, and all other directions denoted by ‘Rest’. (d) PDFs of $\Delta \hat{v}_w$ for 1- and 7-h averaging for wind from ocean and all other directions.

PDFs reveal a tendency of CReSS to predict lower wind speeds than observed. Longer tails on the positive axis reveal that CReSS occasionally overestimate the wind speed at the WF (cf. Fig. 6.6 (a)). Calculated values on RMSE, STDE, bias Σ_{bias} and dispersion ξ are given in Table 6.2. RMSE, STDE and ξ all decrease with longer averaging time, while Σ_{bias} remains small and does not show any clear trend. These results show that dispersion accounts for a large part of the RMSE. As a rough comparison, we note that [123, 128, 130] achieve an RMSE of about 2.4–4 m/s, where [123, 130] used 6-h time-resolution predictions, and [128] up to 2-day-ahead predictions, which were compared against 10-min measurements. Up to 2-day-ahead predictions from the WRF model were

Table 6.2: Statistics of wind speed and direction prediction errors.

	Wind Speeds (m/s)				Direction (deg)	
	RMSE	STDE	Σ_{bias}	ξ	STDE	Σ_{bias}
1 h	3.49	3.48	0.09	2.98	28.07	12.51
3 h	3.32	3.31	0.16	2.74	24.51	13.56
5 h	3.15	3.15	0.08	2.51	21.74	13.45
7 h	2.98	2.98	-0.11	2.33	20.48	13.46
9 h	2.96	2.96	0.02	2.23	16.45	12.58
11 h	2.99	2.98	0.09	2.22	16.05	11.88

evaluated against 10-min measurements in [126], and an RMSE of 2–2.5 m/s was achieved, where the best score was obtained for moderate wind speeds (8–12 m/s). In the light of those results, the results here seem to be normal.

Figure 6.7(c) shows PDFs of wind speeds for two cases: wind from ocean $\phi_{\text{in}} \in [0^\circ, 120^\circ]$, and wind from all other directions, denoted by ‘Rest’. For measurements, the difference between these two cases appears smaller than the difference in the case of predictions. PDFs of prediction errors for these two cases are shown in Fig. 6.7(d). In the case of wind from ocean, the PDFs have negative biases of about 2.5 m/s, while the biases are substantially smaller in the case of other wind directions. This phenomenon is likely caused by the coarse terrain model which does not accurately capture the terrain characteristics around the coastline. More precisely, the simulated wind close to the coast experiences a rougher terrain than in reality for wind coming the ocean, and hence lower wind speeds are predicted. On land, on the other hand, the terrain is likely modeled more smooth than in reality which causes less friction and higher wind speeds, and the wind speed is more likely to be overestimated: see Fig. 6.7(c).

6.5.2 Wind power

In this section, predicted WF outputs with CReSS are compared against measured WF output (estimated in p.u. (per unit) as the average of WT outputs). A power curve which converts hourly wind speeds to WT outputs is derived by first sorting the data according to the following criteria:

1. The absolute difference between nacelle and wind directions is less than 30° for each 10-min measurement.
2. Only one WT is considered (WT No. 7, see Fig. 5.1), and it needs to be in operation.
3. Only cases of wind coming from ocean are included to ensure similar conditions.
4. Wind speeds satisfy $\tilde{v} \geq v_{\text{c-in}} (= 4 \text{ m/s})$.

After sorting according to the above criteria, the data were averaged around $\pm 0.2 \text{ m/s}$ from fixed steps in wind speeds (spaced 0.2 m/s apart) to give an averaged curve. Outliers

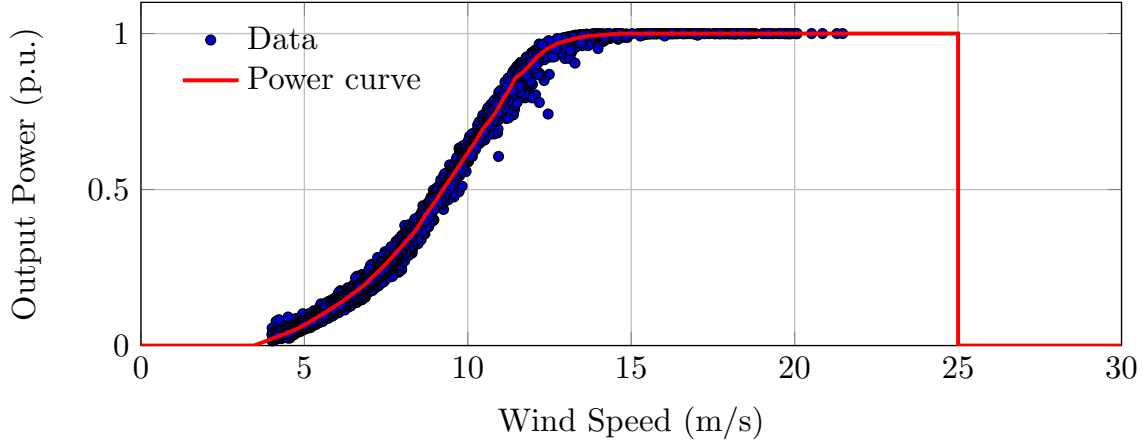


Figure 6.8: Power curve based on 1-h averaged measurements from WT 7 (see Figure 5.1).

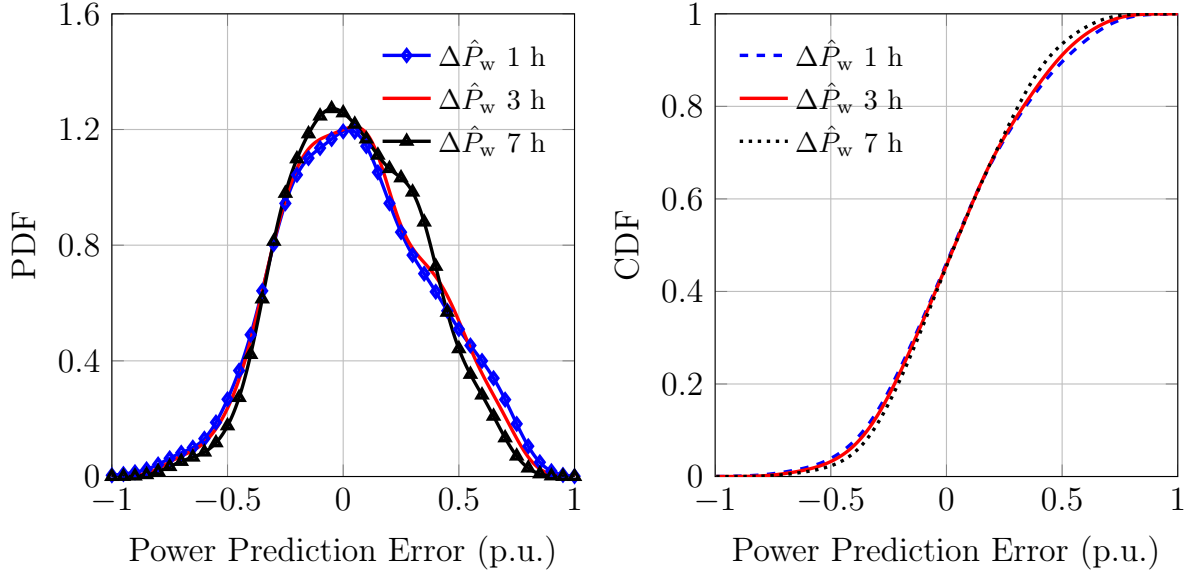


Figure 6.9: PDFs and CDFs of WP prediction errors in p.u. for different averaging periods.

further than 1.5 standard deviations in p.u. from the averaged curve were removed. The remaining data were then again averaged to give a final power curve. The resulting power curve, which is used in the following computations, is shown in Fig. 6.8 together with the sorted wind speed and power data.

The same setting as for the wind speed comparison is now considered. The PDFs and CDFs of the WP prediction errors $\Delta\hat{P}_w$ are shown in Fig. 6.9 for 3-h and 7-h averaging periods in addition to hourly values. Similar to the case of wind speeds, statistical measures of prediction errors are shown in Table 6.3. From these results, we see that the RMSE and dispersion readily decrease with longer time-averaging. Note that the nonlin-

Table 6.3: Statistics of wind power prediction errors in p.u.

	RMSE	STDE	Σ_{bias}	ξ
1 h	0.32	0.32	0.05	0.30
3 h	0.31	0.30	0.05	0.28
5 h	0.29	0.29	0.05	0.27
7 h	0.28	0.28	0.05	0.26
9 h	0.28	0.27	0.04	0.25
11 h	0.26	0.26	0.05	0.24

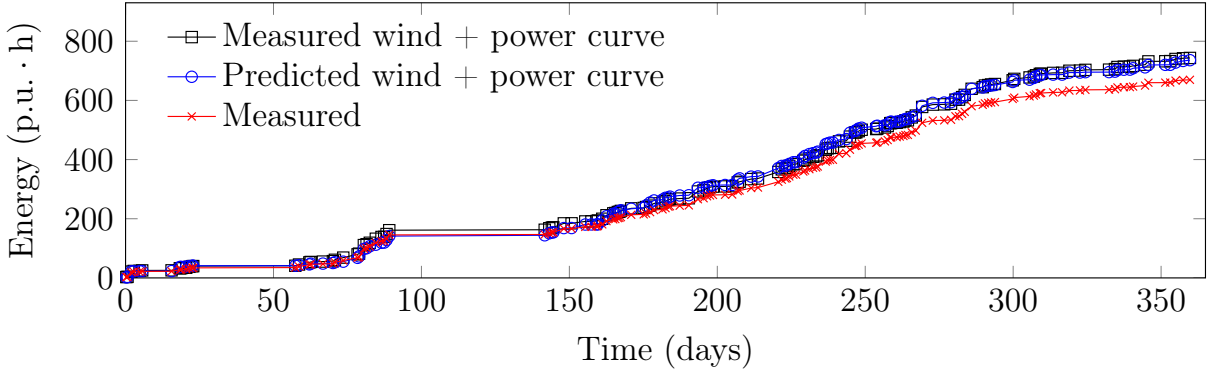


Figure 6.10: Measured and predicted energy production.

ear power curves of WTs or WFs amplify the wind prediction errors and thus give rise to greater errors in power [146].

Predicted and measured accumulative energy productions (p.u. · h) are estimated by piecewise integration of WF-outputs (p.u.). From here on in this section, in order to obtain as long interrupted time-series segments as possible, the data are linearly interpolated when the time-step between available samples does not exceed 4 h. The total energy productions are shown in Fig. 6.10, where a good agreement is achieved between predictions and measurements. The result by using the power curve with measured wind speeds also becomes close to measurements. The results indicate that the utilized power curve is not a large source of error. And also, that the differences in wind speeds evaluated in the previous section do not cause a great difference in the total energy production, which indicates that nacelle measurements do not always provide accurate wind speeds. Note, however, that these results are based on data for only a fraction of every hour of the year—i.e. times when both measurements and predictions are simultaneously available and the measurements satisfy the requirements listed in Section 6.2.

Variability of WP is now evaluated by looking at the power increments $dP_w(t, \tau)$ as defined in (6.10), where τ (hours) is a time-shift. The measured and predicted increments are denoted by $d\tilde{P}$ and $d\hat{P}$ in their abbreviated notation, respectively. The PDF of $dP_w(t, \tau)$ reveals and clearly illustrates the probability of WP fluctuations on time-scales defined by τ . In Fig. 6.11 (a) we show PDFs of $dP_w(t, \tau)$ for $\tau = 1$ h, and of averaged power

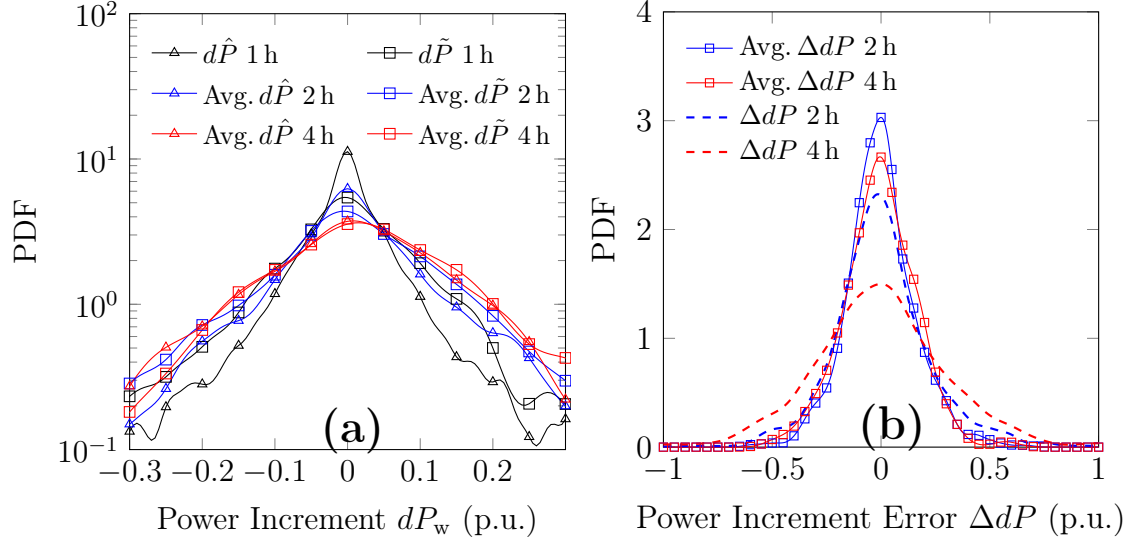


Figure 6.11: (a) PDFs of power increments for different time-shifts τ : for non-averaged increments ($\tau = 1$ h) and for increments computed from averaged power. (b) PDFs of averaged and non-averaged power increment errors for τ set at 2 and 4 h.

increments with $\tau = 2$ h and 4 h. Note that the so-called averaged power increments with e.g. $\tau = 2$ h (or $\tau = 3$ h), are computed with powers averaged by the moving average (6.14) with $d = 1$ (or $d = 2$). I.e. the increments themselves are not averaged. As expected, larger WP fluctuations are more likely for larger τ . For the non-averaged increments with $\tau = 1$ h, measurements and predictions differ considerably (look e.g. at their peaks). On the other hand, the PDFs of the averaged increments become close for τ set at 4 h, indicating the potential to use CReSS to assess WP variability of the WF on this time-scale.

Now $\Delta dP(t, \tau) := d\hat{P}(t, \tau) - d\tilde{P}(t, \tau)$ is defined as the power increment error. The PDFs of power increment errors with averaged powers are shown in Fig. 6.11 (b). The results are compared against ΔdP without averaging, and show substantially better predictions of increments for hourly averaging. For averaged increments with $\tau = 2$ h, $|\Delta dP|$ is below or equal to 0.2 p.u. with a probability of about 80% according to the CDF, which is not shown here.

6.5.3 Wind direction

Now we look at prediction of wind direction which is one of the most important wind characteristics, and affects e.g. the turbulence intensity in inhomogeneous terrains and the WF-output by development of wakes.

Occurrences of wind direction for CReSS predictions and 1-h averaged measurements are shown in Fig. 6.12, for the exact same time-instances. The results illustrate a clear consistency in the occurrence of wind direction. PDFs and CDFs of $\Delta\hat{\phi}_i$ are shown in Figs.

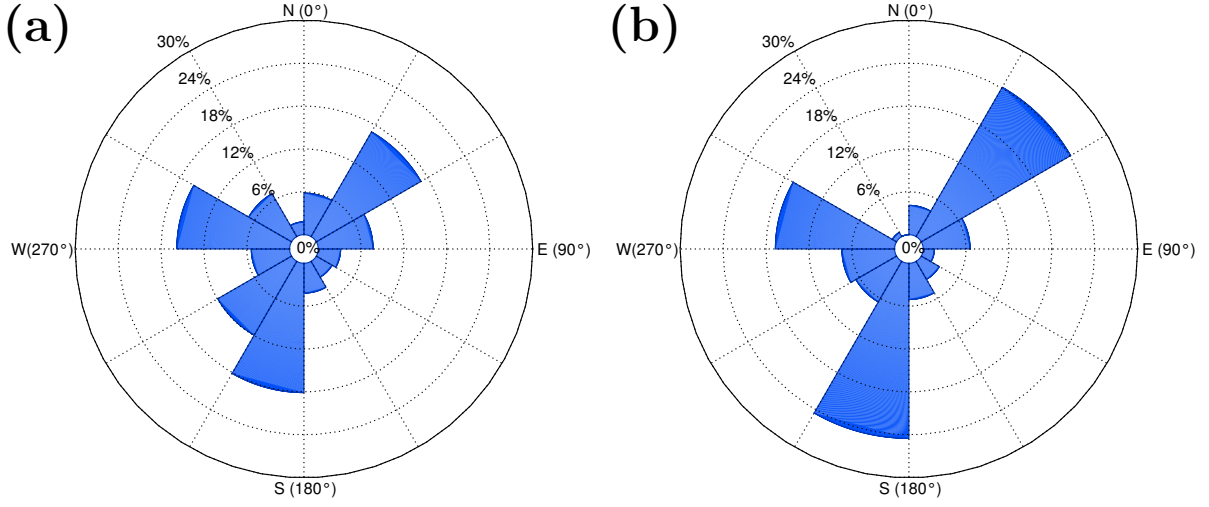


Figure 6.12: Comparison between (a) predicted and (b) measured occurrences of wind directions in the wind farm for synchronized time-instances. For measurements, direction have been averaged over all turbines with (6.3). The diagonal axis indicates the occurrence frequency in percent. The dominant directions are wind from $[30^\circ, 60^\circ]$ and $[180^\circ, 210^\circ]$, clockwise from North.

6.13 (a)–(b) for different maximum allowed tolerances between wind and nacelle directions. The results indicate only minor improvements for smaller tolerances. For a tolerance of 30° , the CDF reveals that CReSS predicts the direction 20° larger than measured with a probability of about 6.8%, and 20° smaller with a probability of about 7.6%. In Table 6.2, STDE and Σ_{bias} of direction errors are shown for different averaging periods. STDE shows a marked improvement as the averaging period increases, and the bias decreases marginally. As rough comparisons, [123] reports STDEs of 13° – 37° for predictions of 6-h time resolution, and [126] achieved an STDE of 35° – 44° for up to two-day ahead predictions by utilizing the WRF model with different boundary and initial conditions. They also achieved an RMSE of direction of 14° – 16° for high wind speeds (above 12 m/s), for which predictions are shown to be much more accurate than for lower wind speeds. A comparison between direction prediction errors for cases of WTs in operation and not in operation is shown in Figs. 6.13 (c)–(d). The figures show great differences in accuracy for these two cases, where the predictions are much more accurate for the case of WTs in operation. The cause of this can be attributed to a weak large-scale wind or mesoscale system in these cases, which increases the variability [156], and makes wind difficult to predict by mesoscale simulations. Similar results were reported in [126]. In addition, the results for high wind speeds ($v_w \geq 8$ m/s) and low wind speeds ($v_w < 8$ m/s) for WTs in operation confirm this. The case of high speeds wind from other directions than from ocean, denoted by ‘High/Rest’ in Figs. 6.13 (c)–(d), gives the smallest prediction errors and suggests the influence from the terrain on the prediction of direction, as in the case of wind speed.

Now we utilize 1-Hz, 5–7 h future predictions, and look at seven data-acquisition

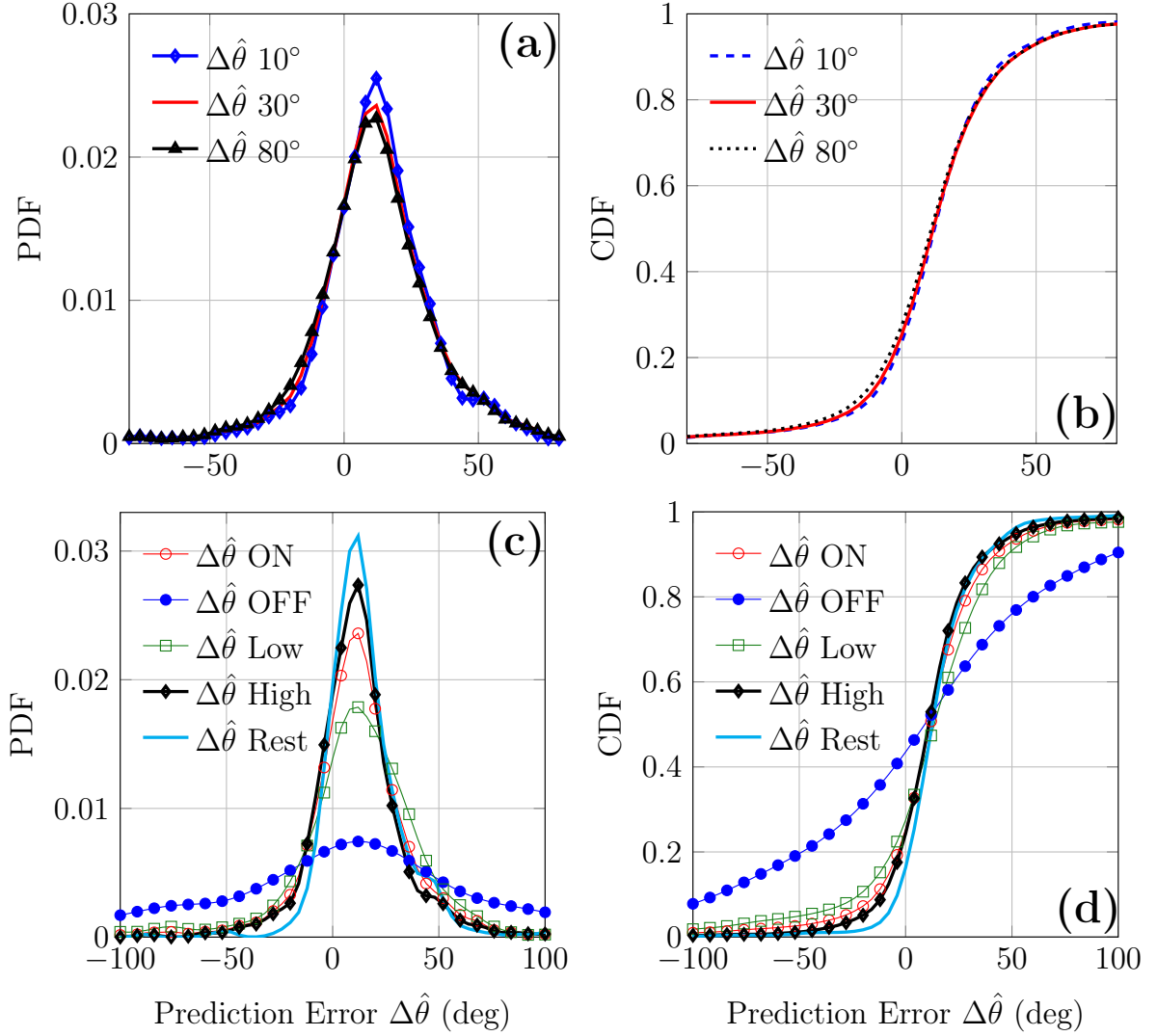


Figure 6.13: PDF and CDF of wind direction prediction errors for (a)–(b) different tolerances between wind and nacelle direction; (c)–(d) cases of WTs in operation (ON) and not in operation (OFF). For WTs in operation, results for high and low wind speeds are provided.

periods where there are overlap between predictions from CReSS and measurements. The data are acquired around three dates, and the wind conditions around the WF at initial times of the obtained CReSS data are shown in Fig. 6.14. At the first date ‘Date 1’, wind comes predominantly from the north; at ‘Date 2’ wind comes from land; and at ‘Date 3’ wind blows from the ocean. STDEs of the wind direction for the 1-Hz data are shown in Fig. 6.15 for 10-min windows of different initial times. In all cases, CReSS predicts the wind direction with an STDE of less than 20° , and in most cases below 10° . This indicates better accuracy than prediction of the hourly wind direction given in Table

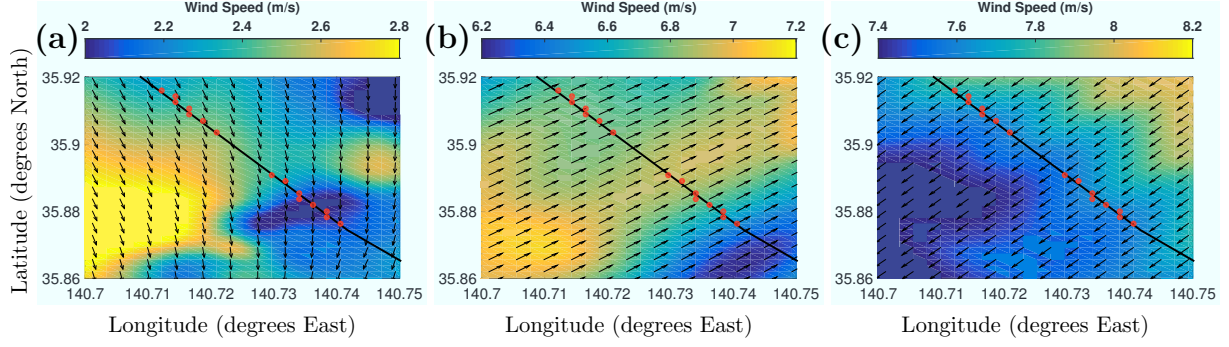


Figure 6.14: Snapshots of wind speed conditions around the WF predicted by CReSS at the three simulated dates ((a)-(c)). Red dots indicate WT locations, and the black line roughly depicts the coast line.

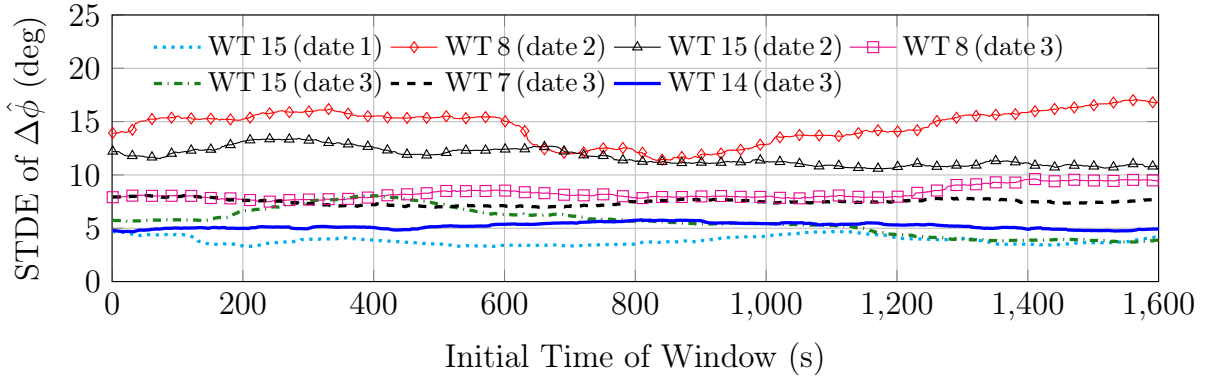


Figure 6.15: STDE of wind direction for 10-min windows with different initial times, for seven cases of overlapping 1-Hz data.

6.3. It is indicated in both Figs. 6.4 and 6.15 that the average wind direction does not vary substantially on an hourly scale. Thus if the initial and boundary conditions of the simulations render initial wind directions close to the measured direction, we can expect not too large changes over a 2-h period and hence that measurements and predictions become close. An indication of this is that all initial STDEs are close to STDEs of the last data-windows in Fig. 6.15. However, more data are required to verify exactly how much better the direction can be predicted with the detailed CReSS simulations.

6.5.4 Spectral coherence

Here, we finally we evaluate spectral coherence of wind speeds for 1-Hz data at three distances d between any two WTs in the WF: $d_1 \approx 3000 \pm 200$ m, $d_2 \approx 1650 \pm 150$ m, and $d_3 \approx 500$ m. There are thirteen possible combinations of WTs in the WF when d_1 occurs, five for d_2 , and nine for d_3 . In total there are 14 occasions for which we have synchronized measurements for d_1 ; three for d_2 ; and none for d_3 . For CReSS there

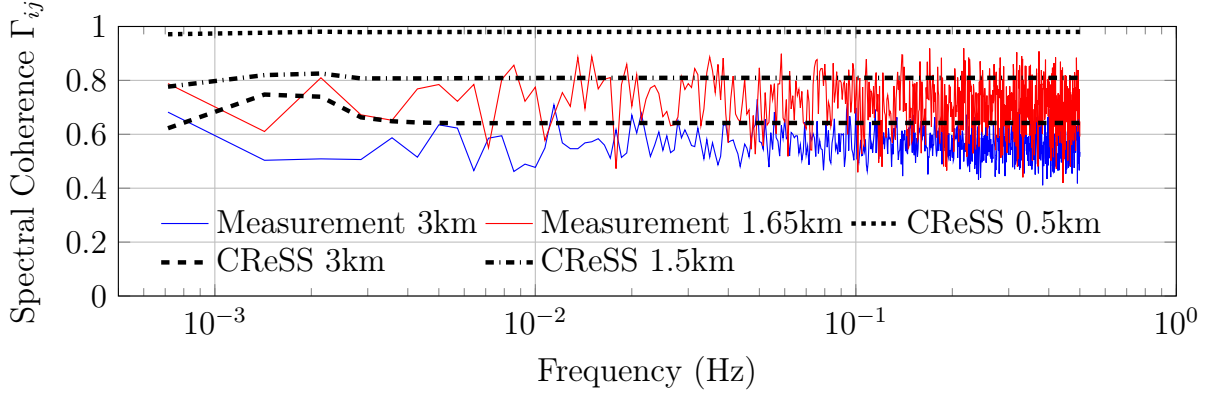


Figure 6.16: Averaged spectral coherence between wind speeds at the locations of WTs for measurements and predictions with CReSS.

are 39 (3 dates $\times$ 13 cases) for d_1 ; 15 (3 dates $\times$ 5 cases) for d_2 , and 27 for d_3 , since data at all WT-positions are synchronously available. Spectral coherence is calculated individually for each pair of time-series, and the results on averaged coherence are shown in Fig. 6.16. From these results, it is clear that the predicted spectral coherence is slightly higher than for the measurements. The slightly higher coherence observed for CReSS is possibly due to spatial and temporal smoothing effects in the simulation system. When uncorrelated turbulence is added to the CReSS data, the coherence would decrease as well. The difference in the averaged value of the coherence between d_2 and d_1 becomes 0.1438 for measurements and 0.1670 for CReSS. This is an indication that the difference in coherence is not only due to uncorrelated turbulence, but also due to coherence at higher frequencies which are not captured by CReSS. NWP models are generally able to capture atmospheric features at the scale of about five times the horizontal resolution [124] which would imply 1 km for these CReSS simulations. The result for the d_3 -case (500 m) which shows a coherence of close to one is thus an expected result since wind speeds at locations separated about two-three times the horizontal resolution are likely to be very close, particularly in the absence of high-frequency turbulence.

6.6 Conclusions

Wind predictions from CReSS were used to evaluate wind and WP at an offshore WF in Ibaraki Prefecture, Japan, where WT measurements are available. Two different data sets from CReSS were used: daily prediction data (up to 30 h) of 2 km spatial resolution and 1 h temporal resolution, and extremely detailed prediction data with a 200-m spatial resolution and 1 s time resolution. Predicted wind speeds, output powers (using a power curve), wind directions, and spectral coherence between wind speeds at different WT locations were compared against real measurements.

It was shown that the spectra of both measured and predicted wind speeds scale ac-

cording to Kolmogorov’s spectrum. In particular, we show that the spectra of 1-h wind speeds become very close, which demonstrates CReSS’s capability of capturing realistic fluctuations at this time-scale. Wind direction is predicted accurately (with STDE of roughly 20°) both in the longer term for coarser simulations and for extremely detailed simulations with predictions up to a few hours. The wind direction is predicted more accurately for higher wind speeds—because the wind is chiefly governed by dominant mesoscale systems in such cases, which CReSS can capture well. In the absence of a dominant mesoscale system, other atmospheric motions on scales larger than turbulence become dominant [156], resulting in an increased variability in direction. Comparing to other studies [123,126,128,130], wind speed and consequently WP are seemingly predicted with slightly lower accuracy than the direction, but the predictions show reductions in RMSE with temporal averaging. Moreover, the accuracy of the predictions is shown to be affected by the direction, which can be attributed to the coarse terrain model. The distributions of wind speed predictions follow the Weibull distribution closely. Wind speeds are measured at the nacelles of WTs and are statistically corrected by a linear model. Nacelle measurements inherently contain errors, but as [129] discovered, they might be closer to NWP than met mast measurements. The long-term energy production is found to be very well estimated by CReSS. The PDF of predicted power increments becomes closer to measurements for larger increments. In particular, the results become close for 4-h increments or larger for averaged power. These results suggest CReSS’s potential to predict WP ramps on an hourly scale. For the 1-Hz data, spectral coherence was obtained slightly higher for wind predictions than measurements. The discrepancy in spectral coherence is likely due to the lack of high amplitude, high frequency components in the simulation data which in the real data contain uncorrelated and correlated fluctuations affecting the coherence depending on the distance between the measurement locations.

This evaluation demonstrated that WP fluctuations on an hourly scale (based on our results on distribution of power increments, prediction of increments, and the wind spectrum) are well captured and predicted by CReSS, and that spectral coherency is achieved similar to real measurements on the scale of one to a few kilometers. In addition to the evaluation of single WF sites, our results suggest the potential and promise of using CReSS to analyze large-scale smoothing effects of WP. Such analyzes are of vital importance in the face of a potential large-scale expansion of offshore WP in Japan. This is investigated in the next chapter.

Chapter 7

Quantifying smoothing effects of wind power via KMD

7.1 Introduction

The continued increase of wind and solar power generation worldwide makes it necessary to understand the characteristics of aggregated renewable generation, and in particular the reduced variability or *smoothing effects* achieved by distributing the generation across a certain area. This dissertation focuses on wind power generation due to its unexploited potential, in particular for offshore installation, for which a rapid increase is expected within the next few years. In Japan, the share of wind power in the power generation mix is currently quite small in comparison with many other nations [106], but is expected to increase in the coming years [43]. Since wind is an intermittent and uncontrollable source of energy, and only a limited amount of electrical energy can practically be stored [55], one way of reducing variability in the aggregated power is to distribute the generation across a large area.

This chapter looks at the smoothing characteristics of wind power, which are determined by the correlation between Wind Farm (WF) outputs at various time-scales. A better understanding of the smoothing characteristics, e.g. understanding the size of the WF distribution area required to achieve a certain degree of smoothing in the aggregated wind power, would contribute to smarter planning and better utilization of equipment and natural energy resources, and greater reliability [157]. A deliberate distribution of renewable generation would hence benefit the development towards cleaner power systems, by facilitating the introduction of large-scale alternative renewables into the systems.

7.1.1 Literature review

In recent decades, smoothing of aggregated wind power has been addressed in numerous studies. Predominantly, various correlation analyses have been carried out, and frequency domain characteristics have been investigated via Power Spectral Density (PSD):

see e.g. [66]. It was shown in [158] how the PSD of the total WF output is estimated via the PSD of the output of a single turbine, and the proposed method showed a good agreement with measurements. Variability reduction in the aggregated power on an hourly scale was demonstrated in [159] for wind power dispersed over a large area (about 50×100 km) in Germany via second-order statistics and PSD. The reduction of variability in the aggregated power for WFs dispersed over a large area in Europe was also discussed in [160]. Based on a statistical analysis, it was concluded in [157] that the correlations between aggregated wind power in different geographical areas are similar to the correlations between WF outputs. In [161], wind power fluctuations were investigated by incorporating weather simulation data. The results showed that the time-scales at which smoothing effects are achieved vary considerably between different geographical regions, and the authors concluded that regional wind characteristics need to be assessed in detail to estimate the smoothing effects of inter-regional wind power. In [27], correlations of hourly power variations in the Nordic countries (Sweden, Norway, Denmark, and Finland) were studied on distances up to 2000 km. The results showed that the hourly variations in aggregated wind power in each country stayed within $\pm 5\%$ of the installed capacity with a probability of 91%–94%. It was shown in [150] that the PSD of wind power scales according to the so-called Kolmogorov spectrum as $S(f) \sim f^{-5/3}$ as discussed in Chapter 6. In [162], this characteristic was exploited by looking at the deviation from the Kolmogorov spectrum in the aggregated power as a measure of smoothing (since the non-aggregated power has been shown to follow it). The authors of [162] demonstrated an 87% reduction in wind power variability by interconnecting four WFs, and that interconnecting 16 additional WFs only gave 8% additional reduction. In [135, 149], smoothing effects in a hypothetical WF and coherence between distant turbines were investigated via PSD with real wind speeds measurements in Japan. A new index to quantify smoothing effects of wind power was proposed in [149], and the results from applying it showed non-significant smoothing effects for fluctuations on the scale of a couple of hours, for locations separated hundreds of kilometers apart. This result implies that wind power fluctuations on the scale of about one hour and faster are well mitigated by geographically distributing the wind power generation.

7.1.2 Proposed idea

Here, two different smoothing indexes of wind power are proposed based on spectral analysis via the so-called Koopman Mode Decomposition (KMD). KMD is applied to wind power time-series, and decomposes the wind power data into a finite number of eigenvalues and *modes*. Modes are here defined as complex valued vectors of the same dimension as the number of observables, each oscillating with a particular frequency for all observables (here wind powers at hypothetical WFs), determined by the eigenvalue. In particular, by utilizing the phase and amplitude information of each observable which characterize the participation in each mode, an indexed quantifying the smoothing via the averaged coherence in KMs is proposed. The proposed index is an extension of a previously proposed method to quantify wind power smoothing in Japan [135, 149] via

PSDs. The performance of the proposed index is exemplified by incorporating large-scale weather simulation data from the Cloud Resolving Storm Simulator (CReSS) [63] to study practical smoothing effects of wind power in Japan. CReSS was evaluated against real measurements from a WF in Chapter 6. The utilized CReSS data in this chapter have 1-h and 1-min temporal resolutions with 2 km spatial resolution. Smoothing has not extensively been investigated on a nationwide scale in Japan except for in [135, 149]. Because correlation of wind power has shown to vary significantly between geographical regions [160, 161], it is important to study every region specifically to acquire accurate knowledge necessary for planning purposes.

7.1.3 Outline

The remainder of the chapter is organized as follows. Section 7.2 introduces the concept of wind power smoothing by simple statistics. Section 7.3 describes a smoothing index calculated via PSDs, used as a comparison to the KMD-based index presented in this chapter. Section 7.4 explains KMD briefly and derives the proposed KMD-based smoothing index. In Section 7.5, we elaborate on the practical range of the smoothing index, and compare the KMD- and PSD-based indexes via simple numerical examples. Section 7.6 first briefly describes the weather simulation data from CReSS, and then presents the experimental results by applying the proposed index to data on wind power in Japan. Finally, conclusions of this study are given in Section 7.7.

7.2 Statistical quantification of smoothing

The concept of wind power smoothing and its quantification by using simple statistics is explained in the following [66]. Let us consider a WF with m number of WTs. Each WT has the *mean* output $P_{\text{turb}} = \langle P(t) \rangle$ for some period of time, e.g. 10 minutes, and a standard deviation $\sigma_{P_{\text{turb}}}$. The average of the total WF output is thus $P_{\text{WF}} = mP_{\text{turb}}$. Now, consider the the following two cases:

- (i) No smoothing effects, i.e. all WTs experience the exact same wind. The standard deviation of the WF output thus becomes $\sigma_{\text{WF}} = m\sigma_{P_{\text{turb}}}$, because all outputs are perfectly correlated.
- (ii) Uncorrelated and normally distributed turbulence. This would result in a reduction of variability in the WF output, and more precisely that $\sigma_{\text{WF}} = m\sigma_{P_{\text{turb}}}/\sqrt{m}$.

We can thus understand that smoothing of the WF output depends on the degree of *correlation* between the individual outputs constituting the total power. In the case of uncorrelated turbulence, the standard deviation becomes $\sigma_{\text{WF}}^n(m) = \sigma_{P_{\text{turb}}}/\sqrt{m}$ (n denoting *normalized*) in per unit (p.u.) of the total WF capacity. In Fig. 7.1, it is shown how the standard deviation of the WF output decreases with the number of WTs, according to both the theory and a numerical test (by summing m number of time-series

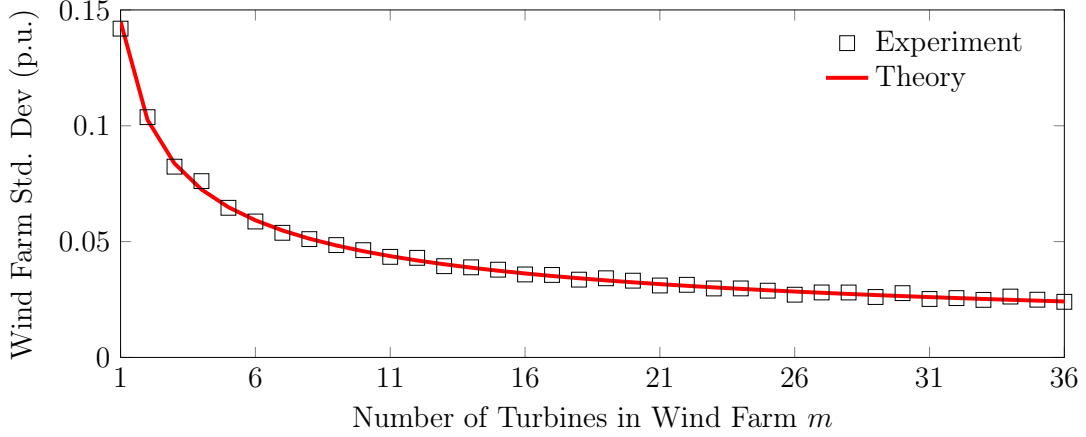


Figure 7.1: Wind farm standard deviation in p.u. for a farm with m turbines with the same mean output and uncorrelated turbulence. ‘Theory’ corresponds to evaluation with (7.1), and ‘Experiment’ to numerical experiments of adding uncorrelated Gaussian noise.

of normally distributed random numbers) verifying the theory. Following [160], we write the standard deviation of the sum of m time-series as

$$\sigma_{\text{WF}}^2 = \frac{1}{m^2} \sum_{i=1}^m \sum_{j=1}^m \sigma_i \sigma_j r_{ij}, \quad (7.1)$$

where r_{ij} is the cross-correlation coefficient between the i -th and j -th time-series, and σ_k the standard deviation of the k -th time-series. In the case (ii) of uncorrelated turbulence we have that $r_{ij} = 0$ for all $i \neq j$, and according to the definition of r we have $r_{ij} = 1$ for all $i = j$. Then $\sigma_{\text{WF}} = \sigma_i / \sqrt{m}$, i.e., same as previously for a WF with m WTs with uncorrelated turbulence. The standard deviation is the most simple way of quantifying the reduction of variability, and r_{ij} provide a simple way of quantifying the correlations between outputs, which are related to the standard deviation of the summed time-series via (7.1). However, these methods do not reveal the time-scales at which smoothing is achieved (if the data have not band-passed filtered to only contain fluctuations of a specific frequency). Therefore, smoothing effects of wind power have often been studied in the frequency domain by Fourier based methods, briefly discussed in the next section.

7.3 Quantification of smoothing via PSD

Smoothing of wind power in the frequency domain has conventionally been described by means of PSD [66, 135, 159]. Here, we are interested in quantifying the amount of smoothing achieved at a WF or in the aggregated power of multiple WFs. Now, let us denote the output power of a turbine or farm i by P_i and its PSD by $S_i(f)$. The aggregated power at a time t becomes $P_{\text{tot}}(t) = \sum_{i=1}^m P_i(t)$, the normalized power $P_{\text{tot}}^n(t) = P_{\text{tot}}(t)/m$ (n again stands for *normalized*, and m is the number of WFs/turbines). Their PSDs are

denoted by $S_{\text{tot}}(f)$ and $S_{\text{tot}}^n(f)$, respectively. In [163], the following smoothing index was considered:

$$s_i(f) := \sqrt{S_{\text{tot}}^n(f)/S_i(f)}, \quad (7.2)$$

where $S_i(f)$ represents the PSD of “a typical” turbine/WF, which e.g. can be thought of as the turbine in a WF with a typical output. The smoothing index (7.2) is in [135] defined as the gain of a filter between the output of a turbine and the WF. In other words it is a measure of how much the total power fluctuates for a certain frequency f compared to “a typical” turbine/WF. When output powers and locations vary significantly, it becomes difficult to judge the correctness of such an index because of the difficulty determining “the typical” turbine/WF. Thus, the following different index was proposed in [163]:

$$s(f) := \text{tmean}_i(s_i(f)), \quad (7.3)$$

where **tmean** represents the 25% truncated mean over all indexes $i = 1, \dots, m$, to disregard the results of the most extreme spectra (outliers). Here, this index is called the *mean smoothing index via PSDs*, since it involves the truncated mean operation.

However, this chapter will focus on another index as a comparison to the KMD-based index that will be proposed later on. First, the total power S_{tot} is given in [149] as follows:

$$S_{\text{tot}}(f) = \sum_{i=1}^m S_i(f) + \sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i}}^m \sqrt{S_i(f)S_j(f)} \cdot \gamma_{ij}(f) \cos(\phi_{ij}(f)), \quad (7.4)$$

where $\gamma_{ij}(f) = |S_{ij}(f)|/\sqrt{S_i(f)S_j(f)}$ is the spectral coherence magnitude, and $\phi_{ij}(f) = \arg(S_{ij}(f))$ the phase between i and j , where $S_{ij}(f)$ is the cross-spectral density between P_i and P_j . Thus, it can be understood that the amount of smoothing is highly dependent on the coherency, including the phase differences between individual powers constituting the aggregated power. Then, the following smoothing index is defined as in [149]:

$$\text{coh}_{\text{PSD}}(f) = \frac{1}{m(m-1)} \sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i}}^m \gamma_{ij}(f) \cos(\phi_{ij}(f)), \quad (7.5)$$

where it has been assumed that the PSDs of power at different locations are *identical*: see [149] for the detailed derivation. Here, this index is called the *averaged smoothing index via PSDs*, since it sums the coherence between any two locations and then takes the average. With this index as a starting-point, a new index is proposed in the next section.

7.4 Smoothing via Koopman modes

Starting from the definition of KMD, which was given and explained in detail in Chapters 2 and 3, we introduce two smoothing indexes, where the one called the *averaged smoothing index via KMD* is used to evaluate smoothing effects of wind power in northern Japan.

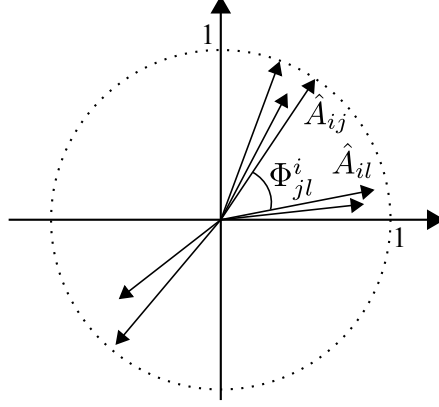


Figure 7.2: Illustration of amplitudes and phases of Koopman Modes (KM).

7.4.1 New smoothing indexes via KMD

Consider $N + 1$ vector-valued temporal snapshots of wind power measurements in p.u. (power/(rated power)) collected at m locations: $\{\mathbf{P}_0, \dots, \mathbf{P}_N\}$, $\mathbf{P}_k \in \mathbb{R}^m$, where k denotes a specific time-instance. Assuming that the residual $\mathbf{r}$ is negligible in (2.8), the data can be decomposed into a finite sum via Arnoldi-like KMD according to (2.8):

$$\mathbf{P}_k = \sum_{i=1}^N \tilde{\lambda}_i^k \tilde{\mathbf{v}}_i, \quad k = 0, \dots, N. \quad (7.6)$$

Modal frequencies are calculated according to $f_i = \text{Im}(\ln(\tilde{\lambda}_i))/(2\pi T_s)$, where T_s is the sampling period. The vector $\tilde{\mathbf{v}}_i = \mathbf{A}_i \angle \boldsymbol{\alpha}_i := [A_{i1} \angle \alpha_{i1}, A_{i2} \angle \alpha_{i2}, \dots, A_{im} \angle \alpha_{im}]^\top$ is here again called the *Koopman Mode* (KM) oscillating with the frequency f_i and contains the magnitudes A_{ij} and phases α_{ij} of power fluctuations corresponding to the m measurements; e.g. outputs of m WFs. To identify lightly damped or undamped oscillations with large magnitude, all N KMs are sorted by $(\tilde{\lambda}_i)^N \|\tilde{\mathbf{v}}_i\|$, where $\|\cdot\|$ denotes the euclidean norm, and higher ranked ones are called *dominant* KMs, which are here used to evaluate smoothing effects. The total (aggregated) power $P_{\text{tot},k}$ can be expressed using (7.6) by

$$P_{\text{tot},k} = \mathbf{P}_k^\top \mathbf{1}_m = \sum_{i=1}^N \tilde{\lambda}_i^k \sum_{j=1}^m [\tilde{\mathbf{v}}_i]_j =: \sum_{i=1}^N \tilde{\lambda}_i^k \bar{\mathbf{v}}_i, \quad (7.7)$$

where $\mathbf{P}_k \in \mathbb{R}^m$ contains the measured powers at time k , $\mathbf{1}_m$ is the m -length vector of ones, $[\tilde{\mathbf{v}}_i]_j$ is the j -th component of $\tilde{\mathbf{v}}_i$, and $\bar{\mathbf{v}}_i \in \mathbb{C}$ is the scalar KM of the total power. That is, a spectral decomposition of the total output is achieved by applying KMD to individual outputs. Now, let $\bar{A}_i = |\bar{\mathbf{v}}_i|$ be the amplitude factor of the total power for the i -th KM, and analogously A_{ij} be the factor for the j -th output power. Then, we define

$$s_i := \frac{\bar{A}_i}{m \cdot \text{tmean}(\mathbf{A}_i)}, \quad (7.8)$$

as a smoothing index, where $\mathbf{A}_i := [A_{i1}, \dots, A_{im}]^\top$, and $\mathbf{tmean}$ is same as in (7.3). Note that m is included in (7.8) to scale the total output to WF p.u. If the amplitude factor $\bar{A}_i$ of the total power is smaller than the truncated mean of amplitudes for individual WFs contained in $\mathbf{A}_i$, then (7.8) becomes smaller than one, implying smoothing effect for the KM with frequency f_i . We call this the *mean smoothing index via KMD*. This index was evaluated against the index (7.3) in [163]. In this chapter however, an additional KMD-based index similar to (7.5) is introduced:

$$\text{coh}_{i,\text{KMD}} = \frac{1}{m(m-1)} \sum_{j=1}^m \sum_{\substack{l=1 \\ l \neq j}}^m \hat{A}_{ij} \hat{A}_{il} \cos(\Phi_{jl}^i), \quad (7.9)$$

where $\Phi_{jl}^i := \alpha_{ij} - \alpha_{il}$, and $\hat{A}_{ij} := A_{ij}/\max(\mathbf{A}_i)$, i.e. components of $\mathbf{A}_i$ normalized by the largest component: see Fig. 7.2 for an illustration. The subscript i refers to the i -th *dominant* KM oscillating with the frequency f_i . We call this the *averaged smoothing index via KMD*, and it can be regarded as a generalization of a previously proposed index [149] which was shown to be effective for quantifying smoothing effects in Japan. The index theoretically takes a value in $[-1, 1]$, where ‘ -1 ’ indicates perfect smoothing and ‘ 1 ’ no smoothing; however, it approximately takes a value in $[0, 1]$ for the practical data that we consider here. This is because, first of all, ‘ -1 ’ is only achievable in the case of *two* signals in antiphase, but here we consider outputs at multiple locations. When the number of locations increases, the value indicating perfect smoothing approaches zero, which is shown in the following section. Secondly, perfect smoothing of wind power is achieved for a large number of outputs with uncorrelated fluctuations or turbulence [66], whereas the number is not significantly large here. For these reasons, a maximum achievable value of smoothing close to zero can be expected.

Additionally, the proposed indexes (7.8) and (7.9) are not limited to the algorithm used to approximate KMD in this chapter (Arnoldi-like KMD, introduced in Chapter 2), and can be computed using any technique producing a decomposition of the same type as (2.8). In fact, there are several different KMD algorithms proposed in literature: see e.g. [17, 18, 62, 78], and a few of them were outlined in Chapter 2. Furthermore, the index takes the same form as (7.5), which follows naturally from the notion and definition of coherency in KMs [13, 19] (see also Chapter 9), and does not rely on the assumptions used to derive (7.5). Thus, we can state that the KMD-based index (7.9) is regarded as a generalization of the PSD-based index (7.5), which is numerically demonstrated in the next section by simple examples.

7.5 Comparison between smoothing indexes

Here, we first look at properties of the formula used to calculate the averaged smoothing via both KMD and PSDs, and second look at two simple numerical examples to clarify the similarities and differences between the two methods. By dropping the index i in

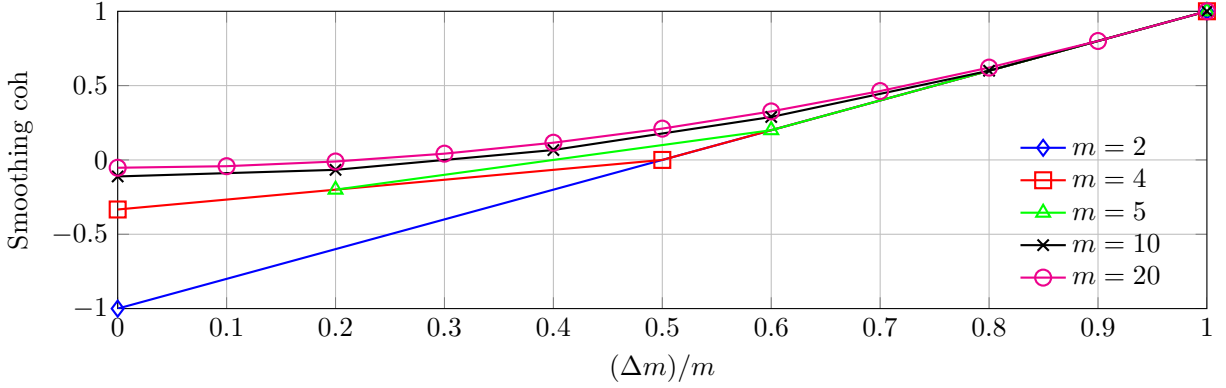


Figure 7.3: Simplified smoothing index coh according to (7.12), evaluated for two groups of phasors of unit length in anti-phase for different m and Δm .

(7.9) and f in (7.5), and only considering a single frequency for simplicity, we rewrite the average coherence on a common form:

$$\text{coh} = \frac{1}{m(m-1)} \sum_{j=1}^m \sum_{\substack{l=1 \\ l \neq j}}^m I_j I_l \cos(\theta_j - \theta_l), \quad m > 1, \quad (7.10)$$

where $I_j \angle \theta_j = \hat{A}_j \angle \alpha_j$ for KMD, or $I_j I_l \cos(\theta_j - \theta_l) = \gamma_{jl} \cos(\phi_{jl})$ for PSDs. Now, we think $I_i \angle \theta_i$ as a phasor as shown in Fig. 7.2 for KMD. The value ‘-1’ on perfect smoothing is achieved when $m = 2$, $I_j = I_l = 1$, and $(\theta_j - \theta_l) = z\pi$ ($z \in \mathbb{Z}$). Thus, we consider a case where $m = 2n$ ($n \in \mathbb{N} > 1$) and $I_i = 1$ for all i , and we let $\theta_i = \beta$ ($i = 1, \dots, m/2$) and $\theta_i = \beta \pm \pi$ ($i = m/2 + 1, \dots, m$), where β is an arbitrary angle. This case implies two groups of unit-length phasors in anti-phase, hence perfect smoothing is achieved. In this case, (7.10) is simplified as

$$\text{coh} = -\frac{m}{m(m-1)} = -\frac{1}{m-1}, \quad m > 1, \quad (7.11)$$

which goes to zero as m approaches infinity. Note that when m is not even as in the previous example (where $m = 2n$), by letting e.g. $\theta_i = \beta$ ($i = 1, \dots, \text{floor}(m/2)$) and $\theta_i = \beta \pm \pi$ ($i = \text{floor}(m/2) + 1, \dots, m$), i.e. allowing a difference of one between the number of anti- and in-phase phasors, (7.11) will closely approximate the smoothing in particular for “large” m . The `floor` function (here we use MATLAB) takes the integer part of a value: e.g. `floor(3/2) = 1`. Still, we consider two groups of phasors of unit lengths in anti-phase, denote the number of phasors in each group by m_1 and m_2 satisfying $m_1 + m_2 = m$. The “imbalance” is defined as $\Delta m = |m_1 - m_2|$. Then, (7.10) can be simplified as

$$\text{coh} = -\left(\frac{m - (\Delta m)^2}{m(m-1)}\right), \quad 0 \leq \Delta m \leq m. \quad (7.12)$$

Here, (7.12) is equivalent to (7.11) for $\Delta m = 0$ (perfect smoothing) and becomes one for $\Delta m = m$ (i.e. no smoothing for only in-phase phasors). In Fig. 7.3, we have evaluated

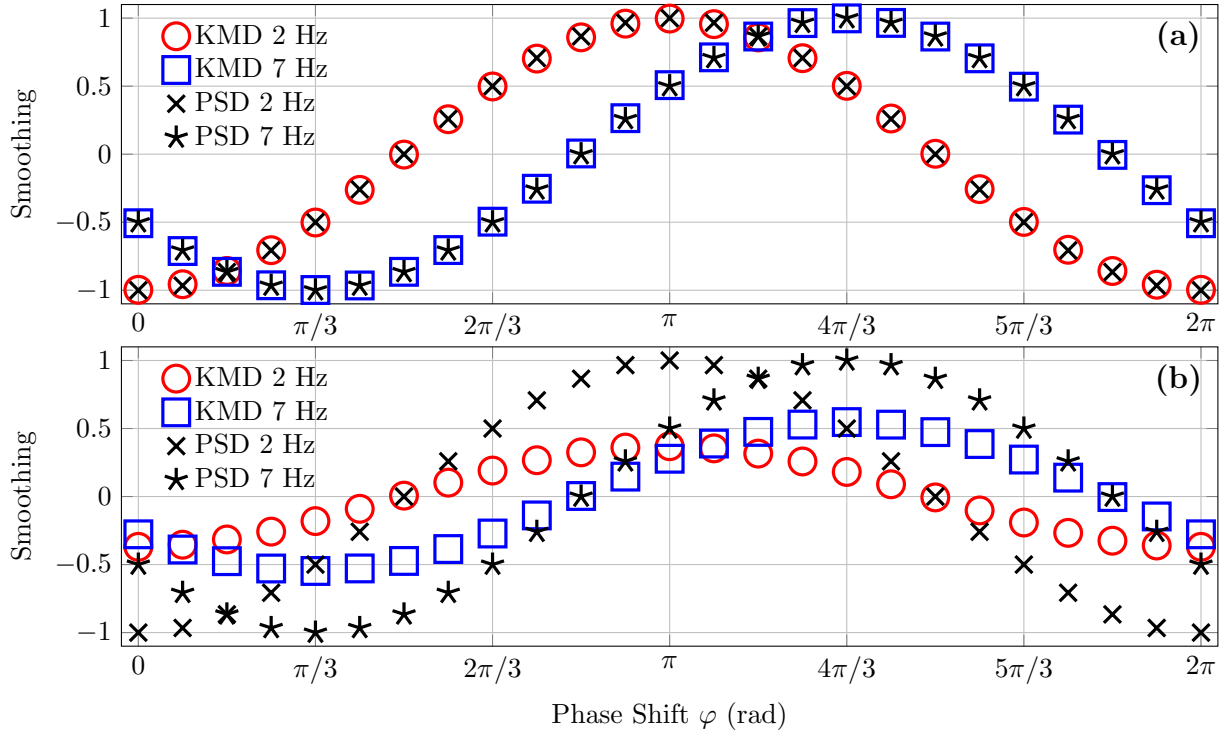


Figure 7.4: KMD- and PSD-based averaged smoothing indexes applied to (7.13) under different phase shifts φ : For (a) Case 1; (b) Case 2.

(7.12) for different m and Δm . The figure shows that for m as low as 10, values on smoothing close to zero are achieved for perfect smoothing ($\Delta m = 0$.) However, results as ideal as in the case of two anti-phase groups of phasors of equal lengths may be unrealistic for our practical data. Also, m is always larger than two for the wind power data considered in this chapter. Thus, we say that the index approximately takes a value in the interval $[0, 1]$, although small negative values are sometimes observed.

Now, we look at simple numerical examples to clarify the differences and similarities between the KMD- and PSD-based averaged smoothing indexes. Let us consider the two signals given by

$$\begin{aligned} x_1(t) &= a_1 \cdot \sin(2\pi f_1 t) + b_1 \cdot \cos(2\pi f_2 t), \\ x_2(t) &= - \left\{ a_2 \cdot \sin(2\pi f_1 t + \varphi) + b_2 \cdot \cos \left(2\pi f_2 t - \frac{\pi}{3} + \varphi \right) \right\}, \end{aligned} \quad (7.13)$$

where $f_1 = 2$ Hz and $f_2 = 7$ Hz. If $a_2 = a_1$ and $b_2 = b_1$, then perfect smoothing is achieved for 2 Hz if $\varphi = n2\pi$, and for 7 Hz if $\varphi = \pi/3 + n2\pi$, where $n \in \mathbb{N}$. Now, we vary φ between 0 and 2π , let the temporal length be 8 s ($N + 1 = 257$), and the sampling frequency $f_s = 32$ Hz. The MATLAB functions `cpsd` (cross-spectral density) and `mscohere` (magnitude squared coherence) are used to compute the PSD-based index (7.5) with no window overlap. The number of `fft` (fast Fourier transform) points is set at 128, and no window function such as Hamming is used.

Smoothing via the proposed (7.9) and the PSD-based index (7.5) is first computed for a case (Case 1) with $a_1 = a_2 = 0.8$, and $b_1 = b_2 = 0.6$. The results are shown in Fig. 7.4 (a), and we see that the results become virtually identical for the two indexes. For example, when $\varphi = 0$, perfect smoothing is achieved for 2 Hz, which corresponds to a smoothing value of -1 . In this case, the PSDs of $x_1(t)$ and $x_2(t)$ should be identical, and thus the derivation leading to the smoothing index (7.5) holds.

Now, let us consider (7.13) with $a_1 = 0.8$, $b_1 = 0.6$, $a_2 = 0.3$, and $b_2 = 1.1$ (Case 2). The smoothing results are given in Fig. 7.4 (b). The PSD-based index here gives the same result as for Case 1, but the KMD-based index generates a slightly different result. The PSD-based index fails to quantify the smoothing correctly since it indicates perfect smoothing for a case where it is obviously impossible because the amplitudes are not equal ($a_1 \neq a_2$, $b_1 \neq b_2$); giving PSDs of different magnitudes. The proposed KMD-based index thus generalizes the PSD-based index for cases where the PSDs are not equal, which can be expected to be the case for the wind power data considered here.

7.6 Test with large-scale wind prediction data

Here, we apply the proposed smoothing index (7.9) to simulated wind power data estimated with wind predictions from CReSS with 1-h and 1-min time-resolutions, using the power curve shown in Fig. 4.3 (b). The 1-h data are from 36-h weather simulations, and the 1-min data are from a 30-h weather simulation. CReSS was evaluated against real measurements and described in detail in Chapter 6.

7.6.1 Data sectioning and mean smoothing with KMD

For example, when applying the KMD-based smoothing index to 1-h sampled wind power predictions, we have a matrix of input data containing temporal snapshots collected over a month's period: $\mathbf{P}_{\text{in}} := [\mathbf{P}_1, \dots, \mathbf{P}_{N_1}] \in \mathbb{R}^{m \times N_1}$; where m denotes the number of locations and N_1 the number of temporal snapshots. Here, we attempt to look at fluctuations of wind power on an hourly rather than daily scale due to their impact on power system operation [159, 164]. Because of this, we arrange the input data into sub-matrices $\mathbf{P}_i$ containing three days of wind speed data to have a reasonable amount of temporal snapshots in each matrix, each time-shifted L (e.g. one day) from the previous one, and define the input to KMD as

$$\mathbf{P}_{\text{new}} := \begin{bmatrix} \mathbf{P}_1 \\ \vdots \\ \mathbf{P}_p \end{bmatrix} \in \mathbb{R}^{(p \cdot m) \times N_2}, \quad (7.14)$$

where N_2 is the number of temporal snapshots in each sub-matrix $\mathbf{P}_i$, and p the number of sub-matrices. Each sub-matrix $\mathbf{P}_i := [\mathbf{P}_{1+(i-1)L}, \dots, \mathbf{P}_{N_2+(i-1)L}]$ can be seen as a subsection of data, and the observables, i.e. the m measurement locations are the same only time-shifted in each section. When applying the KMD-based averaged smoothing

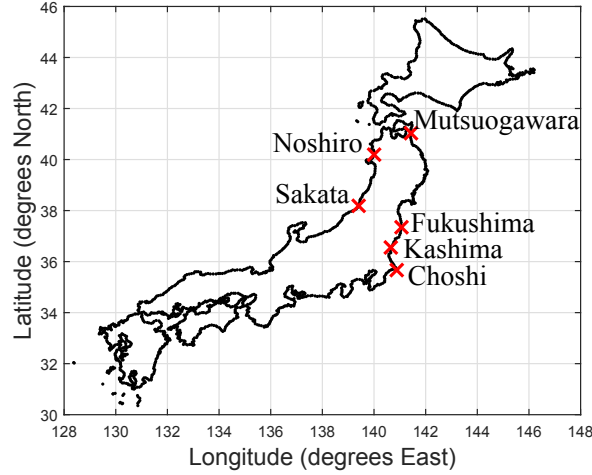


Figure 7.5: Set of wind speed measurement points outlining the coasts of Japan’s main islands, indicated by black dots. Red crosses roughly mark six regions in Northern Honshu, where there are WFs or project/plans: Mutsuogawara, Noshiro, Sakata, Fukushima, Kashima, and Choshi.

index, the smoothing is calculated for each section, and the total result is averaged over all sections. Long time-series of 1-min sampled wind power data are also treated in the same way with this type of sectioning.

7.6.2 Smoothing of aggregated power in Honshu, Japan

In this section, we demonstrate the proposed smoothing index via KMD. Two types of data are used in the demonstration. The first type is daily wind predictions with 1-h time resolution which have been obtained for a period of 3.5 years. However, there are temporal gaps in the data, and hence we will look at months where data from many consecutive days are available. Here, we use data from December 2012 (26 days), February 2015 (28 days), and July 2014 (30 days). The second type is predictions with 1-min time-resolution and temporal length of 30 h, from January 2013.

We will look at wind power around six regions in Japan where there are offshore WFs, or ongoing projects [102]. The six considered regions are shown in Fig. 7.5: Mutsuogawara, Noshiro, Sakata, Fukushima, Kashima, and Choshi. The WF outputs are approximated from wind speeds using a wind turbine power curve.

First, we look at correlations between predicted wind speeds. The decrease in wind speed and wind power correlation with distance is similar and is often modeled by exponential functions [27, 165]. Figure 7.6 shows linear correlation coefficients of wind speeds along the coast of Honshu for both 1-h and minute time-resolution data. The 1-h data are from December 2012, but the results are similar for all months tested. The correlations are computed with respect to a reference point north of Noshiro (see Fig. 7.5), located along the Sea of Japan coast, close to the northernmost point of Honshu. In Fig. 7.6, we

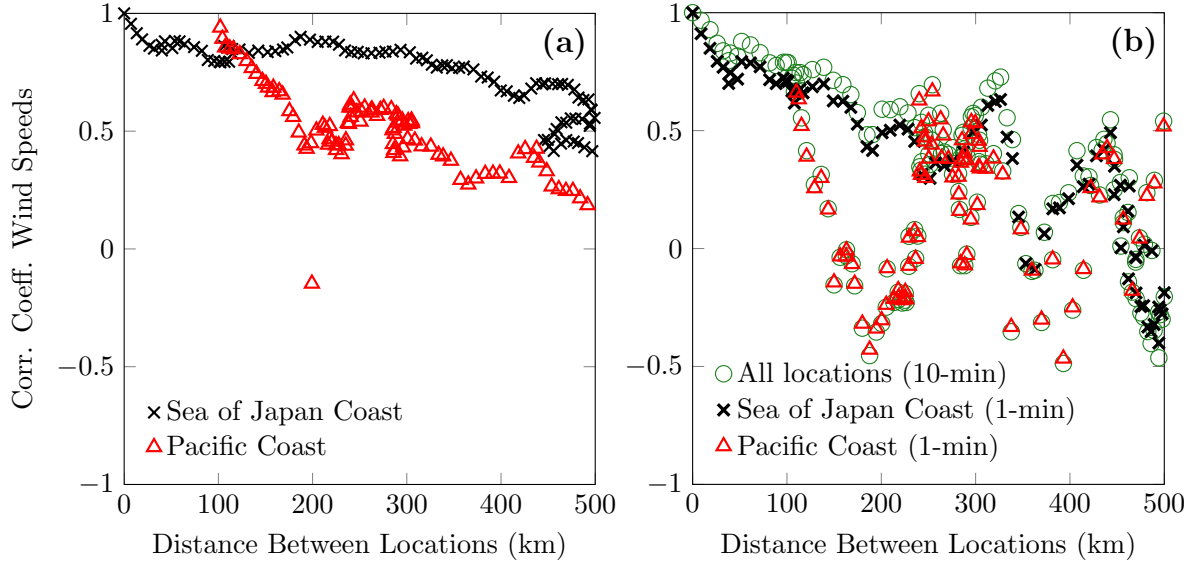


Figure 7.6: Correlation coefficients of wind speeds along the coast of Honshu for (a) 1-h wind speeds; and (b) 1-min and 10-min averaged wind speeds.

highlight locations along the Sea of Japan coast and the Pacific Coast in different colors. In the figure, significantly higher correlations are observed along the Sea of Japan coast, whereas correlations are lower for locations along the Pacific Coast. This difference in correlation is particularly apparent in Fig. 7.6 (a) for the hourly data. A distinct outlier is noted, which may be an effect of the location deep within an inlet. For correlations of minute time-resolution data shown in Fig. 7.6 (b), results for 10-min averages (from 1-min data) generally show increased correlation in comparison with 1-min data. The correlations decrease quickly towards zero and below (anti-correlation). The anti-correlations are likely an effect of the short time-period of the simulations (30 hours) and the weather characteristics on that particular day, and such effects are likely to disappear for longer time-series—since this is not observed with the hourly data which takes the macro-scale correlations on longer time-scales into account. Negative correlations are rare between WFs [157], and [160] reports slightly negative correlations for long distances (1500–3000 km). The observed phenomenon should therefore be examined with more simulation data and real measurements.

Improved smoothing by dispersing wind farms

Now, we look at smoothing effects of aggregated wind power by applying the proposed averaged smoothing index via KMD. In Fig. 7.7 (a), smoothing of aggregated wind power is shown for the December data as the number of WFs around Fukushima (see Fig. 7.5) varies from 4 to 20 in steps of 4. Initially, in Fig. 7.7 (a), the smoothing improves when the number of WFs increases from 4 to 8. However, after this, smoothing does not seem to improve as rapidly with more WFs. Now, using the 1-min data, smoothing results are shown in Fig. 7.7 (b) for input data with temporal length of 14 h, 4 h long sections,

and 2 h overlap. As for the 1-h data, a clear improvement in smoothing is observed by increasing the number of WFs from 4 to 8. Following that, only minor improvements are observed. Here, it should be noted that the 1-h and 1-min data come from different data sets with different dimensions and small differences in the geographical data, thus the positions are only roughly the same. The maximum distances between any two WFs for the cases of 4, 8, and 12 WFs for the 1-h data are 26, 62, and 94 km, and for the 1-min data they are 30, 66, and 105 km. The clear improvement in smoothing from 4 to 8 WFs mainly comes from doubling the maximum distance, causing less coherence between WF outputs.

In Figs. 7.7 (c) and (d), smoothing results for both 1-h and 1-min data are shown for the cases of placing one or two WFs in each of the six regions. The results are compared to the cases of 6 or 12 WFs around Fukushima. Figures 7.7 (c) and (d) clearly show that the smoothing improves substantially when distributing WFs around the six regions in comparison with just one (Fukushima). The smoothing in the cases of 6 and 12 WFs become very similar when they are distributed equally between the regions. However, the smoothing differs more for 6 or 12 WFs around only Fukushima. Intuitively, we can expect this, because increasing the maximum distance between WFs from roughly 50 km for 6 WFs around Fukushima to 100 km for 12 WFs, i.e. a twofold increase, should improve the smoothing due to the decrease in correlation of wind speeds/powers. Nonetheless, even when increasing the number of WFs to 30 (maximum distance 210 km), the smoothing effects are not as good (but almost) as when utilizing six regions. Also, the addition of a second WF in each region a few kilometers away from the first one, did not have as pronounced effect when looking at this temporal (1 h and 1 min) and spatial (2 km) resolutions. Thus, the main origin of the observed improvement of smoothing observed is the vast distribution of WFs across northern Japan, which substantially decreases the correlation/coherence of wind power.

Smoothing dependence on the number and choice of regions

Here, we look at how smoothing depends on the number of geographical regions chosen among the ones indicated in Fig. 7.5, by looking at all combinations of l regions of the set including the $n = 6$ considered regions. That is, we consider in total $\sum_{l=1}^n n!/((n-l)!l!) = 63$ cases. The total number of WFs is fixed at 12. For example, in the case of $l = 5$, there will be 2 WFs in four regions and 4 WFs in one region. In all other cases, the 12 WFs will be evenly distributed between the l regions. To quantify the smoothing in each case c , we consider the smoothing integrated over frequency: $\Sigma_c := \int_{f=0}^{f_\infty} s_c(f) df$, calculated with the trapezoidal method, where $s_c(f)$ ($f \in [0, f_{\infty, c}]$) represents the calculated smoothing from the KMD-based averaged smoothing index, and $f_\infty := \min([f_{\infty, 1}, \dots, f_{\infty, 63}])$. Then, we look at the normalized smoothing defined by $\hat{\Sigma}_c := \Sigma_c / \max([\Sigma_1, \dots, \Sigma_{63}])$, where the maximum Σ_c is naturally achieved for a case where $l = 1$ (one region), where we anticipate less smoothing since the correlation of wind powers within a region is likely higher than between regions.

Results from calculating this “normalized integrated smoothing” plotted vs. the num-

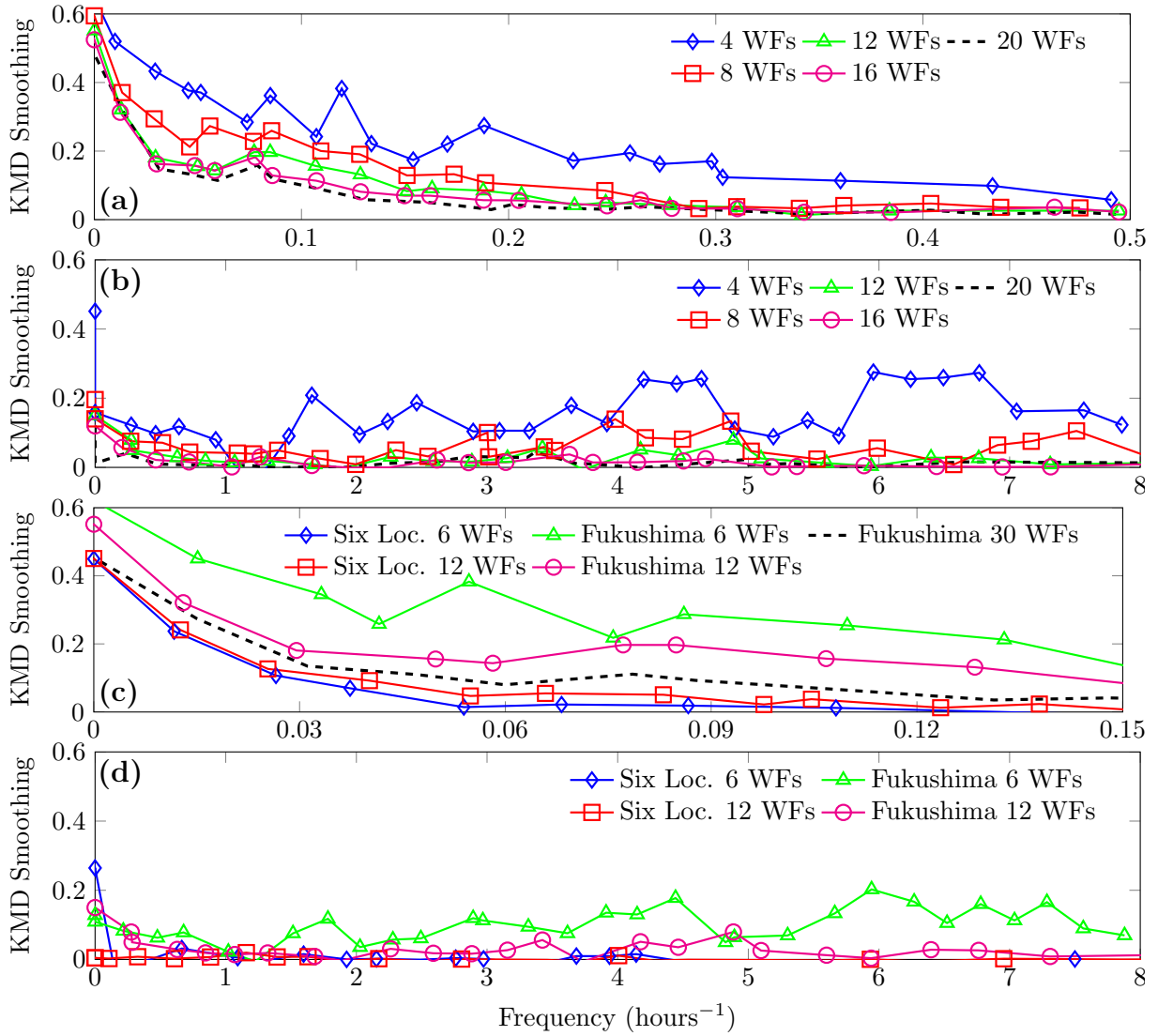


Figure 7.7: Smoothing around Fukushima as the number of WFs vary from 4 to 24 for (a) 1-h wind prediction data from December 2012 as input and (b) 1-min data. Smoothing for 6 or 12 WFs in Fukushima and WFs distributed over the six regions (locations) shown in Fig. 7.5, for (c) 1-h wind prediction data from December 2012 as input and (d) 1-min data.

ber of regions are shown in Fig. 7.8(a) for 1-h data of February, July, and December. Here, smaller improvements in smoothing are observed for a larger number of regions l , while the smoothing varies more for $l = 1, 2$, and 3. These results confirm the intuitive notion that the smoothing will depend highly on the particular regional characteristics when looking at only one or a couple of regions, while these characteristics tend to cancel out when looking at multiple regions. Note that the results for all months become close, which is clearly seen by looking at the mean results. The smoothing vs. maximum dis-

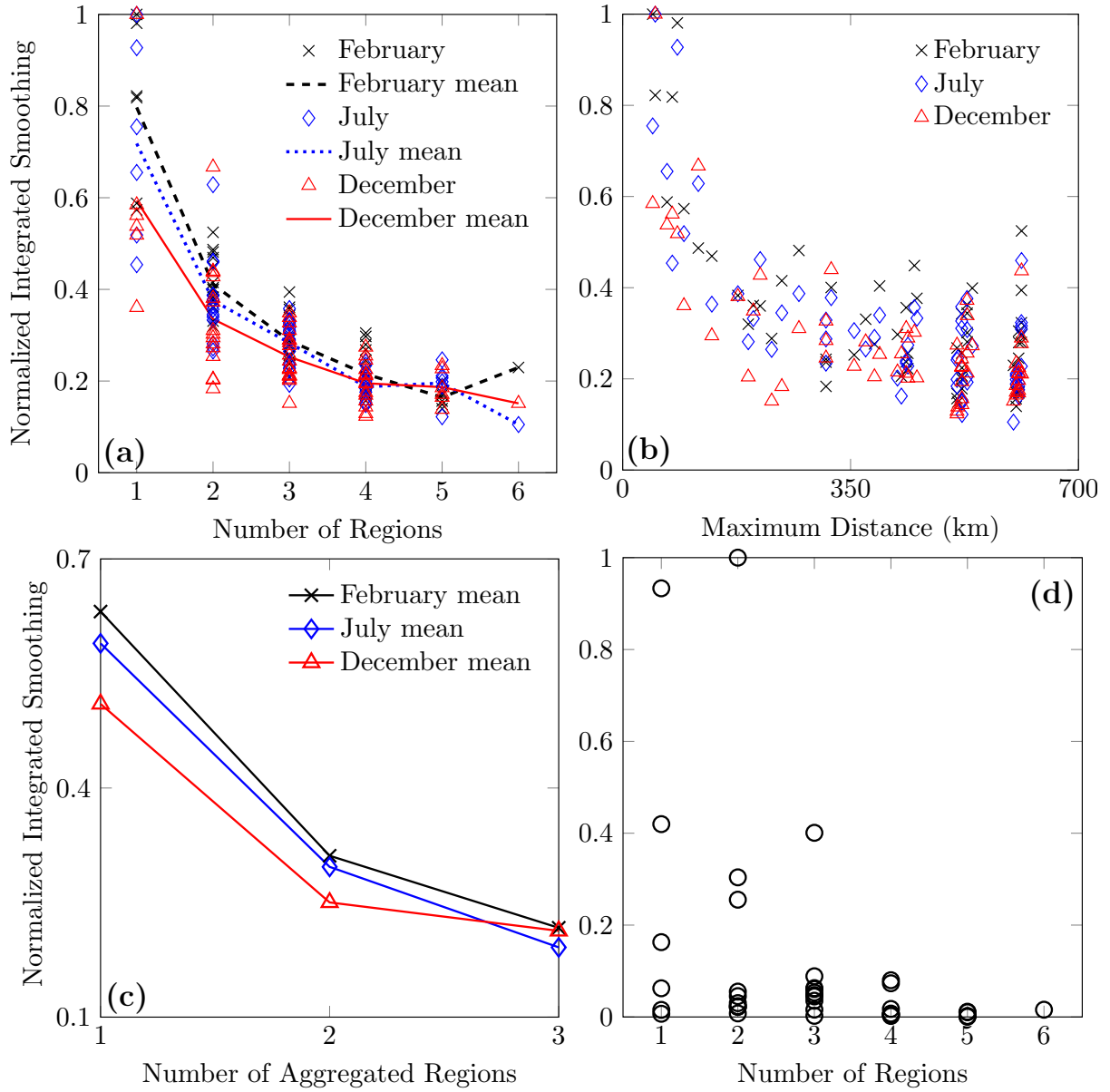


Figure 7.8: Normalized integrated smoothing vs.: (a) the number of regions for 1-h data, and the mean results; (b) distance; (c) aggregated regions where the results have been averaged for each region; (d) number of regions for 1-min data.

tance between WFs is shown in Fig. 7.8 (b). A clear, exponential decrease in smoothing is observed with increased distance for all months. Exponential decrease in correlation with increased distance has been shown in numerous studies—see e.g. [157]—which is a likely cause of this. Our results indicate that increasing the maximum distance up to about 200 km improves the smoothing significantly, whereas further increments provide less noticeable improvements. Here, it should be noted that when the distance reaches about 200 km, smoothing results for inter-regional distributions are included, such as

WFs both at the west (Sea of Japan) and east (Pacific Ocean) coasts of Japan, whereas distances below about 200 km only include one region, or multiple “close” regions such as Kashima and Choshi. To more clearly illustrate the improvements in smoothing for inter-regional distribution of WFs, we reduce the number of regions to three very rough (aggregated) regions as follows: Mutsuogawara by itself in the north, Noshiro and Sakata along the Sea of Japan coast, and Fukushima, Kashima, and Choshi along the Pacific Coast. As an example, if all WFs are located around Noshiro and Sakata, the number of included aggregated regions only becomes one. In Fig. 7.8(c) we have plotted the mean smoothing results (using the same data as in Fig. 7.8(a)) vs. the number of aggregated regions. The results clearly indicate similar improvements for all months by dispersing the generation inter-regionally over at least two regions. The cause of this can be attributed to the different correlation characteristics that were discussed in Section 7.6.2.

The smoothing results for the 1-min data are shown in Fig. 7.8(d). These results differ significantly from the results using the 1-h wind data, in that much larger smoothing effects are observed overall. In fact, most of the cases yield smoothing close to zero, which indicates almost perfect smoothing. The result suggests that the distance between WFs in one or two regions only is enough to make the wind powers uncorrelated at this time-scale. The rapid decrease in correlation for smaller spatial and temporal scales was demonstrated in Fig. 7.6, and has been shown in e.g. [27].

7.7 Conclusions

In this chapter, a new smoothing index of wind power based on the KMD was proposed, which decomposes the complex spatio-temporal data on wind power into modes oscillating with single frequencies. It was shown that the proposed smoothing index is regarded as a generalization of a previously proposed index based on power spectral densities. The results from applying the proposed index to wind predictions from CReSS indicated that by distributing wind power across only one to three regions in Northern Japan, smoothing can vary considerably; and, as the number of regions increases, the particular choice of regions matters less. In particular, our results demonstrated that a distribution of WFs between the Sea of Japan coast and the Pacific Coast results in greater improvements in smoothing than WFs placed along the same coast. This highlights the practical importance of deliberately selecting regions for large-scale wind power production to more effectively smooth the aggregated power. Furthermore, an exponential-like improvement in smoothing with increasing distance was observed.

Several direct extensions of this work are possible. This study mainly looked at 1-h data, while data of higher temporal resolution may be highly interesting for studies on short-term power system operation. Applying the proposed index to real measured data is a future work of great practical importance. In Chapter 6, the applicability of the CReSS was briefly demonstrated for wind power analysis in Japan, but many additional studies incorporating CReSS-data are possible; to name a few: wind power prediction as done in the previous chapter, economic dispatch with high penetration of wind power, and detailed case studies to evaluate potential WF locations.

Chapter 8

Predicting wind speeds via KMD

8.1 Introduction

The purpose of this chapter is two-fold. First, to investigate the application of Koopman Mode Decomposition (KMD) and related algorithms as tools for *very short-term* wind speed prediction, which has not been investigated before. Second, to highlight and explain connections between KMD/Koopman theory and commonly used prediction methods. Note that short-term or very short-term forecasting of wind speed and wind power often refers to prediction on a time-scale of minutes to hours [124, 166, 167], and is closely related to power system operation and short-term planning. Here, we are interested in forecasting wind speeds on the time-scale of *seconds*, which is related to Wind Turbine (WT) operation and control [168–170]. For example, turbulent gusts pose significant danger to the operation of WTs, and prediction of such gusts would allow WTs to execute controls to minimize the detrimental effects on equipment caused by sudden wind speed changes [168]. Hence, there is a high economic incentive to perform very short-term prediction on the time-scale of seconds to protect WTs from damage and increase their longevity. Also, in [169], by using a nonlinear model, it was shown how prediction of two variables: wind speed and direction, can increase the output power of WTs. This was achieved by using the wind direction predictions to properly align the WT and the wind direction, resulting in improved energy capture. Here, similar to [169], the potential utility of the wind speed predictions is to achieve an increased output of WTs. First and most importantly, however, this chapter focuses on the capabilities of KMD-related algorithms for very short-term wind speed prediction, and clarifies how Koopman operator theory and associated methods are related to commonly used prediction methods.

In the following, a literature review on this topic is presented. In [168, 171] it was noted that a so-called *persistence* or *naive predictor*, which assumes the future value $\hat{y}_{i+\tau}$ of a measured quantity to be equal to its last observed value y_i , is actually a quite good predictor of the future wind speed, in comparison to a nonlinear method based on *nearest neighbors*, when using standard statistical performance metrics (root mean square of errors etc.). Note that τ indicates a fixed time, or number of samples, ahead of present time

or sample i . However, “naive predictions” are by definition useless for predicting sudden gusts, since the *increment*, or difference between the future and present value is always zero: i.e. $\hat{y}_{i+\tau} - y_i := 0$. Thus, more sophisticated models or methods are evidently needed. To construct an appropriate model or prediction scheme, knowledge of the underlying dynamics giving rise to observed behavior is needed. There are many different tests for determinism in time-series data: see e.g. [4, 172, 173], which provide clues for choosing an appropriate prediction method. Boundary layer wind speeds have been shown to be non-stationary and turbulent [26, 112], and wind speed time-series exhibit nonlinear features, justifying the prediction by nonlinear models [169]. However, the proposed nonlinear model in [169] showed only minor improvements compared to a linear model. These predictions, however, allowed for a significant improvement in the power output of the WT. In [174] it was shown that locally constant predictors in the space of time-delayed coordinates [175] can be used to capture nonlinear deterministic characteristics in wind surface data. Despite the predictor not showing significant improvements in the overall prediction error, it was able to predict sudden gusts with significantly higher accuracy than the naive predictor. In [170], an artificial neural network was combined with a Markov chain to predict wind speeds on the time-scale of seconds. The method’s superiority compared to prediction with only an artificial neural network was demonstrated. For weakly turbulent data, it was shown that improvements in prediction can be achieved by using a locally linear prediction scheme instead of stochastic models [4].

The remainder of this chapter is organized as follows. Section 8.2 explains the motivation behind this study, and states two important questions that arise from the problem formulation. Section 8.3 investigates wind speed characteristics by applying a test of determinism [173] to the measured wind speeds. Section 8.4 explains and outlines a number of prediction methods, including the proposal of two new KMD-related methods, and a couple of methods for comparison. It also discusses and clarifies connections between Koopman operator theory/KMD and two common prediction methods. Section 8.5 presents and discusses the prediction results. Finally, conclusions are given in Section 8.6.

8.2 Motivation and problem formulation

The most important feature of a WT is to extract as much power as possible from the wind and to feed this power into the grid. To do so, MPPT (Maximum Power Point Tracking) [65, 176] techniques have been developed whose purpose is to control the rotor speed ω_r (rad/s) of the WT as close to its optimal value as possible. Such a technique is described and discussed in Chapters 4 and 5. The optimal value changes depending on the wind speed v_w (m/s), according to the Tip Speed Ratio (TSR)

$$\lambda := \frac{\omega_r r}{v_w}, \quad (8.1)$$

where r (m) denotes the radius of the WT. For all WTs, depending on their aerodynamic characteristics, there is an optimal value of λ , denoted by λ_{opt} : see e.g. Fig. 4.2.

However, the wind speed changes quickly, and WTs cannot change their speed as fast as the wind due to their inertia. This will render WTs unable to follow the fast wind speed variations, and consequently being unable to perform at their optimal TSR. Various MPPT controls have been proposed: see e.g. [176] for a review. The so-called TSR-control, which adjusts the rotor speed to achieve optimal TSR based on measurements of wind speed and rotor speed, has many advantages in comparison with other controls with the drawback that wind speed can be difficult and costly to measure. In this chapter, we attempt to predict the wind speed itself a few seconds in the future, and the predictions could possibly be used to improve the power extraction. To this end, this chapter investigates an application of KMD to short-term wind speed prediction. As was discussed in [168], sudden gusts cause structural load and damage to WT structures, which can possibly be mitigated by utilizing very short-term wind speed predictions to initiate protection measures in WTs such as adjusting the pitch angles [66] of the WTs' blades. That is another possible application of the predictions investigated here. This chapter will discuss very short-term wind speed prediction by attempting to answer the following two questions:

- (i) Is it possible to predict short-term wind speed fluctuations, and is there any determinism governing them?
- (ii) On which time-scale do KMD-based prediction techniques and common prediction techniques work in an effective manner, and when can they beat the naive predictor?

Question (i) is fundamental to prediction: if the measurements are by their nature or by noise contamination unusable, i.e. future values are not predictable, it is meaningless to attempt prediction. However, if some determinism is found, then it gives an idea about the underlying dynamics which can be exploited for prediction. Question (ii) is important for potential practical applications of the prediction methods in WT control, where their usage depends on their prediction accuracy on various time-horizons. The answers to Questions (i) and (ii) are provided in Sections 8.3 and 8.5, respectively.

1-Hz measurements from WTs in the wind farm described in Chapters 5 and 6 are used here. In Fig. 8.1 (a), measured and simulated WT-outputs are plotted against the wind speed, and shown together with estimated power curves. As mentioned before, the figure clearly illustrates how the measured and simulated wind power-speed trajectories fluctuate around the power curves. Here, the power curves represent optimal relation between output power and wind speed, and thus in the ideal case, we would like the trajectories to stay on the power curves. In Fig. 8.1 (b), the simulated output power for the same time-period is plotted against low-pass filtered wind speeds in two cases: normal and low WT inertia. Normal inertia means inertia constants set as follows: $H_t = 6$ s, $H_g = 0.6$ s (see (4.6)), and in the low inertia case: $H_t = 0.6$ s, $H_g = 0.1$ s. In the case of low inertia, it can be seen that the trajectory almost remains on the power curve 'Mod. Curve', while the WT in the case of normal inertia is not able to extract as much power. In this chapter, short-term wind speed prediction is attempted, with the application of being used in WT control in mind, such that the power-speed trajectory stays closer to the optimal curve, which allows for more power to be extracted from the wind.

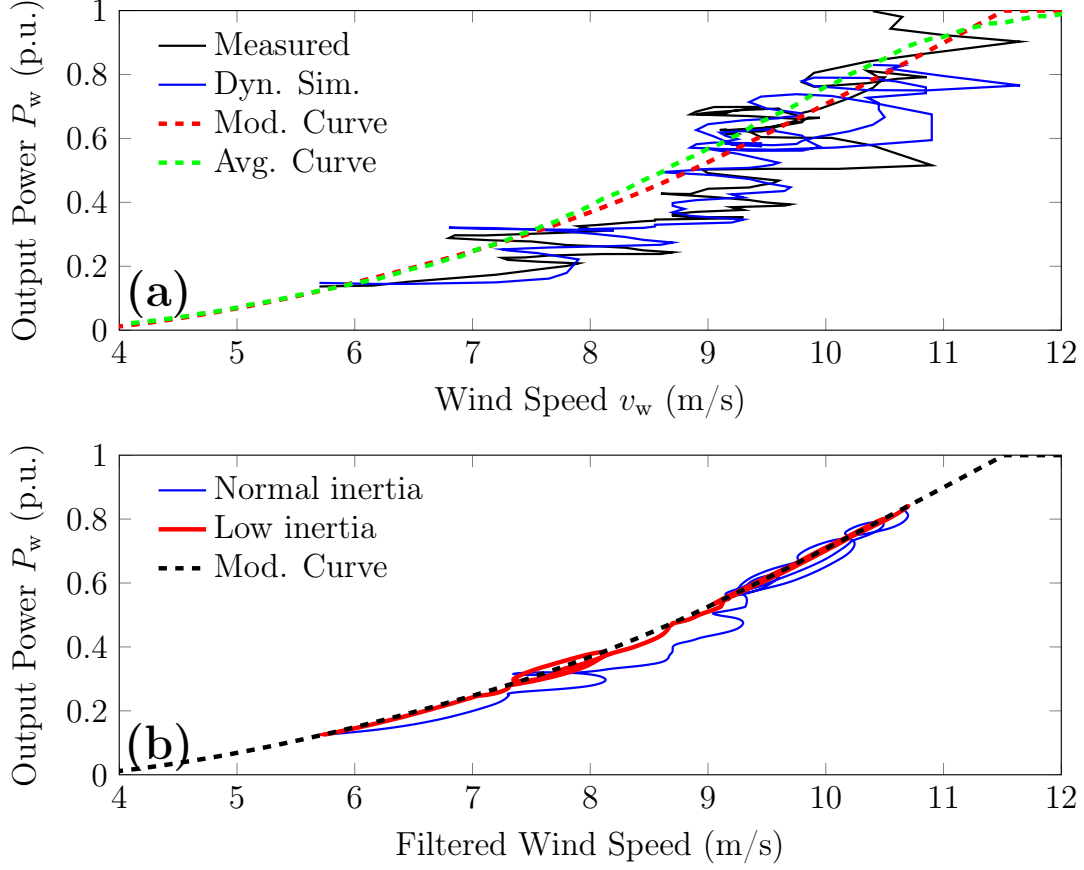


Figure 8.1: (a) Measured and simulated power fluctuations around power curves: these plots also appeared in Fig. 5.6 (a). (b) WT outputs from simulations run with normal and very low values on WT inertia, plotted against low-pass filtered wind speeds.

8.3 Characterization of wind speed fluctuations

Before attempting to predict future values of a signal, it is necessary to understand the underlying dynamics giving rise to the observed signal. Since here we attempt to predict short-term wind speed fluctuations, it is noted that they are often modeled by a normal (Gaussian) distribution [113, 153, 177]. This is arguably quite accurate when considering normalized fluctuations from a 10-min average [112]: see the models (5.1) and (5.3), and Fig. 5.7 (a) in Chapter 5. However, it was shown in Chapter 5 that this assumption fails when looking at increment statistics, which clearly indicate non-Gaussian characteristics. Thus, the description of wind speed fluctuations as Gaussian noise superimposed on the averaged wind speed is a simplification which ignores the complex dynamics of turbulence [110]. If information about the deterministic properties could be derived from data, then it could be exploited for more accurate prediction. Therefore, it is important to determine any deterministic feature in the wind speed time-series considered here. First,

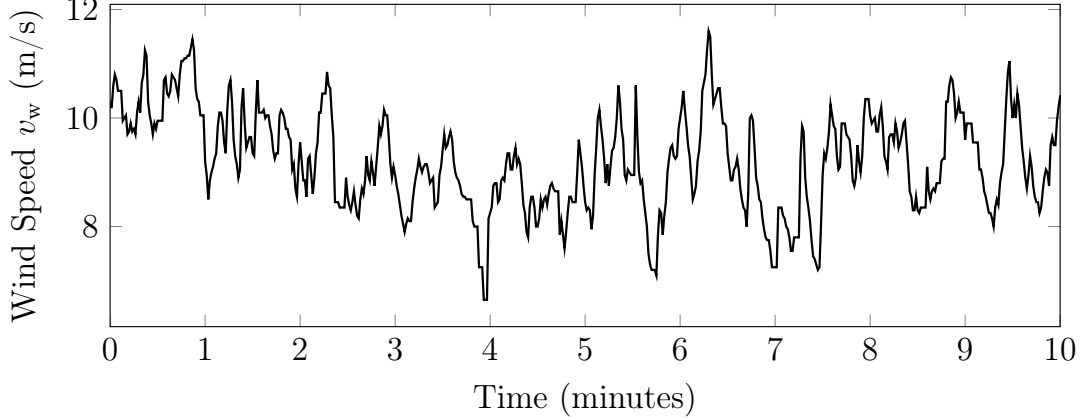


Figure 8.2: Example of 10-min time-series of measured wind speeds.

to remove possible measurement noise, wind speeds are filtered through a low-pass filter which represents the natural filtering of high frequency components over the rotor surface [92]. The time constant of the filter is set to 4 s as recommended in [92]. Thus, in addition to the removal of noise, the filtered wind speeds can also be expected to be closer to what the WT experiences.

To check if the signal evolves according to a rule (determinism), some notation needs to be introduced. Sampled wind speeds are now denoted by u . In total there are N measurements: $u_1, \dots, u_N$. Then, snapshot-vectors are constructed as in (3.5) ($n \geq 1$):

$$\mathbf{y}_n(m, L) := [u_n, u_{n+L}, \dots, u_{n+(m-1)L}]^\top \in \mathbb{R}^m, \quad (8.2)$$

where L is the lag in number of samples, and $m \geq 1$ the vector length, or spatial dimension. The parameters L and m need to be fixed. The lag L is determined by utilizing the so-called average mutual information [178, 179]:

$$I(L) = \sum_{u_n, u_{n+L}} p(u_n, u_{n+L}) \log_2 \left[\frac{p(u_n, u_{n+L})}{p(u_n) p(u_{n+L})} \right], \quad (8.3)$$

where $p(\cdot, \cdot)$ denotes a conditional probability density, evaluated by first binning the data using the histogram function (`hist`) in MATLAB, and $\log_2$ the base 2 logarithm. In particular, L is chosen where (8.3) has its first minimum. This is regarded as the lag L where sufficient independence from the previous measurement has been achieved, and becomes an appropriate choice for a so-called time delayed coordinate [178]. Note that (8.3) can be viewed as a nonlinear alternative to autocorrelation [179].

As an example, (8.3) is evaluated for a 5-min time series of measured wind speeds such as shown in Fig. 8.2, after they have first been filtered through the low-pass filter previously mentioned, and having their mean-value subtracted. The result is shown in Fig. 8.3 (a) which shows that $L = 9$ is an appropriate choice, in this case. Second, with L determined, the two indexes introduced in [173] are utilized, denoted by $E1(m)$ and

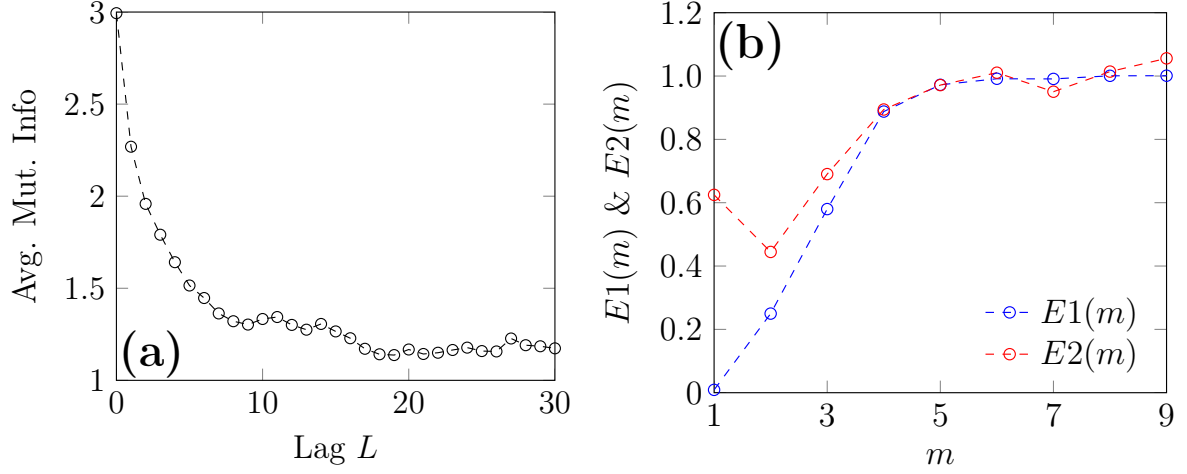


Figure 8.3: Example of evaluation of wind speed measurements by applying (a) average mutual information according to (8.3); (b) $E1$ and $E2$ indexes: see Appendix 8.A.

$E2(m)$ (explained in Appendix 8.A). The results of applying $E1$ and $E2$ to the filtered wind speeds, and utilizing snapshots according to (8.2) are shown in Fig. 8.3 (b). The results indicate two things: first, that the wind speed time-series is not purely stochastic (on the time-scale of several seconds) because $E2(m) \neq 1$ for low m , and second, that a choice of spatial dimension $m = 5$ or 6 is appropriate (where $E1(m)$ reaches 1) in this specific case.

As mentioned before, boundary layer wind speeds are *non-stationary and turbulent* [26, 112, 180], meaning that results such as in Fig. 8.3 might vary, and that the predictability may vary depending on the atmospheric (wind) conditions. This implies that the effectiveness or suitability of a particular prediction method can vary over time. To exemplify the non-stationary characteristics of the measured wind speeds, turbulence intensities TI calculated according to (5.2) for non-overlapping 600-s windows over a 3-day period are shown in Fig. 8.4. While TI often exhibits rather small variations (such as in the interval 2–2.5 days) between consecutive windows, rapid changes such as a doubling of its value are also seen.

One could suspect that the perceived low-dimensional deterministic nature of wind speed fluctuations is caused by linear correlations which are amplified by the low-pass filtering (see e.g. [181]), rather than nonlinear determinism. To help clarify the nature of the short-term fluctuations, “synthetic” wind speed fluctuations according to [110] will be generated with a Stochastic Differential Equation (SDE). To do so, it is first assumed that the autocorrelation of normalized wind speed fluctuations $z'(t)$ (see (5.3) and (5.7)), can be approximated by

$$R_{z',z'}(\tau) = \langle z'(t + \tau)z'(t) \rangle \approx \exp(-\gamma\tau), \quad (8.4)$$

for time-shifts τ up to some value, where γ is a decay (or diffusion) coefficient in an exponential approximation of the autocorrelation. The coefficient γ can be determined

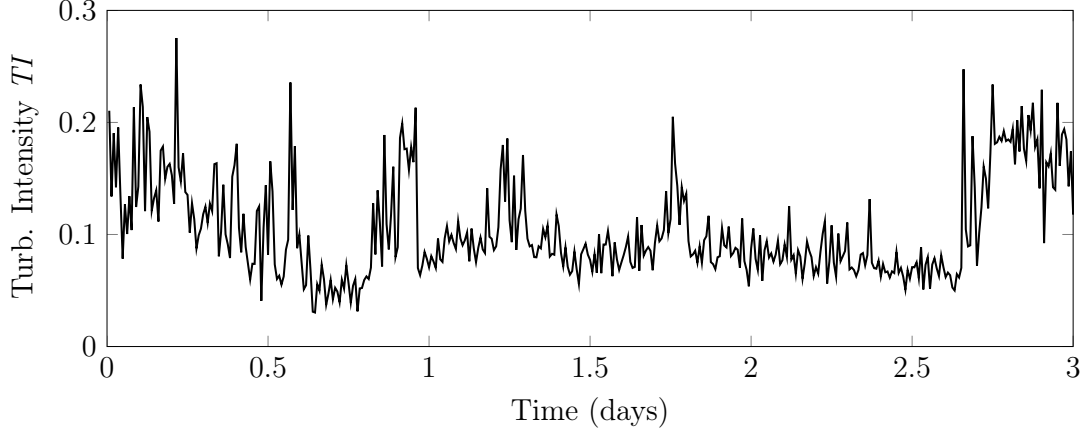


Figure 8.4: Turbulence intensities calculated for non-overlapping 10-min windows.

by means of curve-fitting using (8.4) and measured wind speeds decomposed according to (5.3). By doing so, γ is here fixed at 0.05. Then, synthetic wind speed fluctuations are generated by simulating the following SDE:

$$\frac{dz'(t)}{dt} = -\alpha\gamma z'(t) + \sqrt{\gamma} \cdot \Gamma(t), \quad (8.5)$$

which is called a Ornstein–Uhlenbeck process [110], where α is the friction coefficient [10], and $\Gamma(t)$ denotes Gaussian noise with variance $\sigma^2 = 2$. For now, we set $\alpha = 1$ in accordance with [110], and solve the SDE (8.5) according to

$$z'(t + \Delta t) = z'(t) - \gamma z'(t)\Delta t + \sqrt{\Delta t\gamma} \cdot \Gamma(t), \quad (8.6)$$

where Δt is the time-step set to 1 s in order to match the sampling frequency of the measured wind speeds. With $\gamma = 0.05$ and $\Delta t = 1$, (8.6) can be rewritten as

$$z'_{k+1} = (1 - \gamma)z'_k + \sqrt{\gamma} \cdot \Gamma_k = 0.95z'_k + \sqrt{\gamma} \cdot \Gamma_k, \quad (8.7)$$

where k denotes the discrete time with each time-step corresponding to 1 s. The model (8.7) is a first-order autoregressive process (since z'_{k+1} only depends on z'_k and the noise).

Now, the indexes $E1$ and $E2$ are evaluated for measured and synthetic wind speed fluctuations of low-pass filtered wind speeds. In Fig. 8.5 (a) the results of averaging the results over one hundred 600-s windows show that the characteristics of these two different data are nearly identical. It is also noted that the results are very close to the results shown in Fig. 8.3 (b). As suggested in [182], to emphasize nonlinear features of the data, and to attenuate the effects of linear correlations, increments $dz'(t) := z'(t+1) - z'(t)$ are now considered instead of fluctuations $z'(t)$. If the data are generated by a deterministic nonlinear system, it should not matter much whether dealing with z' or dz' [182]. The results of computing $E1$ and $E2$ for wind speed increments $dz'(t)$ are shown in Fig. 8.5 (b). In this case too, the results are almost identical for the two data sets, and we see that $E2 \approx$

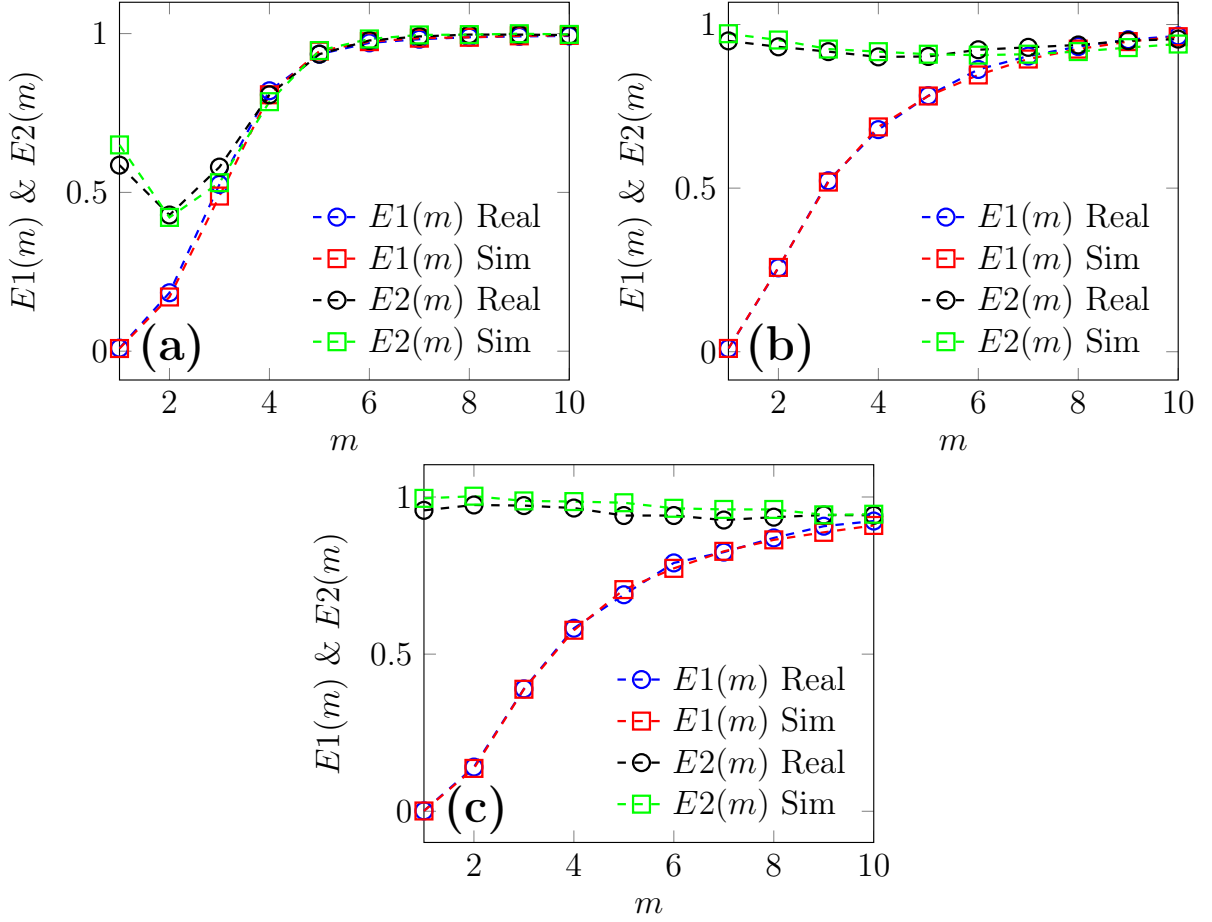


Figure 8.5: $E1$ and $E2$ for measured ‘Real’ and simulated (synthetic) ‘Sim’ wind speeds, both filtered through a low-pass filter with a time-constant of 4 s. (a) Averaged results of one hundred 600-s long time-series of fluctuations. (b) Averaged results of one hundred 600-s long time-series of wind speed increments. (c) Results of 20000-s long time-series of increments.

1 for all m , indicating that the dynamics of wind speed increments are stochastic rather than deterministic. This implies that the signs of determinism found in the fluctuations of filtered wind speeds are due to linear correlations, and that the measured wind speeds are described well by (8.5), i.e. a stochastic process with an exponentially decaying autocorrelation function. The results of $E1$ and $E2$ applied to 20000-s long time-series of wind speed increments are shown in Fig. 8.5 (c) and are similar to Fig. 8.5 (b).

It was shown in [10] that application of Extended DMD (EDMD) (see Section 2.2.6) to data generated by (8.5) accurately outputs the eigenvalues and eigenfunctions of the Koopman operator associated with (8.5). Thus, we realize that EDMD applied to measured wind speeds (if they are well-described by (8.5)), lets us qualitatively compare the data-derived eigenfunctions with the theoretical ones, to confirm the correctness of the SDE. This is however outside the scope of this study.

8.4 Prediction methods

Here, several prediction methods are introduced and discussed. First, however, the subsection that follows introduces some notations and definitions. Also, two basic prediction methods are introduced for the purpose of comparison: the naive predictor mentioned before, and prediction based on the trend of the time-series. After that, a prediction method based on nearest neighbors [168, 179] is described, which is a commonly used method for nonlinear time-series. It is shown how the method can be interpreted in terms of Koopman operator theory. Following that, the well-known autoregressive (AR) model [111, 168] and its connections to Dynamic Mode Decomposition (DMD) and KMD are discussed. Furthermore, two new methods are introduced: a combination between nearest neighbor and DMD, and a new alternative variant of DMD. The following list summarizes all prediction methods evaluated in this chapter:

- (1) Standard DMD, abbreviated ‘DMD’: see (3.6).
- (2) Arnoldi-like KMD, abbreviated ‘Arnoldi’: see (3.6).
- (3) Vector Prony KMD, abbreviated ‘Prony’: see (3.6).
- (4) Reshaped DMD, abbreviated ‘DMD-Re’: see (8.36).
- (5) DMD and nearest neighbor prediction, abbreviated ‘DNN’: see (8.34).
- (6) Nearest neighbor prediction, abbreviated ‘NN’: see (8.13).
- (7) Linear trend predictor, abbreviated ‘Trend’: see (8.12).

Methods (1)–(4) have been chosen in order to clarify the predictive capabilities of the algorithms that have been discussed throughout this dissertation. In particular, Method (4) is a new method proposed here, and is a variant of standard DMD. Method (5) is the second method proposed here, and acts as an interesting hybrid between a commonly used prediction method for nonlinear time-series and DMD. Method (6) is chosen because it has been used for short-term wind speed prediction [168, 171, 174], and has been applied to other nonlinear data: see e.g. [182] where a closely related method is used. Method (7) acts as a simple method for comparison.

8.4.1 Preliminaries and simple prediction methods

We consider N wind speeds measurements $u_1, \dots, u_N$ sampled with a fixed time-step, from which vector-valued snapshots are constructed similarly to (8.2):

$$\mathbf{y}_k(m, L) := [u_{1+(k-1)L}, u_{1+kL}, \dots, u_{1+(k+m-2)L}]^\top \in \mathbb{R}^m, \quad (k, m \geq 1), \quad (8.8)$$

giving N_1 vector-valued snapshots as input to the prediction methods:

$$\mathbf{y}_1, \dots, \mathbf{y}_{N_1}. \quad (8.9)$$

The Δn -step ahead wind speed prediction is denoted by $\hat{u}_{N+\Delta n}$, and its true value is $u_{N+\Delta n}$. The prediction error is then defined by

$$\Delta \hat{u}_{N+\Delta n} := \hat{u}_{N+\Delta n} - u_{N+\Delta n}. \quad (8.10)$$

Two very simple prediction methods are applied for the sake of comparison. First, the naive predictor, which was already used in Chapter 3, and explained in the introduction of this chapter, is used, and which predicts the future wind speed as follows:

$$\hat{u}_{N+\Delta n} = u_N. \quad (8.11)$$

Second, for one-step ahead prediction, a simple model utilizing the trend of the time-series is used, defined in [168] as follows:

$$\hat{u}_{N+1} = 2u_N - u_{N-1}. \quad (8.12)$$

In [168], (8.12) is argued to be better suited for prediction of smooth data than turbulent data such as wind speed time-series.

8.4.2 Nearest neighbor prediction

Here, the popular prediction method for nonlinear data following [168, 179] is described, which we call the nearest neighbor prediction method. The method utilized in [182] is similar. The idea is to find snapshots $\mathbf{y}_n$ that satisfy $\|\mathbf{y}_N - \mathbf{y}_n\|_\infty < \epsilon$ (ϵ is a small constant) and form a neighborhood $\mathcal{N}_\epsilon$ around $\mathbf{y}_N$ (the last snapshot). The idea is then to look up the “predictions”, i.e. the snapshots Δn steps following $\mathbf{y}_n \in \mathcal{N}_\epsilon$, denoted by $\mathbf{y}_{n+\Delta n}$. These can be considered as estimates of $\mathbf{y}_{N+\Delta n}$. The Δn -step future prediction can then be computed by

$$\hat{\mathbf{y}}_{N+\Delta n} = \frac{1}{|\mathcal{N}_\epsilon|} \sum_{\mathbf{y}_n \in \mathcal{N}_\epsilon} \mathbf{y}_{n+\Delta n}, \quad (8.13)$$

where $|\mathcal{N}_\epsilon|$ denotes the number of utilized snapshots, i.e. number of neighbors to $\mathbf{y}_N$. In other words, (8.13) can be interpreted as estimating the expectation value of $\mathbf{y}_{N+\Delta n}$ by averaging over the Δn -step future values of the $|\mathcal{N}_\epsilon|$ neighbors $\mathbf{y}_n \in \mathcal{N}_\epsilon$ to $\mathbf{y}_N$. Furthermore, the optimal predictor for Markov-chains corresponds to the nearest neighbor method [168]. This method’s connection to Koopman operator theory will now be discussed. Following [14, 18, 183], the action of the *stochastic* Koopman operator $\tilde{\mathbf{U}}$, for the discrete-time Markov process

$$\mathbf{x}_{k+1} = \mathbf{T}(\mathbf{x}_k, \boldsymbol{\omega}_k), \quad (8.14)$$

is defined as

$$(\tilde{\mathbf{U}}g)(\mathbf{x}) = \mathbb{E}[g(\mathbf{T}(\mathbf{x}_k, \boldsymbol{\omega}_k))], \quad (8.15)$$

where $\mathbf{x}$ is the state evolving according to the map $\mathbf{T}$ on the state space $\mathbb{X}$, $\mathbb{E}[\cdot]$ denotes the expectation value, $\boldsymbol{\omega}_k \in \Omega_s$ is a point in the probability space of the stochastic dynamics Ω_s [18], and g an observable. Thus, we have that $(\tilde{\mathbf{U}}g)(\mathbf{x}_k) = \mathbb{E}[g(\mathbf{x}_{k+1})]$ according to

(8.14) and (8.15), and hence that the application of $\tilde{\mathbf{U}}$ provides the expectation value of g evaluated at the next time-step: $g(\mathbf{x}_{k+1})$, assuming that $\mathbb{E}[g(\mathbf{x}_{k+1})]$ exists (for some cases $\mathbb{E}[\cdot]$ does not exist). Consequently, (8.13) can be interpreted as

$$\hat{\mathbf{y}}_{N+\Delta n} = (\tilde{\mathbf{U}}^{\Delta n} \mathbf{g})(\mathbf{x}_N) := \begin{bmatrix} \mathbb{E}[u_{1,N+\Delta n}] \\ \mathbb{E}[u_{2,N+\Delta n}] \\ \vdots \\ \mathbb{E}[u_{m,N+\Delta n}] \end{bmatrix} \approx \frac{1}{|\mathcal{N}_\epsilon|} \sum_{\mathbf{y}_n \in \mathcal{N}_\epsilon} \mathbf{y}_{n+\Delta n}, \quad (8.16)$$

in the case that the dynamics are given by (8.14), $\mathbf{y}_i := \mathbf{g}(\mathbf{x}_i) = [g_1(\mathbf{x}_i), \dots, g_m(\mathbf{x}_i)]^\top = [u_{1,i}, \dots, u_{m,i}]^\top$, $\tilde{\mathbf{U}}$ acts on each component $g_j(\mathbf{x}_i)$ of $\mathbf{g}$ individually according to $(\tilde{\mathbf{U}} g_j)(\mathbf{x}_i) := \mathbb{E}[u_{j,i+1}]$, and if the number of neighbors $|\mathcal{N}_\epsilon|$ is appropriately chosen. In particular, for time-shifted measurements, we have that $g_1 = g_2 = \dots = g_m$, i.e. $\mathbf{y}_i = [g_1(\mathbf{x}_i), \dots, g_1(\mathbf{x}_{i+m-1})]^\top := [u_1, \dots, u_{i+m-1}]^\top$.

8.4.3 Autoregressive (AR) model

This section discusses the well known autoregressive model. The relations between it and DMD/KMD are discussed in the two sections that follow. Again, consider N samples of e.g. wind speeds: $u_1, \dots, u_N$. Then, as in [111], a zero-mean autoregressive (AR) model of order p , denoted by $\text{AR}(p)$, is defined as follows:

$$u_n = \sum_{j=1}^p a_j u_{n-j}, \quad n = 1, \dots, N. \quad (8.17)$$

In the case that we want to predict the next value u_{N+1} we have

$$\hat{u}_{N+1} = \sum_{j=1}^p a_j u_{N+1-j}, \quad (8.18)$$

i.e. we are utilizing the p previous observations scaled by the coefficients a_j determined from (8.17). If we want to predict u_{N+2} we repeat the procedure (8.18), now including our previous prediction $\hat{u}_{N+1}$ assuming it is equal to u_{N+1} . To determine the a_j coefficients of (8.17), (8.17) is rewritten into matrix form as follows:

$$\begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} = \begin{bmatrix} u_0 & \dots & u_{1-p} \\ \vdots & \ddots & \vdots \\ u_{N-1} & \dots & u_{N-p} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}. \quad (8.19)$$

However, since the sampled data are only $u_1, \dots, u_N$ (u_0 or earlier samples are not available), the p top rows of (8.19) are disregarded and (8.19) is rewritten as

$$\begin{bmatrix} u_{1+p} \\ \vdots \\ u_N \end{bmatrix} = \begin{bmatrix} u_p & \dots & u_1 \\ \vdots & \ddots & \vdots \\ u_{N-1} & \dots & u_{N-p} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}. \quad (8.20)$$

Now, the notation of snapshots from (8.2) is used: $\mathbf{y}_n(m, L)$. Setting $L = 1$ and $m = N - p$ gives e.g. $\mathbf{y}_1 = [u_1, \dots, u_{N-p}]^\top$ and $\mathbf{y}_p = [u_p, \dots, u_N]^\top$. In other words, with this notation, (8.20) can be rewritten as follows:

$$\mathbf{y}_{p+1} = \begin{bmatrix} \mathbf{y}_p & \dots & \mathbf{y}_1 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}, \quad (8.21)$$

and re-ordered as:

$$\mathbf{y}_{p+1} = \begin{bmatrix} \mathbf{y}_1 & \dots & \mathbf{y}_p \end{bmatrix} \begin{bmatrix} a_p \\ \vdots \\ a_1 \end{bmatrix} =: \mathbf{X}\mathbf{a}, \quad (8.22)$$

and a least-squares solution is given by $\mathbf{a} = \mathbf{X}^\dagger \mathbf{y}_{p+1} \in \mathbb{R}^p$, where $\dagger$ denotes the pseudo-inverse. Note that the matrix $\mathbf{X}$ in (8.22) has a Hankel-matrix-like structure as discussed in Chapters 2 and 3. Then, the prediction $\hat{u}_{N+1}$ is given by evaluating (8.18) with the coefficients contained in $\mathbf{a}$. A more common method to estimate $\mathbf{a}$ is by utilizing the so-called Yule-Walker (YW) equations which involve computation of autocovariances [111]. In MATLAB, the AR-coefficients are obtained (with negative signs compared to $\mathbf{a}$ here) via a method based on the YW equations by executing the command `aryule`.

8.4.4 On standard DMD and AR

This section discusses the relation between DMD and AR for time-shifted measurements. The DMD utilizes two matrices of snapshots (still using same notation as in (8.22)): $\mathbf{X} = [\mathbf{y}_1, \dots, \mathbf{y}_p]$, and $\mathbf{Y} = [\mathbf{y}_2, \dots, \mathbf{y}_{p+1}]$, $\mathbf{y}_i \in \mathbb{R}^{m_1}$, where the spatial dimension $m_1 := N - p$ as in (8.21) and (8.22). The relation between $\mathbf{X}$ and $\mathbf{Y}$ is exactly like in Chapter 2 governed by the matrix $\mathbf{A}$ defined as follows:

$$\mathbf{Y} = \mathbf{A}\mathbf{X}. \quad (8.23)$$

Then, a least-squares solution is given by

$$\mathbf{A} = \mathbf{Y}\mathbf{X}^\dagger \in \mathbb{R}^{m_1 \times m_1}. \quad (8.24)$$

Assuming $\text{rank}(\mathbf{X}) = m_1$ such that $\mathbf{A}$ has m_1 non-zero eigenvalues, the structure of $\mathbf{A}$ becomes as follows:

$$\mathbf{A} = \begin{bmatrix} \mathbf{0}_{m_1-1} & \mathbf{I}_{m_1-1} \\ \mathbf{b}^\top & \end{bmatrix}, \quad (8.25)$$

where $\mathbf{0}_{m_1-1} \in \mathbb{R}^{m_1-1}$ is a column vector of zeros, and $\mathbf{I}_{m_1-1}$ the $(m_1 - 1) \times (m_1 - 1)$ identity matrix. Note that the vector $\mathbf{b} \in \mathbb{R}^{m_1}$ could alternatively be solved as

$$\mathbf{b} = (\mathbf{X}^\top)^\dagger [u_{N-p+1}, \dots, u_N]^\top, \quad (8.26)$$

where $[u_{N-p+1}, \dots, u_N]^\top = \mathbf{y}_{N-p+1}(p, 1)$ from (8.2). We can convince ourselves about the structure of (8.25) by considering the following example:

$$\begin{bmatrix} u_2 & u_3 & u_4 & u_5 \\ u_3 & u_4 & u_5 & u_6 \\ u_4 & u_5 & u_6 & u_7 \end{bmatrix} = \mathbf{A} \begin{bmatrix} u_1 & u_2 & u_3 & u_4 \\ u_2 & u_3 & u_4 & u_5 \\ u_3 & u_4 & u_5 & u_6 \end{bmatrix}, \quad (8.27)$$

where naturally the least-squares solution of $\mathbf{A}$ has the form of (8.25):

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ b_1 & b_2 & b_3 \end{bmatrix}, \quad (8.28)$$

and the constants b_i according to (8.26) are components of $\mathbf{b} = (\mathbf{X}^\top)^\dagger [u_4, u_5, u_6, u_7]^\top \in \mathbb{R}^3$, where $\mathbf{X}$ in this example is the data-matrix in the right hand side of (8.27). Explicitly, the following regression problem was solved to determine $\mathbf{b}$:

$$\begin{bmatrix} u_1 & u_2 & u_3 \\ u_2 & u_3 & u_4 \\ u_3 & u_4 & u_5 \\ u_4 & u_5 & u_6 \end{bmatrix} \mathbf{b} = \begin{bmatrix} u_4 \\ u_5 \\ u_6 \\ u_7 \end{bmatrix}. \quad (8.29)$$

It is therefore understood that the vector $\mathbf{b}$ contains coefficients of an $\text{AR}(N - p)$ -model, i.e. same order as the spatial dimension $m_1 = N - p$, in comparison with the $\text{AR}(p)$ above in (8.17). Note that in the example (8.27), $N = 7$ (measurements) and $p = 4$ (four snapshots/columns in $\mathbf{X}$), giving $\mathbf{b} \in \mathbb{R}^3$. Also note that $\mathbf{A}^\top$ of (8.25) and (8.28) are companion-type matrices: see the next subsection. Finally, the one-step future prediction is given as follows:

$$\hat{u}_{N+1} = \mathbf{y}_{p+1}^\top \mathbf{b} =: \mathbf{y}_N^\top \mathbf{b},$$

with $\mathbf{y}_N = \mathbf{y}_{p+1}$ denoting the last snapshot.

8.4.5 On Arnoldi-like KMD and AR

This section connects AR with Arnoldi-like KMD for time-shifted measurements. In Arnoldi-like KMD, following Chapter 2, we have the following equation:

$$\mathbf{y}_{p+1} = \mathbf{X}\mathbf{c} + \mathbf{r}, \quad \mathbf{r} \perp \text{span}(\mathbf{X}), \quad (8.30)$$

where $\mathbf{c} = [c_1, \dots, c_p]^\top$ is a vector of coefficients, $\mathbf{r}$ a vector of residuals, and one solution is $\mathbf{c} = \mathbf{X}^\dagger \mathbf{y}_p$ [78]. Now, we notice that $\mathbf{c}$ is exactly the vector of AR coefficients $\mathbf{a}$ solved earlier in (8.21). The relation between $\mathbf{X}$ and $\mathbf{Y}$ (see the previous subsection) can then be written

$$\mathbf{Y} = \mathbf{X}\mathbf{C} + \mathbf{r}[0, \dots, 0, 1], \quad (8.31)$$

where $\mathbf{C}$ is the companion matrix defined as follows:

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & \dots & 0 & c_1 \\ 1 & 0 & & 0 & c_2 \\ 0 & 1 & & 0 & c_3 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & c_p \end{bmatrix}.$$

From these examples, it is understood that the regression problem of KMD algorithms results in AR-like regression of various orders, which can be viewed as vector autoregression [72] (VAR). In particular, for Hankel-like matrices obtained for time-shifted measurements, AR-type models are achieved (not VAR), and their spectral analyses are basically Prony analysis [184], which was discussed in Chapter 2 in the case of vector-valued observables. In the common DMD and KMD procedures, spectral decompositions (of e.g. $\mathbf{A}$ in DMD) are computed and dominant modes are identified: see Chapters 2 and 3. Hence, the matrices $\mathbf{A}$ and $\mathbf{C}$ are often not explicitly used, and instead the methods aim to identify a reduced-order model (decomposition) of the following type:

$$\mathbf{y}_k = \sum_{i=1}^r \lambda_i^k \mathbf{v}_i,$$

constituted by r dominant pairs of eigenvalues $\lambda_i \in \mathbb{C}$ and modes $\mathbf{v}_i \in \mathbb{C}^{m_1}$.

8.4.6 DMD and nearest neighbor prediction

A method is proposed in the following which combines DMD and the idea of nearest neighbors, and is reminiscent of the “locally linear” model in [168] and the method proposed in [185]. Suppose we have identified a k number of snapshots $\mathbf{y}_n \in \mathcal{N}_\epsilon$ ($n = 1, \dots, k$) forming a neighborhood around the last observation $\mathbf{y}_N$. Their Δn -step predictions (future values) are now denoted by $\hat{\mathbf{y}}_{n+\Delta n}$. Then, the following linear-regression problem can be constructed and solved:

$$[\hat{\mathbf{y}}_{1+\Delta n}, \dots, \hat{\mathbf{y}}_{k+\Delta n}] = \mathbf{A}_k [\mathbf{y}_1, \dots, \mathbf{y}_k], \quad (8.32)$$

or more simply written as

$$\hat{\mathbf{Y}}_k = \mathbf{A}_k \mathbf{X}_k, \quad (8.33)$$

and solved according to $\mathbf{A}_k = \hat{\mathbf{Y}}_k \mathbf{X}_k^\dagger$. Then, the $n \times \Delta n$ -step prediction is given by

$$\hat{\mathbf{y}}_{N+n\Delta n} = (\mathbf{A}_k)^n \mathbf{y}_N, \quad (8.34)$$

since $\mathbf{A}_k$ estimates $\mathbf{y}$ ’s evolution Δn -steps forward at each iteration.

8.4.7 Reshaped DMD via optimal hard threshold

Here, another method is proposed, based on first reshaping the matrix of input data based on a threshold for singular values, and second, by applying DMD to the reshaped data-matrix. The method is outlined as follows. First, apply the Singular Value Decomposition (SVD) to the matrix containing all measurements: $\mathbf{D} := [\mathbf{y}_1, \dots, \mathbf{y}_N]$, where $\mathbf{y}_k \in \mathbb{R}^m$ is the snapshot at the k -th discrete time instance, as usual. Then, based on the optimal hard threshold for singular values [80], r number of “dominant” (large) singular values are identified. Now, instead of using these singular values, the data matrix $\mathbf{D}$ is reshaped into a new matrix $\mathbf{D}_r$ of different size: $\mathbf{D}_r := [\tilde{\mathbf{y}}_1, \dots, \tilde{\mathbf{y}}_{N_1}] \in \mathbb{R}^{r \times N_1}$ ($r \leq m$), i.e. a matrix constructed by N_1 number of snapshots $\tilde{\mathbf{y}}_k \in \mathbb{R}^r$. I.e. exactly the same size as the number of dominant singular values. Finally, the DMD-matrix is calculated as

$$\mathbf{A}_{\text{Re}} = \mathbf{Y}_{\text{Re}} \mathbf{X}_{\text{Re}}^\dagger, \quad (8.35)$$

where $\mathbf{X}_{\text{Re}} = [\tilde{\mathbf{y}}_1, \dots, \tilde{\mathbf{y}}_{N_1-1}]$, and $\mathbf{Y}_{\text{Re}} = [\tilde{\mathbf{y}}_2, \dots, \tilde{\mathbf{y}}_{N_1}]$. The n -step prediction is consequently found as the last component of the following vector:

$$\hat{\mathbf{y}}_{N_1+n} = \mathbf{A}_{\text{Re}}^n \tilde{\mathbf{y}}_{N_1}. \quad (8.36)$$

The idea of this method is that the number of dominant singular values is linked to the dimensionality of the dynamical system [173]. Thus, the derived matrix $\mathbf{A}_{\text{Re}}$ should be of appropriate (optimal) size with regards to dimension. Furthermore, the method is adaptable, since r will change depending on the prevailing characteristics of $\mathbf{D}$.

8.5 Numerical test of wind speed prediction

The predictive capabilities of the seven methods listed in the introduction of Section 8.4 will now be evaluated by applying them to wind speed fluctuations z' (of low-pass filtered wind speeds) according to (5.3) ($T = 5$ h) at a WT in the WF discussed in Chapters 5 and 6. A number n_p of “prediction points” or samples, and a time-lag L between consecutive samples are fixed. For every choice of L , at each point, future predictions for a number of combinations of spatial dimension m and snapshots N_s are computed with each method (done exactly in the same fashion as in Chapter 3). Example time-series of predictions from several methods are shown in Fig. 8.6. In Fig. 8.6 (a) it is observed that ‘DMD/Arnoldi’ and ‘DMD-Re’ are very close and perform much better than the naive predictor, and in Fig. 8.6 (b) that ‘DMD-NN’ and ‘NN’ are very close, but ‘DNN’ looks slightly better (the advantage of ‘DNN’ is shown later on). These results indicate that the methods are capable of following the fluctuations better than the naive predictor.

Sample results for different combinations of m and N_s are shown in Fig. 8.7, for L set to 2 samples (2 s) and $n_p = 1000$. The coloring indicates the reduction in the RMSE compared to the naive predictor. For example, if the naive predictor achieves an RMSE value of say, r_n , and some prediction method achieves an RMSE value of r_m , then the

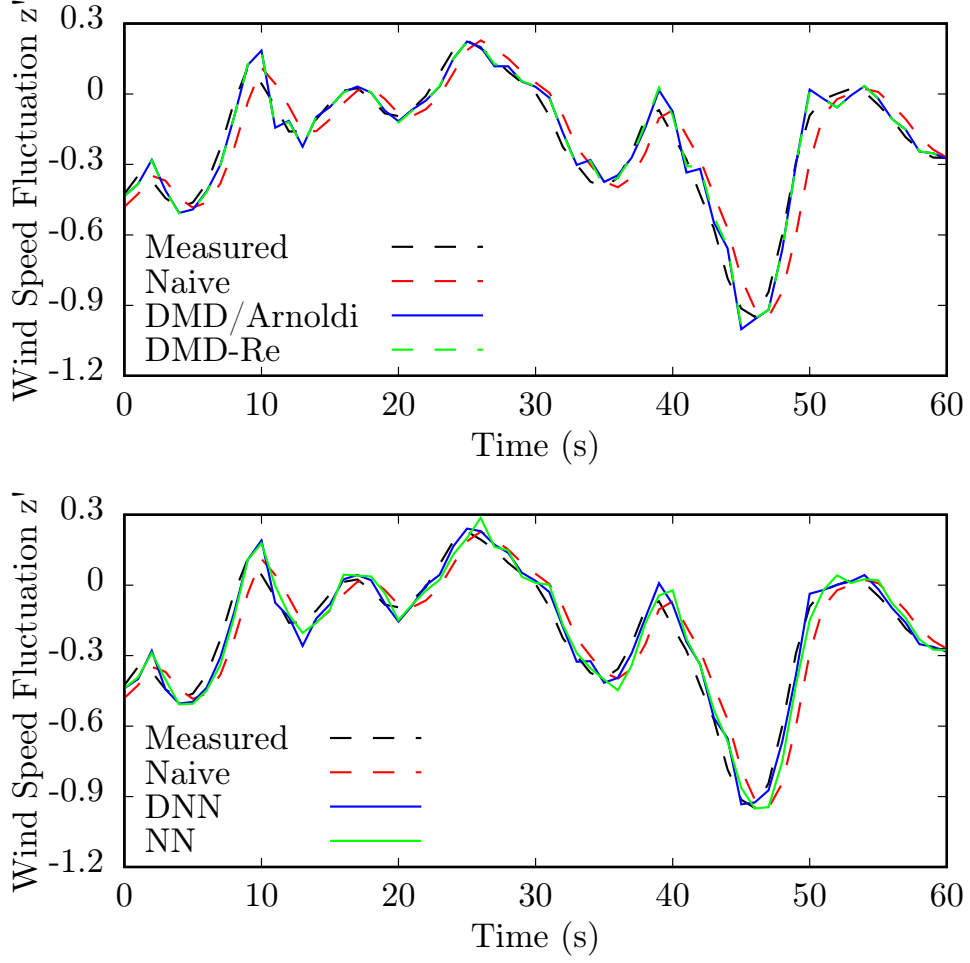


Figure 8.6: Example time-series for 2-s predictions.

displayed value is r_m/r_n . For clarity, coloring only indicates the prediction accuracy in the interval $[0, 1]$. A value ‘1’ would indicate an equal or worse result than the naive predictor. The results show that predictions with DMD and Arnoldi-like KMD are close to equally accurate, and are reminiscent of the results in Chapter 3, Fig. 3.11. DMD, Arnoldi-like KMD, reshaped DMD ‘DMD-Re’, and the combination of DMD and nearest neighbor ‘DNN’, yield the best results. The striking feature of the ‘DMD-Re’-results is that as long as the spatial dimension is increased sufficiently, we can approximately achieve the same prediction result by appropriate reshaping the input data, regardless of the number of snapshots. This is a huge practical advantage since it guarantees consistent results dependent on only one parameter: m . Furthermore, the method does not need as many snapshots as the other methods to output accurate predictions. Prony KMD also gives reduction in the RMSE, but not as large reduction as the other KMD-like methods. ‘DNN’ shows improvements for most combinations, while the nearest neighbor method ‘NN’ only shows a minor reduction in RMSE, for cases where only a few snapshots were

used. The setting for those methods was as follows: both utilized a history of the 2000 last samples, and the number of neighbors was defined as the number of snapshots. It is possible that their results would improve if the number of neighbors was set differently. For example, one could set a tolerance for the maximum distance ϵ from the latest snapshot $\mathbf{y}_N$, and also require a lowest number of neighbors, and then increase the tolerance if too few were included. Alternatively, one can use the distance to the last snapshot as an indicator of the importance of each neighbor, and weight their future predictions accordingly [182, 186]. The trend predictor ‘Trend’, computed with (8.12), only depends on the two last measurements and does not depend on the spatial dimension. Hence, its prediction results are same regardless of the choices of m and N_s .

Here, we are interested in how much better than the naive predictor we are able to predict. To this end, the best RMSE values are saved for all methods, for a number of choices of L . Figure 8.8 shows the best RMSE results plotted against the prediction horizon determined by L , varied between 1 and 5 s. The results for $L = 2$ s are consistent with Fig. 8.7. The best results were achieved with DMD/Arnoldi-like KMD, ‘DMD-Re’, and ‘DNN’. For $L = 1$ s, they achieve about 65% reduction in RMSE compared to the naive predictor, and about 45% reduction for $L = 2$ s. The performance of all methods is close to the naive predictor for $L = 5$ s (an improvement of about 5% is still achieved for the most accurate methods). It is worth nothing that the average mutual information (8.3) decreased substantially in Fig. 8.3 (a) for $L = 5$ s. As previously shown, DMD and Arnoldi KMD can be considered as AR or VAR prediction if the matrices $\mathbf{A}$ and $\mathbf{C}$ are used for prediction, assuming snapshots of time-delayed measurements (8.9). Furthermore, a locally linear model such as ‘DNN’ with a “large” number of neighbors, can be viewed as an autoregressive model [4]. This explains the close resemblance between this method and the DMD-like methods.

Now, it is demonstrated how well the methods predict the increments:

$$\hat{du}_{N+i} := \hat{u}_{N+i} - u_N, \quad (8.37)$$

where $\hat{u}_{N+i}$ is the i -samples ahead predicted wind speed, and u_N the last wind speed observation. In Fig. 8.9, $\hat{du}_{N+i}$ are plotted against $du_{N+i} = u_{N+i} - u_N$ (the actual increment) for three of the methods. The points would fall on a line if the method predicts the increment accurately. From the figure, we see a big difference in one- and three-step ahead predictions. In particular, we see that ‘DMD-Re’ and ‘DNN’ again give the best prediction results. Note that the naive predictor would always predict $\hat{du}_{N+i} = 0$.

8.6 Conclusions

First, an evaluation of wind speed characteristics was conducted which indicated that wind speed fluctuations are represented well by a stochastic process. After that, several prediction methods were described and discussed, and relations between autoregression and KMD/DMD-like methods were clarified for time-shifted data. Following that, several prediction methods were evaluated with real wind speed measurements, and substantial

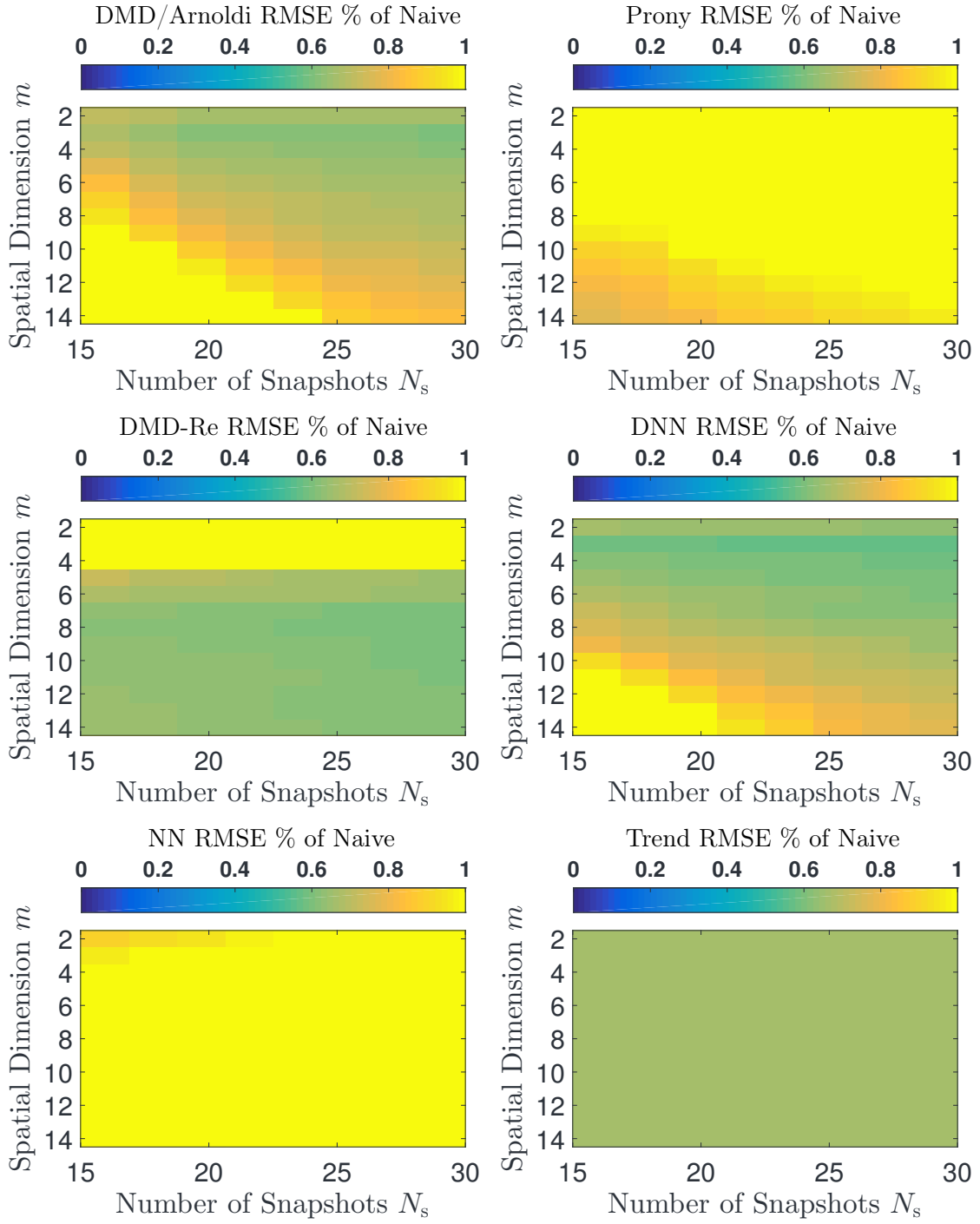


Figure 8.7: Improvements in RMSE of future predictions of wind speed with $L = 2$ s, with all the considered methods, compared with the naive predictor.

improvements in performance compared to the naive predictor were demonstrated for future predictions up to a few seconds ahead. Specifically, the performance of all methods became close to the naive predictor for predictions about 5-s ahead. The prediction results

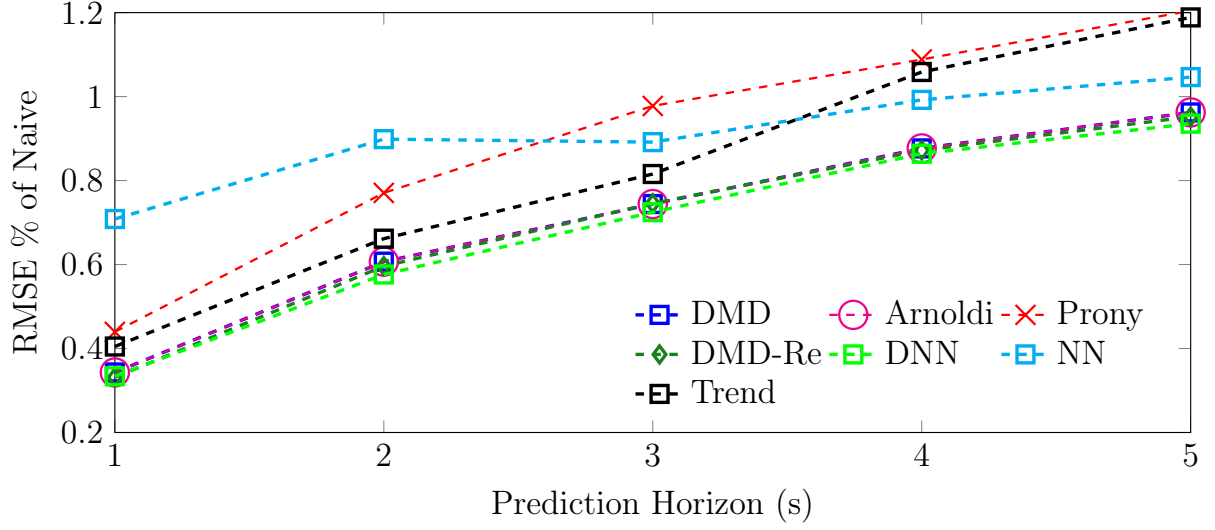


Figure 8.8: Best RMSE of prediction methods vs. prediction horizon.

with KMD/DMD-related methods were superior to other methods. These results further confirm the description of the wind speed fluctuations in terms of a stochastic process with non-zero autocorrelation, and that we are only able to exploit linear short-term correlations for our predictions. Despite being only to accurately predict few seconds into the future, the predictions could very well be incorporated in wind turbine control such as in a Model Predictive Control (MPC) to optimize the energy capture from wind, since even small improvements on the scale of few seconds are useful for WT operation and control [168,169].

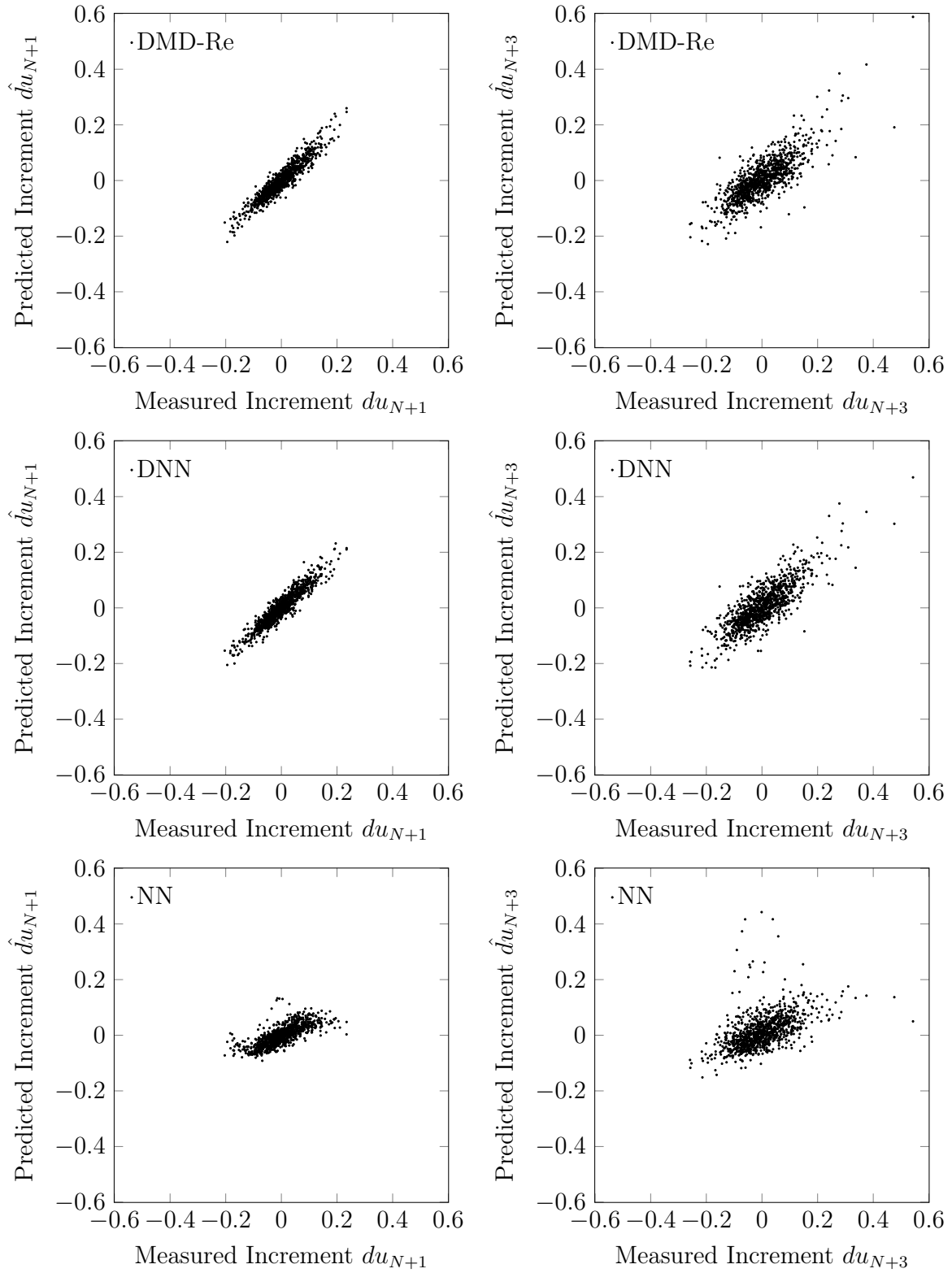


Figure 8.9: Predicted vs. measured increments for three of the considered prediction methods.

Appendices

8.A Cao's test for dimension and determinism

Here, we briefly outline a simple test of dimensionality and determinism of a time-series proposed by [173] and which is used in this chapter to evaluate wind speed characteristics. The test is based on two indexes. To introduce the first one, we start by defining the following measure to quantify the change in distance between two points in the space of time-delayed measurements, as the dimension m of the vector of measurements (8.2) increases by one:

$$a(n, m) = \frac{\|\mathbf{y}_n(m+1) - \mathbf{y}_{z(n,m)}(m+1)\|_\infty}{\|\mathbf{y}_n(m) - \mathbf{y}_{z(n,m)}(m)\|_\infty}, \quad (8.38)$$

where:

- $n = 1, \dots, N - mL$,
- $1 \leq z(n, m) \leq N - mL$, where $z(n, m)$ is an integer chosen such that $\mathbf{y}_{z(n,m)}(m)$ and $\mathbf{y}_n(m)$ are closest neighbors in m -dimensional space, and $z(n, m) \neq n$,
- $\|\cdot\|_\infty$ is the maximum norm: $\|\mathbf{u}\|_\infty = \max(|u_1|, |u_2|, \dots, |u_N|)$.

Now, the mean of $a(n, m)$ for each m is denoted by

$$E(m) = \frac{1}{N - mL} \sum_{n=1}^{N-mL} a(n, m), \quad (8.39)$$

and finally the first index is defined by

$$E1(m) = \frac{E(m+1)}{E(m)}. \quad (8.40)$$

The idea of $E1(m)$ is that when m reaches some value m_d , which can be thought of as the minimum embedding dimension related to the dimensionality of the dynamical system, $E1(m)$ should saturate at one. The reason is as follows. For small m , closeness in the space of time-delayed coordinates between two points (neighbors) can be due to the projection of the attractor on a smaller space, and as the attractor is *unfolded* by increasing m until $m = m_d$, no further unfolding occurs, and the neighbors are *true neighbors* [179], i.e. $E1(m+1) \approx E1(m)$. Furthermore, [173] defines the following quantity:

$$E^*(m) = \frac{1}{N - mL} \sum_{n=1}^{N-mL} |u_{n+mL} - u_{z(n,m)+mL}|, \quad (8.41)$$

and proposes a second index defined as follows:

$$E2(m) = \frac{E^*(m+1)}{E^*(m)}. \quad (8.42)$$

While $E1$ was proposed to determine the minimum embedding dimension of time-series, it was observed that it can vary slightly in practice even for random time-series [173]. Therefore, [173] proposed $E2$ to distinguish data generated by stochastic and deterministic systems. If the time-series is stochastic, $E2$ stays constant at one independent of m . This is because random signals do not have serial dependence (the next measurement does not depend on the previous one), and the choice of m should therefore not matter when computing (8.41). In such a case we have $E^*(m+1) \approx E^*(m)$, yielding $E2(m) \approx 1$.

Chapter 9

Partitioning power networks via KMD

9.1 Introduction

Power system dynamics are influenced by the loading and stress on critical transmission lines, which may be affected in highly variable systems incorporating large amounts of renewables. Quickly extracting information of the system's state from measurements following faults and taking appropriate action are considered important measures to ensure safe operation [31]. This can be conjectured to play an even greater role in future power systems. A strategy often proposed as a way to protect against large-scale blackouts and limit cascading failures in power systems is the so-called *controlled islanding strategy* [187–189]. Such a strategy aims to separate the network into several disjoint parts (islands) in a timely manner to avoid an impending blackout. Hence, the key part of the controlled islanding strategy is the partitioning problem, i.e. to determine where to appropriately separate the power network. There are many different ways to derive a suitable partition of the network: power flow analysis, graph theoretical analysis of the network, analysis of the underlying differential algebraic equations, data analysis, or any combination of them. Here, a new method utilizing the Koopman Mode Decomposition (KMD) is proposed for power network partitioning, based solely on measurements of bus-voltage angles.

Power network partitioning is based on the premise that a power network consists of tightly connected groups of buses (buses: interconnection points in the network), or clusters, that are loosely connected to each other. Existing methods can be classified into two types: (static) graph-based and dynamics-based. For the first one, the network partitioning problem has attracted interest from a purely graph theoretical point of view. If the network is preferably separated by identifying clusters loosely connected to one another, relevant cutsets (sets containing the separation points between sub-networks) can be identified by applying methods such as spectral graph theory. For the second one, a large power system exhibits inter-area oscillations (or modes) that cause groups of generators to oscillate in anti-phase. An *in-phase group* of generators is called coherent [190]. Partitioning of a target network can be derived based on the notion of coherency.

Numerous papers have proposed methods for partitioning power systems. In [188] a

strategy based on the concept of *slow-coherency* [70] is proposed where a search algorithm is used to determine optimal separation points. The same method is used in a case-study similar to the 2003 blackout in North America in [189], and it is shown that a controlled islanding strategy can improve the network’s fault response. The standard slow-coherency technique is extended to include load-buses in [191] to directly provide a partitioning of the network. In [192] a strategy is proposed based on so-called *ordered binary decision diagram*. In [193] a simulation study of the same strategy is given, and it is modified and extended to an effective real-time strategy in [194]. Spectral graph theory is an integral part in identifying network partitions in [195] where power flow and generator coherency constraints are included. Also, in [67], where the so-called graph Laplacian is evaluated with different weights on graph edges (transmission lines) and a dendrogram is used.

9.1.1 Proposed idea and contributions

The contribution of this work is the demonstration of a new methodology for power network partitioning based on applying KMD to bus-voltage angles following a system fault. Despite only using measurements, two relevant properties are derived from the proposed partitioning method. First, graph theoretical properties of a target network are captured. Second, similarities with slow-coherency are identified. It is thus expected that dynamically significant partitions are generated by the proposed method, which is shown to capture versatile features from data without models or simplifications.

The advantage of the proposed methodology compared with previously proposed methods is that the proposed methodology is model-free and only relies on actual data on the system dynamics in the network following a disturbance to identify a suitable partitioning of the network. On the other hand, graph theoretical methods or linearization techniques are per default constrained by modeling simplifications and/or potentially require substantial amounts of detailed power system data and parameters to provide accurate results. The main disadvantage of the proposed method is that it is completely dependent on the utilized algorithms and quality of data—thus for corrupt data or algorithmic instability or inaccuracy, the method could theoretically give a misleading result. Hence, further investigation on the reliability of the procedure is needed for a practical implementation.

The chapter is organized as follows. Section 9.2 explains the idea and methodology of slow coherency in power systems, and explains fundamental concepts in spectral graph theory. Section 9.3 presents the proposed data-driven partitioning via KMD. Section 9.4 demonstrates the methodology with two benchmark test systems. Section 9.5 elaborates on an alternative method to choose an appropriate number of groups of buses for the partitioning, and provides additional comparisons. Conclusions are given in Section 9.6.

9.2 Theoretical backgrounds

This section provides the theoretical backgrounds of this chapter. First, we briefly revisit the basic modeling of power systems. Then, we describe the two methods which

are compared to the proposed method: slow coherency which is a popular method for identifying coherent groups of generators or areas in a power system, and spectral graph theory which characterizes the network structure in terms of eigenvectors and eigenvalues of the so-called graph Laplacian.

9.2.1 Modeling of power system dynamics

In this chapter, power system equations are solved in MATLAB with the Power System Analysis Toolbox (PSAT) [98]. As already discussed in Chapters 1 and 3, power system dynamics are commonly described by nonlinear Differential Algebraic Equations (DAEs). If the loads in the power system are assumed to be of so-called constant impedance type, then system dynamics are described by only a set of differential equations. For the most simple generator model, given by the so-called swing equation, the two differential equations of each generator are:

$$\begin{aligned}\frac{d\delta_j}{dt} &= \omega_j, \\ \frac{d\omega_j}{dt} &= \frac{1}{M_j} (P_{m,j} - P_{e,j} - D_j\omega_j),\end{aligned}\tag{9.1}$$

where δ_j are rotor angles and ω_j angular frequency deviations (from nominal system frequency) of generators, respectively, $P_{m,j}$ is the mechanical power (here assumed to be constant), $P_{e,j}$ the electrical power which is a nonlinear function of δ_i ($i = 1, \dots, m_g$, and m_g denoting the number of generators), and D_j the generator's damping constant. The system (9.1) is called a reduced network model since the network information contained in $\mathbf{g}$ (see (1.1)) has been lost, and it can be written in its abbreviated form as

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}),\tag{9.2}$$

where $\mathbf{x} = [\delta_1, \dots, \delta_{m_g}, \omega_1, \dots, \omega_{m_g}]^\top$.

9.2.2 Slow coherency

One common way to identify coherency in the power system dynamics is by the concept of slow coherency [70,196–198]. Slow coherency has been incorporated in controlled islanding strategies [188,189,191,199] to identify coherent areas or groups of generators. The method is based on time-scale separation, which is derived by so-called singular perturbation of the system equations [200]. In the following, we outline the practical method of identifying slowly coherent generators according to [198]:

- (i) The second order, undamped (all $D_j = 0$), reduced network model (9.1) is linearized and used:

$$\frac{d^2\Delta\boldsymbol{\delta}}{dt} = \mathbf{A}_s\Delta\boldsymbol{\delta},\tag{9.3}$$

where $\Delta\boldsymbol{\delta}$ is the vector of linearized rotor angles.

- (ii) Eigenvalues and eigenvectors of $\mathbf{A}_s$ are calculated. The $(m_g \times r)$ matrix $\mathbf{V}_r$ is constructed, containing the r eigenvectors associated with the r slowest modal frequencies derived from the eigenvalues.
- (iii) Gaussian elimination is performed on $\mathbf{V}_r$ to identify the r most linearly independent rows corresponding to generators that now are considered as “reference generators”.
- (iv) $\mathbf{V}_r$ is re-ordered as $\mathbf{V}_r := [\mathbf{V}_1^\top \ \mathbf{V}_2^\top]^\top$, where the $(r \times r)$ matrix $\mathbf{V}_1$ contains rows of reference generators and $\mathbf{V}_2$ contains rows of remaining generators of $\mathbf{V}_r$.
- (v) The so-called grouping matrix $\mathbf{L}_d$ is calculated from $\mathbf{V}_1^\top \mathbf{L}_d^\top = \mathbf{V}_2^\top$. Generators are grouped based on the most positive entry in each row of $\mathbf{L}_d$, that corresponds to the strongest linear dependency on one of the reference generators.

To include all network buses and not only the generators’ internal buses, [191] proposed a method based on slow coherency to identify coherent groups by utilizing a variant of the DAE model (1.1):

$$\begin{aligned} \mathbf{M} \frac{d^2 \boldsymbol{\delta}}{dt} &= \mathbf{f}(\boldsymbol{\delta}, \mathbf{V}), \\ 0 &= \mathbf{g}(\boldsymbol{\delta}, \mathbf{V}), \end{aligned} \quad (9.4)$$

where $\mathbf{M}$ is an $m_g \times m_g$ diagonal matrix of inertia constants, and $\mathbf{V} = [\mathbf{V}_r^\top \ \mathbf{V}_x^\top]^\top$ is a $2N_b \times 1$ vector of the real $\mathbf{V}_r$ and imaginary parts $\mathbf{V}_x$, of bus voltages. The number of buses is N_b . The DAE system (9.4) is then linearized around its equilibrium point which yields,

$$\begin{bmatrix} \frac{d^2 \Delta \boldsymbol{\delta}}{dt} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{J}_A & \mathbf{J}_B \\ \mathbf{J}_C & \mathbf{J}_D \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{\delta} \\ \Delta \mathbf{V} \end{bmatrix}, \quad (9.5)$$

where $\mathbf{J}_A, \mathbf{J}_B, \mathbf{J}_C, \mathbf{J}_D$ denote the Jacobian matrices of the DAEs in (9.4) with respect to $\boldsymbol{\delta}$ and $\mathbf{V}$. From (9.5), if $\mathbf{J}_D^{-1}$ exists, we construct the so-called state matrix $\mathbf{A}_s$ by

$$\mathbf{A}_s = \mathbf{J}_A - \mathbf{J}_B \mathbf{J}_D^{-1} \mathbf{J}_C. \quad (9.6)$$

Note that this corresponds to the matrix $\mathbf{A}_s$ in the previously outlined slow-coherency procedure. With (9.6) we can write the eigenvalue problem as

$$\mathbf{A}_s \mathbf{u}_i = \nu_i \mathbf{u}_i, \quad (9.7)$$

where ν_i is the eigenvalue and $\mathbf{u}_i$ the corresponding eigenvector (mode). Utilizing (9.5), the eigenvalue problem can instead be rewritten as

$$\begin{bmatrix} \mathbf{J}_A - \nu_i \mathbf{I} & \mathbf{J}_B \\ \mathbf{J}_C & \mathbf{J}_D \end{bmatrix} \begin{bmatrix} \mathbf{u}_i \\ \mathbf{z}_i \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad (9.8)$$

where $(\nu_i, \mathbf{u}_i)$ is an eigenpair of $\mathbf{A}_s$, too. We also have vectors $\mathbf{z}_i$ that quantify the modal shapes of $\mathbf{u}_i$ in the power system network voltages. The vectors $\mathbf{z}_i$ are then used to group network buses as in the previously outlined slow-coherency methodology. Refer to [191] for the exact procedure.

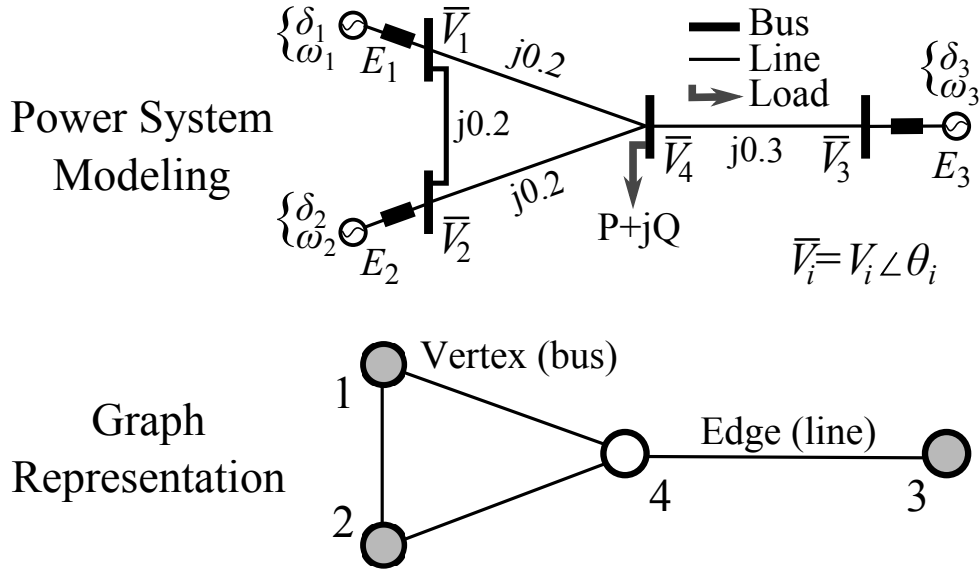


Figure 9.1: Simple test system with three generator buses and one load bus. Two representations are shown; the upper figure depicts typical power system modeling and the lower one depicts a graph representation with vertices as generator buses (filled circles) and one load bus (hollow circle) connected by edges (transmission lines).

9.2.3 Spectral graph theory for network partitioning

Here, fundamental concepts in spectral graph theory are introduced based on [69]. Spectral graph theory is the study of networks in terms of eigenvalues and eigenvectors of matrices associated with the network. A graph is given by the tuple $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$, where $\mathcal{V}$ is the set of vertices, $\mathcal{E}$ the set of edges, and $w : \mathcal{E} \rightarrow \mathbb{R}$ the weight function of each edge (e.g. see Fig. 9.1). For a power network with m buses and an arbitrary number of transmission lines, $\mathcal{V}$ and $\mathcal{E}$ correspond to the sets of buses labeled by integers from 1 to m and transmission lines, respectively. The notation (i, j) is used to represent the edge connecting the vertices i and $j \in \mathcal{V} = \{1, \dots, m\}$. The weight function w usually assigns every edge unity ($= 1$) which is the setting here, or weight according to line admittance, or steady-state power flow on the line [67]. A graph is said to be *connected* if there exists a path between any two vertices. The *adjacency matrix* $\mathbf{A}$ for the graph $\mathcal{G}$ is defined as

$$[\mathbf{A}]_{ij} := \begin{cases} w(i, j), & \text{if } (i, j) \in \mathcal{E}. \\ 0, & \text{otherwise.} \end{cases} \quad (9.9)$$

By counting the *degree* of vertex i as $d_i := \sum_{j=1}^m [\mathbf{A}]_{ij}$, the so-called *graph Laplacian* $\mathbf{L}$ is defined as the following symmetric matrix:

$$\mathbf{L} := \text{diag}(d_1, d_2, \dots, d_m) - \mathbf{A}. \quad (9.10)$$

Here, we are interested in looking at the eigenstructure of this matrix. For a connected graph, $\mathbf{L}$ has one zero eigenvalue, and all of the other eigenvalues are positive. In the

following, they are listed based on increasing magnitude, that is, $\lambda_1 = 0 < \lambda_2 < \dots < \lambda_m$.

The eigenstructure of $\mathbf{L}$ has been exploited for graph partitioning. In [201] the so-called *algebraic connectivity* is defined and associated with the pair of the 2nd eigenvalue and associated eigenvector, $(\lambda_2, \mathbf{V}_2)$, where $\mathbf{V}_2$ is known as the *Fiedler vector*. For a connected graph $\mathcal{G}$, a partitioning of $\mathcal{G}$ into two disjoint parts, called the *graph bisectioning*, can be determined by evaluating the sign of each component in $\mathbf{V}_2$ as follows:

$$i \in \mathcal{V}_1 \text{ if } [\mathbf{V}_2]_i \geq 0, \quad i \in \mathcal{V}_2 \text{ if } [\mathbf{V}_2]_i < 0, \quad (9.11)$$

where $\mathcal{V}_1$ and $\mathcal{V}_2$ are the two disjoint sets of vertices. For a multilevel partitioning using Laplacian eigenvectors, we can consider two simple approaches. The first one is to apply (9.11) iteratively on the increasingly smaller partitions until a desired amount of partitions is obtained. The second one described in [202] is to consider a matrix $\mathbf{W} := [\mathbf{V}_2, \dots, \mathbf{V}_{k>2}]$, where the rows of $\mathbf{W}$ can be interpreted as points in a $(k-1)$ -dimensional space. One can now apply a k -means clustering method [203] to obtain k partitions. Here the function `kmeans` in MATLAB is used for this.

9.3 Proposed data-driven partitioning method

This section explains the notion of coherency in Koopman Modes (KM), and outlines the proposed methodology to partition a power network into several disjoint parts by applying KMD to bus-angle dynamics following a disturbance.

9.3.1 Coherency in Koopman modes

The notion of coherency in KMs is explained in the following. First, the Arnoldi-like KMD outlined Chapter 2 is utilized, and applied to bus-angle snapshots:

$$\{\boldsymbol{\theta}_0, \dots, \boldsymbol{\theta}_N\},$$

where the snapshots $\boldsymbol{\theta}_i \in \mathbb{R}^m$ are sampled with a fixed sampling period. Application of KMD yields the following decomposition of the N first samples in terms of $(\tilde{\lambda}_i, \tilde{\mathbf{v}}_i)$:

$$\boldsymbol{\theta}_k = \sum_{i=1}^N \tilde{\lambda}_i^k \tilde{\mathbf{v}}_i, \quad k = 0, \dots, N-1, \quad (9.12)$$

where $\tilde{\mathbf{v}}_i$ and $\tilde{\lambda}_i$ are called the KMs and Koopman Eigenvalues (KE), respectively. The following is based on [19], which first proposed data-driven coherency identification via KMD. Suppose a set of KMs $\{\mathbf{v}_1, \dots, \mathbf{v}_N\}$ is obtained such as in (9.12), from application to a window of bus angle dynamics in power system with m buses, see Fig. 9.2(a). First, consider the j -th element of the i -th KM, $[\mathbf{v}_i]_j = A_{ij} \angle \alpha_{ij}$, $i \in \mathcal{N}$, $j \in \{1, \dots, m\}$, where A_{ij} is called the *amplitude factor* and α_{ij} the *initial phase*. In a *coherent group* of buses for a given KM $\mathbf{v}_i$, all buses are oscillating with a small phase difference between them.

More rigorously, according to [13] by taking a positive constant ϵ_α , which we call the phase tolerance, and choosing two measurement locations (buses) $k, l \in \{1, \dots, m\}$, we define the following condition for phase coherency:

$$|[\boldsymbol{\alpha}_i]_k - [\boldsymbol{\alpha}_i]_l| < \epsilon_\alpha, \quad (9.13)$$

where $\boldsymbol{\alpha}_i$ is the initial phase vector for $\mathbf{v}_i$. If (9.13) is fulfilled, then the buses k and l are called ϵ_α -coherent for the i -th KM, see Fig. 9.2(b). When coherency for a collection of dominant KMs $\{\mathbf{v}_1, \dots, \mathbf{v}_d\}$ with associated initial phases $[\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_d]$ is concerned, then (9.13) has to hold for all $\boldsymbol{\alpha}_i$ simultaneously which can be expressed as

$$|[\boldsymbol{\alpha}_i]_k - [\boldsymbol{\alpha}_i]_l| < \epsilon_\alpha, \quad \forall i \in \{1, \dots, d\}. \quad (9.14)$$

The coherency condition for multiple KMs given by (9.14) is exploited at the end of this chapter in a proposed multi-way partitioning strategy. Related to coherency in KMs, we define the sum of Complex Conjugate (CC) (oscillatory) KEs scaled by their KMs by

$$\begin{aligned} [\mathbf{S}_{\text{CC},i}]_j^k &= \lambda_i^k [\mathbf{v}_i]_j + (\lambda_i^c)^k [\mathbf{v}_i]_j^c \\ &= 2A_{ij} |\lambda_i|^k \cos(\alpha_{ij} + k\phi_i), \quad \lambda_i = |\lambda_i| \angle \phi_i, \end{aligned}$$

where setting $k = 0$ (initial time) yields

$$[\mathbf{S}_{\text{CC},i}]_j = 2A_{ij} \cos(\alpha_{ij}), \quad (9.15)$$

which provides a quantitative measure on how the oscillatory KM pair is excited at each measurement location. The real-valued vector $\mathbf{S}_{\text{CC},i}$ will be referred to as the i -th *spatial shape* of oscillatory KMs.

9.3.2 Outline of KMD-based partitioning strategy

First of all, two assumptions are made:

- (i) Graph information $\mathcal{G}$ of a target power network is given.
- (ii) Bus-angle dynamics are observed for *every* bus.

The network's configuration can be obtained from the state estimator's topology processor [204], and hence (i) is reasonable. Assumption (ii) is based on two reasons. First, within foreseeable future we can expect an extensive deployment of synchronized measurement units. Second, with the dynamic state estimator, full observability is achievable even with a limited amount of PMUs. In fact, placing PMUs at about one third of the total number of buses is generally sufficient to achieve full observability [58]. Until such a system is in place, it is possible to construct a state estimator utilizing both SCADA and PMU measurements [204].

The proposed data-driven partitioning methodology demonstrated in this chapter is outlined in the following. Its pseudocode is provided in Algorithm 1.

1. Under the stated assumptions (i) and (ii), consider the finite-time data on bus-angle dynamics under uniform sampling, given by

$$\{\boldsymbol{\theta}_0, \dots, \boldsymbol{\theta}_N\},$$

where $\boldsymbol{\theta}_k = (\theta_{0,k}, \dots, \theta_{m,k})^\top$ is the m -dimensional snapshot of bus-angles at the k -th time instance and the number of available snapshots corresponds to $N + 1$ (see Fig. 9.2(a)). By applying the Arnoldi-like algorithm to the finite-time data, N pairs $(\tilde{\lambda}_i, \tilde{\mathbf{v}}_i)$ ($i \in \mathcal{N}$) of KEs and KMs are obtained (KMD).

2. A set of *dominant* KMs, denoted by $\mathcal{N}_d \subset \mathcal{N}$, is identified by arranging them in descending order of GR and norm (DominantKMD).
3. For every pair $(\tilde{\lambda}_i, \tilde{\mathbf{v}}_i)$ of dominant KMs, the initial phase vector $\boldsymbol{\alpha}_i := (\alpha_{i1}, \dots, \alpha_{im})^\top$ is calculated (InitialPhaseVector).
4. By plotting all the components of $\boldsymbol{\alpha}_i$ on the unit circle, the phase coherency in bus-angles is detected. Below, we use the same notation α_{ij} to represent its plotted point. When a k -means clustering algorithm is applied to the m points on the unit circle, k center points called *centroids* C_i^ℓ ($\ell \in \mathcal{K}$) and associated *clusters* (subsets of buses) $\mathcal{A}_i^\ell$ are located (kmeansClustering).
5. For each $\mathcal{A}_i^\ell$, the distance between a contained point α_{ij} and its centroid C_i^ℓ is denoted by $\text{dist}(\alpha_{ij}, C_i^\ell)$ with the standard Euclidean norm. Thus, we compute the *Cluster Coefficient* CIC for each cluster $\mathcal{A}_i^\ell$ as

$$CIC_i^\ell := \frac{\sum_{j \in \mathcal{A}_i^\ell} \text{dist}(\alpha_{ij}, C_i^\ell)}{\#(\mathcal{A}_i^\ell)}, \quad (9.16)$$

where $\#(\mathcal{A}_i^\ell)$ stands for the number of elements belonging to $\mathcal{A}_i^\ell$ (ComputeCIC). A small CIC indicates a tightly clustered group of buses, and is deemed “small” in comparison to other dominant KMs and previous data.

6. The number of dominant KMs is now decreased from $\#(\mathcal{N}_d)$ such that only KMs with small CICs and clear phase separation θ_{sep} between centroids are included (see Fig. 9.2(b); SelectBestPartition).
7. Since from assumption (i) the complete network structure is given, the cutsets are identified by first assigning an index to each bus that represents one of the k groups and second identify the lines connecting buses of different groups (IdentifyCutset). In the case that a singular bus or few buses are detected incoherent with surrounding buses (see Fig. 9.2(c)), the so-called *scattered buses* may be discarded depending on the preferred partitioning setting.

Algorithm 1 Data-Driven Partitioning via KMD

Require: $\{\theta_0, \dots, \theta_N\}$: bus-angle dynamic data; $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$: power network graph;
 k : order of k -means clustering algorithm

Ensure: Cutset

- 1: $\{(\tilde{\lambda}_i, \tilde{\mathbf{v}}_i); i \in \mathcal{N}\} \leftarrow \text{KMD}(\boldsymbol{\theta})$
 - 2: $\mathcal{N}_d \leftarrow \text{DominantKMD}(\tilde{\lambda}, \tilde{\mathbf{v}})$
 - 3: $\{\boldsymbol{\alpha}_i; i \in \mathcal{N}_d\} \leftarrow \text{InitialPhaseVector}(\tilde{\lambda}, \tilde{\mathbf{v}}, \mathcal{N}_d)$
 - 4: $\{(C_i^\ell, \mathcal{A}_i^\ell); i \in \mathcal{N}_d, \ell \in \mathcal{K}\} \leftarrow \text{kmeansClustering}(\boldsymbol{\alpha})$
 - 5: $\{ClC_i^\ell; i \in \mathcal{N}_d, \ell \in \mathcal{K}\} \leftarrow \text{ComputeClC}(C, \mathcal{A})$
 - 6: $i^* \leftarrow \text{SelectBestPartition}(ClC)$
 - 7: $\text{IdentifyCutset}(\tilde{\lambda}_{i^*}, \tilde{\mathbf{v}}_{i^*}, \mathcal{G})$
-

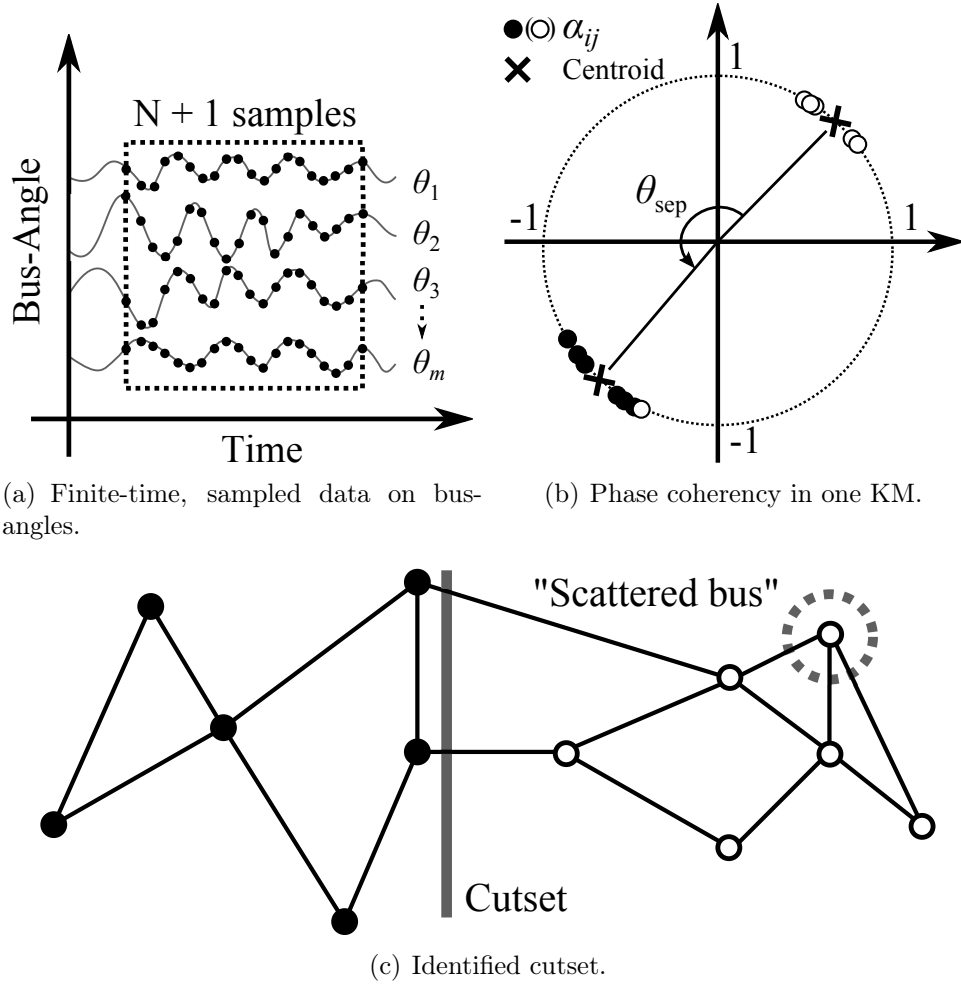


Figure 9.2: Illustration of data-driven network partitioning via Koopman mode analysis described in Algorithm 1.

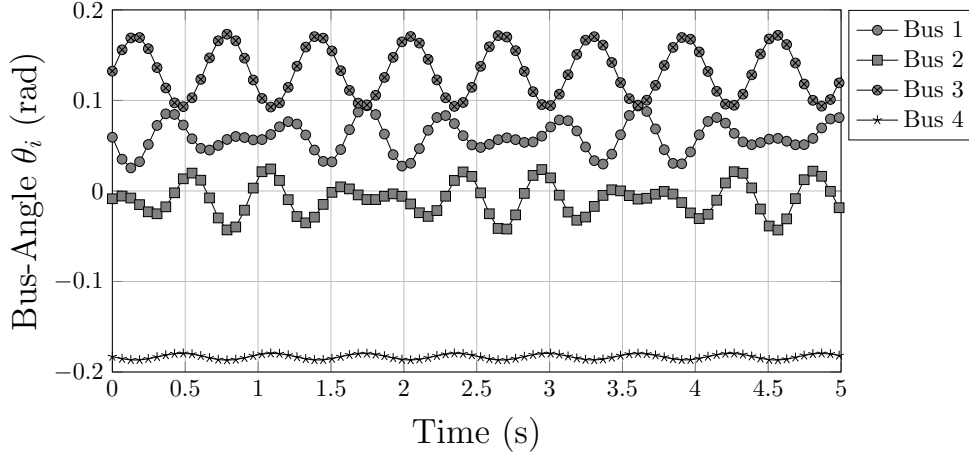


Figure 9.3: Dynamics of bus-angles θ_i for the simple test system following a small disturbance of initial generator speed ω_i^0 for $i = 2, 3$.

9.4 Demonstration

In this section, the data-driven partitioning method is demonstrated for both a very simple system, and then for a larger benchmark system. A connection between partitioning results obtained with KMD and spectral graph theory is also analytically revealed.

9.4.1 Benchmark systems and simulation setting

Two benchmark systems are utilized in this chapter, and are both simulated using PSAT by the DAEs in (1.1). The number of buses for each system is m and the number of generators is n . The first test system is depicted in Fig. 9.1 and is a simple four-bus system consisting of three generator buses and one load bus. The second one is the IEEE 118-bus test system [205] depicted in Fig. 9.6 and is simulated with 19 generators whose parameters are chosen according to [193]. The setting of the numerical simulations is as follows. Throughout the simulations, the sampling frequency $f_s = 60$ Hz was used. Bus-voltage angles θ were sampled and observed with respect to the mean-angle $\bar{\theta}_k$, which is similar to the Center-Of-Inertia (COI) [7] and is defined as follows: for every time k ,

$$\bar{\theta}_k := \frac{1}{m} \sum_{j=1}^m \theta_{j,k}. \quad (9.17)$$

According to simulations, this produces a moving reference similar to the COI reference frame for generator rotor angles. Consequently, a change of reference frame from the proposed one to COI yields similar results on partitioning. Also, by using the proposed reference frame, no additional measurements from generators are required.

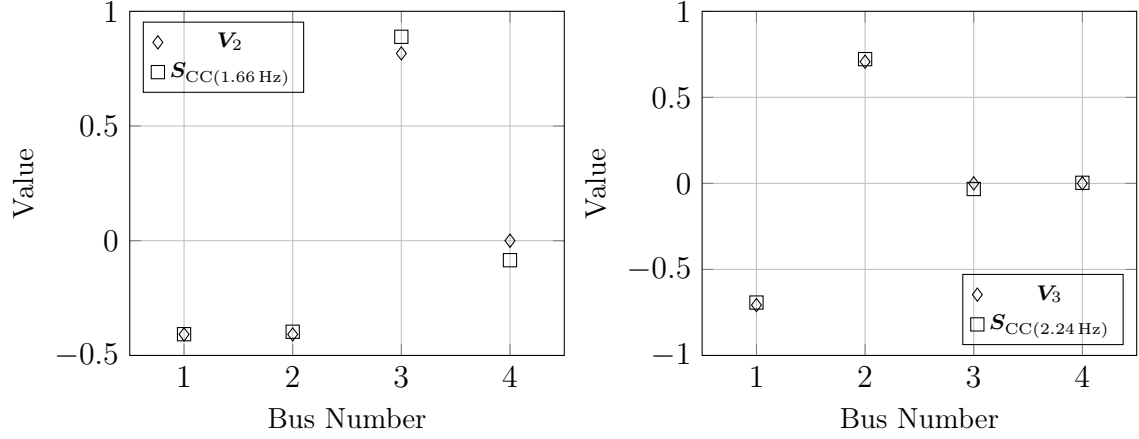


Figure 9.4: Comparison between eigenvectors $\mathbf{V}_2$, $\mathbf{V}_3$ of the unweighted Laplacian, and spatial shapes $\mathbf{S}_{CC(1.66 \text{ Hz})}$, $\mathbf{S}_{CC(2.24 \text{ Hz})}$ of the dominant KMs identified from dynamics shown in Fig. 9.3.

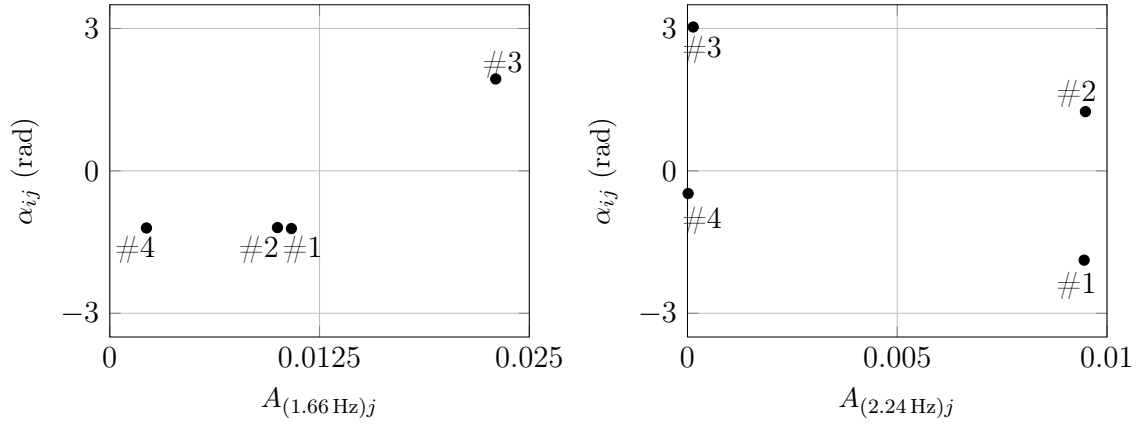


Figure 9.5: Initial phase α_{ij} vs. amplitude factor A_{ij} plots for $\mathbf{v}_{1.66 \text{ Hz}}$ and $\mathbf{v}_{2.24 \text{ Hz}}$, where $\#i$ beside a point denotes the bus number.

9.4.2 Simple test system

The proposed methodology is demonstrated by first considering a small test system (Fig. 9.1) which comprises three generator buses and one load bus. Parameters of the test system are as follows: base power S_b and generator ratings S_g are set to 100 MVA, nominal frequency $f_n = 60 \text{ Hz}$ and a constant impedance load $S_L = P_L + jQ_L = 3 + j0.05 \text{ p.u.}$. Generator parameters are chosen in a simple manner: transient reactances $x'_{d,i} = 0.2 \text{ p.u.}$, inertia constants $M_i = 7 \text{ s}$, damping $D_i = 0.1$, for all generators and mechanical power $P_m = \{P_{m,1}, P_{m,2}, P_{m,3}\} = \{1.5, 0.5, 1\} \text{ p.u.}$ Two complex conjugate oscillatory mode pairs with the frequencies 1.64 Hz and 2.27 Hz are identified from linearization of (9.4).

For initial generator frequencies ω_i^0 , a perturbation is initiated as $\omega_i^0 + \Delta\omega_i$, $i = 2, 3$

with $\Delta\omega_i = 0.005$ p.u., and the resulting dynamics are shown in Fig. 9.3. KMD is applied to 5 s of disturbance data to give 300 KMs. Two oscillatory KMs with frequencies 1.66 Hz and 2.24 Hz, respectively, are identified as dominant in the decomposition. Note that the frequencies of the identified KMs are almost identical to the linear modes. Their spatial shapes denoted by $\mathbf{S}_{\text{CC}(1.66 \text{ Hz})}$ and $\mathbf{S}_{\text{CC}(2.24 \text{ Hz})}$ are displayed in Fig. 9.4 together with $\mathbf{V}_2$ and $\mathbf{V}_3$, derived from the graph Laplacian of the test system. The spatial shapes $\mathbf{S}_{\text{CC}}$ are scaled with appropriate factors for the sake of comparison. The values of elements quantify the connectivity. For $\mathbf{S}_{\text{CC}(1.66 \text{ Hz})}$, buses 1 and 2 are strongly connected (coherent) and separated from bus 3. By inspecting the initial phases of $\mathbf{v}_{1.66 \text{ Hz}}$ in Fig. 9.5 it is concluded that bus 4 is also coherent with buses 1 and 2. Thus, the partition according to the 1.66 Hz KM leads to two groups of buses: $\{1, 2, 4\}$ and $\{3\}$. This is intuitive by inspection of the network's structure. The bus-number dependence of elements are close between $\mathbf{S}_{\text{CC}(1.66 \text{ Hz})}$ (or $\mathbf{S}_{\text{CC}(2.24 \text{ Hz})}$) and $\mathbf{V}_2$ (or $\mathbf{V}_3$). From this observation, it can be speculated that the spectral property of the network is extracted from sampled dynamics of bus-angles via the spectrum of the Koopman operator.

9.4.3 An analytical evaluation

The previous section numerically showed a prominent similarity between the Laplacian eigenvectors and spatial shapes of Koopman modes and hence in the dynamics of the simple test system. Here, we theoretically clarify the connection between KMD and spectral graph theory in terms of bus-angle dynamics in a general power network model.

Now we introduce the equations of motion for a general power network. Let us denote by $\mathcal{V}_G$ the set of buses with synchronous generators and by $\mathcal{V}_L$ the set of buses with no generator. For a generator connected to bus $i \in \mathcal{V}_G$, the voltage behind reactance is denoted by $E_i \angle \delta_i$, and the bus-voltage by $V_i \angle \theta_i$. The differential equations in the DAE model (1.1) are as follows: for $i \in \mathcal{V}_G$,

$$\frac{d\delta_i}{dt} = \omega_i, \quad M_i \frac{d\omega_i}{dt} = P_{m,i} - P_{e,i} - D_i \omega_i, \quad (9.18)$$

with the electrical output power $P_{e,i}$ given by

$$P_{e,i} = \frac{E_i V_i}{x'_{d,i}} \sin(\delta_i - \theta_i), \quad (9.19)$$

where $x'_{d,i}$ is the transient reactance. According to the standard argument of short-term rotor stability [7], constant voltages and active/reactive power decoupling are assumed. The algebraic variables $\mathbf{y}$ coincide with the bus-angles $\boldsymbol{\theta}$. Thus, the active power part of $\mathbf{g}$ is considered: for $i \in \mathcal{V} = \mathcal{V}_G \cup \mathcal{V}_L$,

$$0 = \sum_{j \in \mathcal{V} \setminus \{i\}} V_i V_j |Y_{ij}| \cos(\theta_j - \theta_i + \phi_{ij}) - P_i(\mathbf{x}, \theta_i), \quad (9.20)$$

where $|Y_{ij}| \angle \phi_{ij}$ is the admittance of lines connecting bus i and j (if no line exists, we regard its modulo as zero) and $P_i(\mathbf{x}, \theta_i)$ represents injected power ($P_{e,i}$) for $i \in \mathcal{V}_G$ or

constant consumed power ($P_{L,i}$) for $i \in \mathcal{V}_L$. Now, following [206], we use the singular perturbation technique for (9.4) and introduce a sufficiently small, positive parameter ϵ as follows:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \boldsymbol{\theta}), \quad \epsilon \frac{d\boldsymbol{\theta}}{dt} = \mathbf{g}(\mathbf{x}, \boldsymbol{\theta}). \quad (9.21)$$

By introducing the new independent variable $\tau := t/\epsilon$, we have

$$\frac{d\mathbf{x}}{d\tau} = \epsilon \mathbf{f}(\mathbf{x}, \boldsymbol{\theta}), \quad \frac{d\boldsymbol{\theta}}{d\tau} = \mathbf{g}(\mathbf{x}, \boldsymbol{\theta}). \quad (9.22)$$

By taking the $\epsilon \rightarrow 0$ limit, the boundary layer system is constructed as

$$\frac{d\boldsymbol{\theta}}{d\tau} = \mathbf{g}(\mathbf{x}, \boldsymbol{\theta}), \quad \mathbf{x} \text{ fixed}, \quad (9.23)$$

which captures the dynamical feature of (fast) bus-angle dynamics in an analytical, self-consistent manner. Here, by linearizing (9.23) around a feasible equilibrium point $(\mathbf{x}^*, \boldsymbol{\theta}^*)$ of the DAE system (9.4)², we derive

$$\frac{d\boldsymbol{\theta}}{d\tau} = (\mathbf{L}_{P,\theta} + \mathbf{K}_{P_e,\theta}) \Delta \boldsymbol{\theta}, \quad (9.24)$$

with $\mathbf{L}_{P,\theta}$ defined as

$$[\mathbf{L}_{P,\theta}]_{i,j} := \begin{cases} \sum_{\ell \in \mathcal{V} \setminus \{i\}} V_i V_\ell |Y_{i\ell}| \sin(\theta_\ell^* - \theta_i^* + \phi_{i\ell}), & i = j, \\ -V_i V_j |Y_{ij}| \sin(\theta_j^* - \theta_i^* + \phi_{ij}), & i \neq j, \end{cases} \quad (9.25)$$

where $\mathbf{L}_{P,\theta}$ is a graph Laplacian matrix of linearized power flow for every bus where the state variables $(\mathbf{x})$ are viewed as constant. The matrix $\mathbf{K}_{P_e,\theta}$ is a diagonal matrix and represents the linearized injected power from generators given as follows:

$$[\mathbf{K}_{P_e,\theta}]_{i,i} := \begin{cases} \frac{E_i V_i}{x'_{d,i}} \cos(\delta_i^* - \theta_i^*), & i \in \mathcal{V}_G, \\ 0, & i \in \mathcal{V}_L, \end{cases} \quad (9.26)$$

which can be viewed as a perturbation to $\mathbf{L}_{P,\theta}$. The magnitude of perturbation becomes small in particular if the number of buses is much larger than the number of generators. Equation (9.24) determines the behavior of bus-angle dynamics close to the equilibrium point and is influenced by the graph Laplacian $\mathbf{L}_{P,\theta}$. For the data-driven partitioning method based on bus-angle dynamics proposed here, it makes sense to compare the approach based on the Koopman operator with the network analysis based on spectral graph theory.

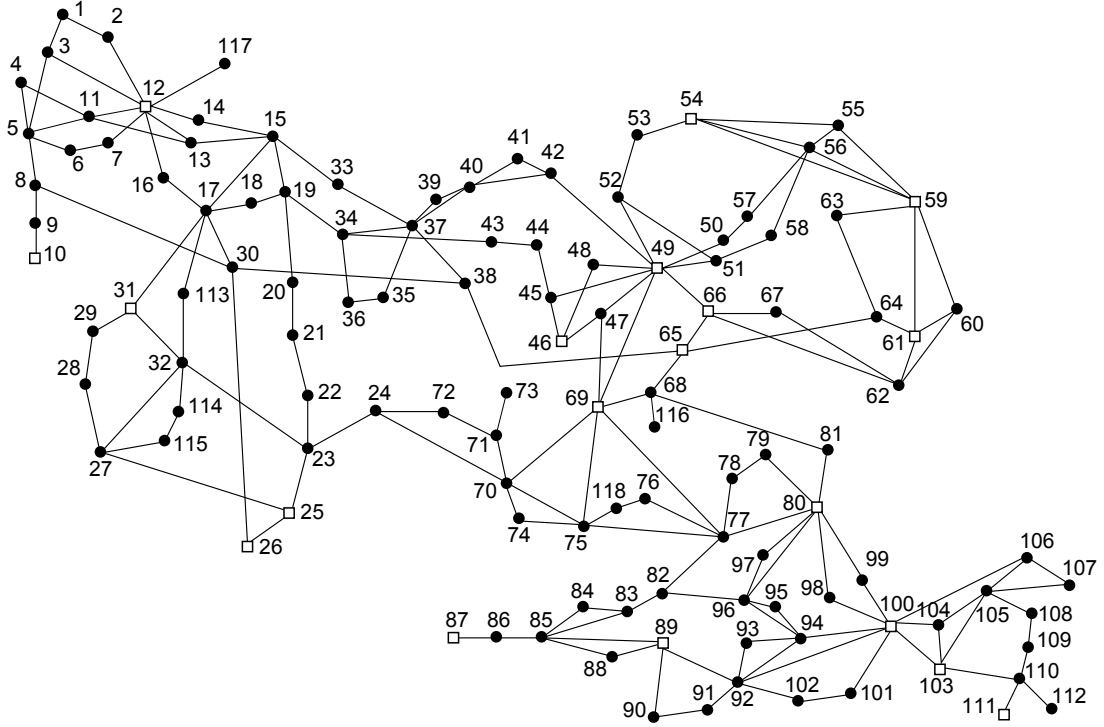


Figure 9.6: IEEE 118-bus test system [205]. The dynamics of the system are simulated with the classical model of 19 generators. Generator buses are indicated by hollow squares, and filled circles represent load and intermediate buses.

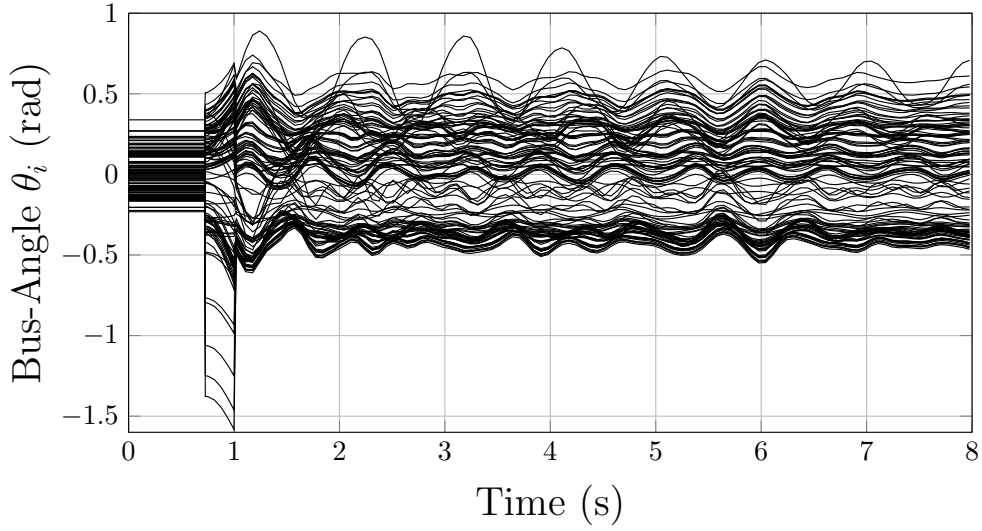
9.4.4 IEEE 118-bus test system

Next, the data-driven partitioning method is applied to a larger test system, the IEEE 118-bus test system (see Fig. 9.6). To obtain finite-time data on bus-angle dynamics, the system is perturbed by three-phase-faults according to following two cases:

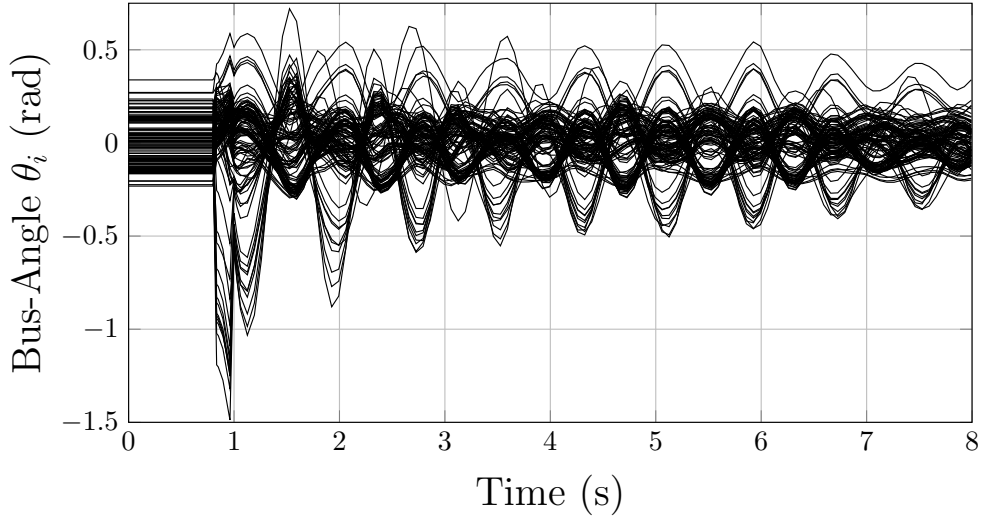
- (i) Fault applied to bus 17 with a clearing time $t_c = 280$ ms, slightly below the critical clearing time t_{cc} .
- (ii) Fault applied to bus 100 with a clearing time $t_c = 180$ ms, slightly below its t_{cc} .

Load flow parameters used are from [205] and generator parameters are identical to [193]. The dynamic responses of all bus-angles for the two cases (i) and (ii) are shown in Fig. 9.7. For each case, KMD is applied to the post-fault data on bus-angles during [1 s, 8 s] and provides 420 KMs. Dominant KMs identified in the decompositions are listed in Table 9.1. The 0 Hz KMs with unit GR appear for both cases and correspond to the time-averaged values of the bus-angle measurements in the system. Three KM pairs for each fault case are chosen and used for partitioning the system. The chosen KMs have large norms among

²An equilibrium point $(\mathbf{x}^*, \boldsymbol{\theta}^*)$ of the DAE system (9.4) is said to be *feasible* if the Jacobian matrix of $\mathbf{g}$ with respect to $\mathbf{y}$ is regular inside a neighborhood of $(\mathbf{x}^*, \boldsymbol{\theta}^*)$.



(a) Fault at bus 17 (Case (i)).



(b) Fault at bus 100 (Case (ii)).

Figure 9.7: Dynamics of bus-angles θ_i in the IEEE 118-bus test system for (a) fault at bus 17 (Case (i)) and (b) fault at bus 100 (Case (ii)).

the dominant oscillatory KMs and all show small ClCs and clear separation θ_{sep} . Initial phase vectors for chosen KMs for both disturbance cases are depicted in Fig. 9.8. For Case (i), KM-3 (1.24 Hz) clearly displays 3 large groups of buses, whereas remaining KMs for both the disturbances predominantly contain 2 large groups of buses in anti-phase. Algorithmically speaking, this is identified by applying the partitioning algorithm with first $K = 2$ and then $K = 3$ and note that the obtained ClCs decrease for the case with $K = 3$ while a clear θ_{sep} is maintained. Thus for the KM-3, 3-means clustering is applied, otherwise the 2-means one is used. According to the algorithm, dominant KMs provide partitioning of the system into two or more sub-systems (depending on cohesiveness of

Table 9.1: Dominant Koopman modes obtained for the data on voltage angle dynamics shown in Fig. 9.7. Rows marked in gray indicate KMs that have been chosen for the network partitioning.

Case #	No. j	GR $ \tilde{\lambda}_j $	Freq. [Hz] $\text{Im}[\ln \tilde{\lambda}_j]/(2\pi T_s)$	Norm $\ \tilde{\mathbf{v}}_j\ $	ClC ClC_j^ℓ	θ_{sep} (deg)
	1	1	0	3.29	0, 0	180
	2	0.9977	± 1.04	0.28	0.009, 0.008	177
	3	0.9976	± 1.24	0.25	0.036, 0.018, 0.028	77, 103, 180
(i)		0.9975	± 3.29	0.05	0.130, 0.074	173
		0.9970	± 0.21	0.02	0.091, 0.004	185
		0.9965	± 2.59	0.03	0.164, 0.246	168
		0.9959	± 2.37	0.16	0.081, 0.055	179
	4	0.9950	± 2.02	0.19	0.039, 0.043	180
	1	1	0	1.02	0, 0	180
	2	0.9992	± 1.00	0.17	0.042, 0.104	167
	3	0.9976	± 1.40	0.33	0.007, 0.044	178
(ii)		0.9970	± 0.83	0.1	0.079, 0.086	173
		0.9970	± 1.59	0.1	0.128, 0.049	176
	4	0.9966	± 1.23	1.73	0.007, 0.004	180
		0.9966	0	0.27	0, 0	180
		0.9965	2.46	0.04	0.016, 0.011	179

coherent buses). Partitions for the chosen dominant KMs based on phase coherency are given in Figs. 9.9 (a) and 9.9 (b). For Case (i), KM-3 provides a partition that shares the cutset obtained from spectral graph bisectioning but contains additional cutsets. KM-2 for both the disturbance cases are almost identical in terms of provided partitioning and frequency. A comparison between the Laplacian eigenvector $\mathbf{V}_2$ and the spatial shapes $\mathbf{S}_{\text{CC}(1.04 \text{ Hz})}$ and $\mathbf{S}_{\text{CC}(1.00 \text{ Hz})}$ of KM-2 is given in Fig. 9.10 (spatial shapes for buses 86 and 87 are not shown due to large magnitudes compared to other buses). It is clear that the

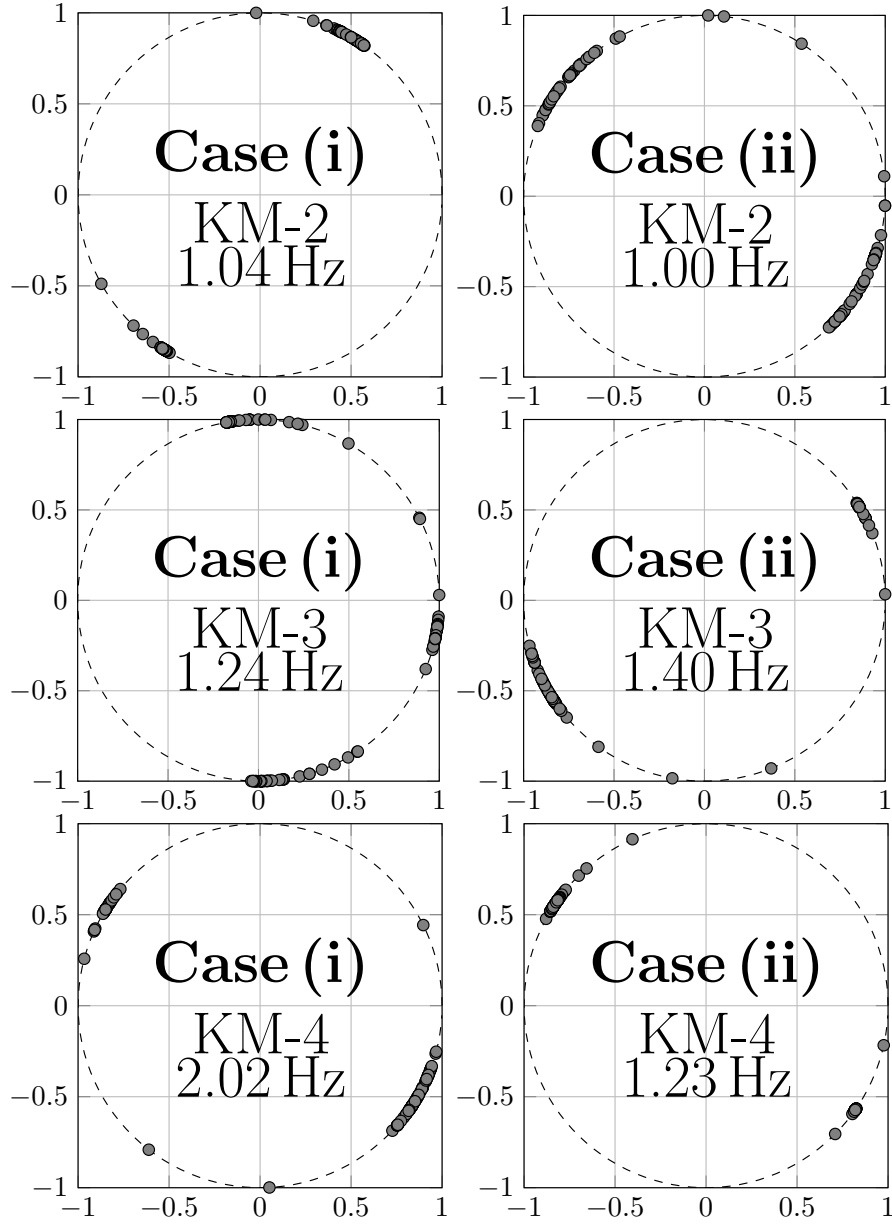


Figure 9.8: Initial phase vectors α_j of selected oscillatory KMs listed in Table 9.1.

same mode is captured for both the disturbance cases and that its shape is similar to $\mathbf{V}_2$. KM-4 for Case (i) provides a cutset similar to cutsets for KM-3 and 4 for Case (ii).

Here, it should be noted that the partition depends on if buses in a coherent group are geographically adjacent. For instance, it is possible that buses in the north and south of the system are swinging coherently against the central buses for a certain mode. Thus, even though they are detected as coherent in terms of KM, applying 2-means clustering and identifying cutsets will yield a partitioning into 3 groups: see e.g. the *broken green* circle in Fig. 9.9 (a), that encircles generator bus 31 which is incoherent with surrounding

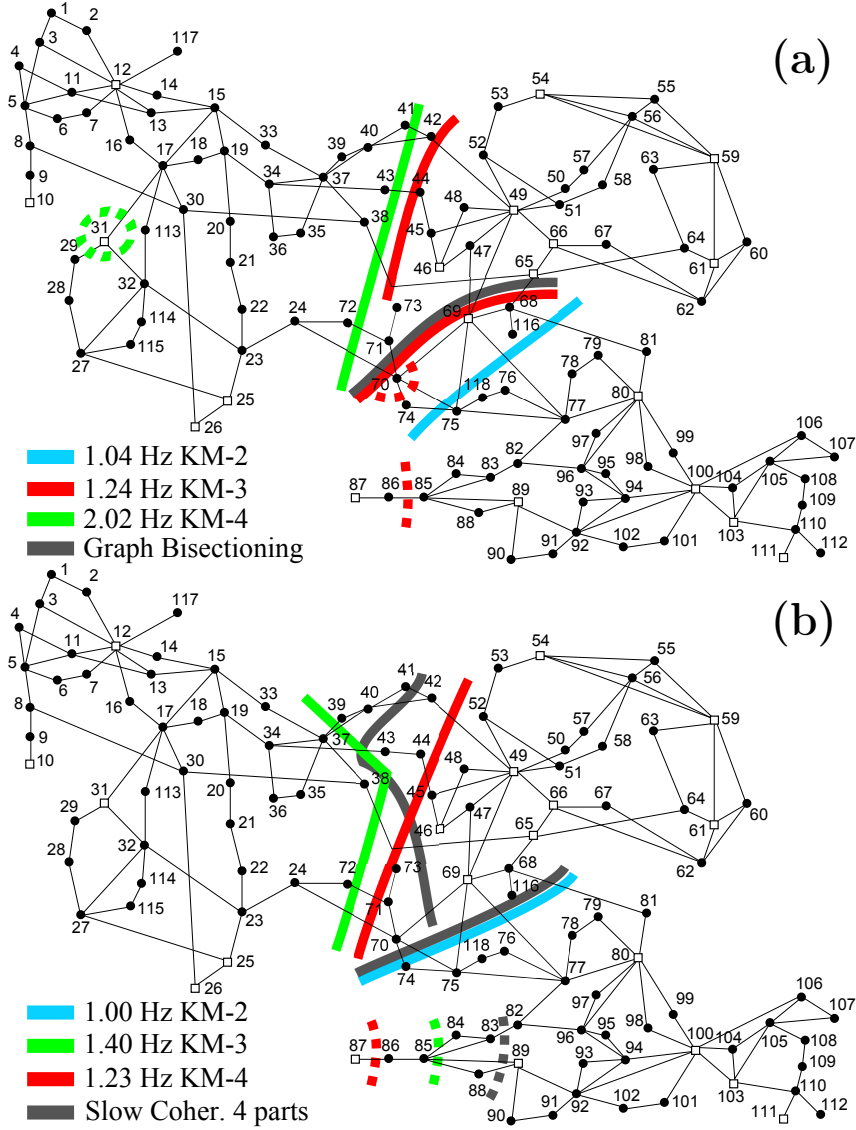


Figure 9.9: Partitioning of the IEEE 118-bus test system based on phase-coherency in KMs for (a) fault Case (i) (bus 17) and spectral graph bisectioning using V_2 ; and (b) fault Case (ii) (bus 100) and a slow-coherency technique including load buses. The cutsets are indicated by colored lines.

buses. A special attention should be given to these occurrences.

9.4.5 Comparison with slow-coherency-based islanding

Slow-coherency-based partitioning has widely been used, as reviewed in Sections 9.1 and 9.2. Here, as a comparison, we partition a power network based on the method proposed in [191] described in Section 9.2, and which extends the method in [198] to include load

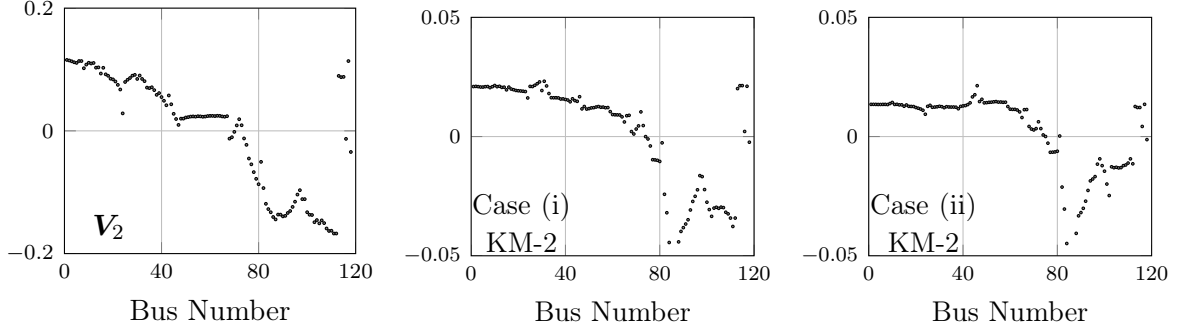


Figure 9.10: Comparison between the second eigenvector $\mathbf{V}_2$ of the graph Laplacian and spatial shapes $\mathbf{S}_{CC}$ of KM-2 for both disturbance cases.

buses and to not only group generators behind transient reactances. The resulting 4-way partitioning for the slow-coherency based method is depicted in Fig. 9.9 (b). As was pointed out earlier for Figs. 9.9 (a)–(b), the KMD-based partitioning identifies generator bus 87 as incoherent with the adjacent buses. In agreement with this, the slow-coherency-based method identifies bus 87 as an incoherent bus with respect to the adjacent buses. A more refined slow-coherency-based partitioning (not shown) reveals a new partition around bus 31, as was identified by KM-4 for Case (i). On the other hand, graph theoretical methods do not take into account the peculiar behavior and interaction of dynamical components in the large test system. Additionally, KM-2’s partition for Case (ii) is identical to one of the cutsets obtained from the slow-coherency technique. Also, KM-2 and 3 from Case (i) as well as spectral graph bisectioning provides a similar cutset. In fact, the partitioning obtained from KM-3 (1.24 Hz) for Case (i) corresponds to a large extent to the 4-way partitioning using the slow-coherency technique. The same result is also achievable by combining the cutsets of two or more KMs.

9.4.6 Computational considerations

An important practical aspect of using the proposed KMD-based method is the window length of the acquired data. Here, we briefly demonstrate its influence on the results by a numerical example using the data on bus-angle swings shown in Fig. 9.7 (b). For the original window length in terms of samples; $l_w = N + 1 = 421$, we examine the results of KMD with the slightly perturbed window length $l'_w = l_w + \Delta l_w$ with Δl_w set to ± 6 and ± 30 (here, with our used f_s , $\Delta l_w = 30$ corresponds to 0.5 s). Note that the “starting point” of the window is the same for all cases. In Table 9.2, KMs are listed which closely correspond to the KMs previously used for the partitioning. In one case (for $\Delta l_w = -6$), a KM close to ‘KM-4’ could not be picked up which is marked by ‘-’ in the table. To demonstrate how the partitioning result can change with varied window length, partitioning for ‘KM-3’ is given in Fig. 9.11 for all window lengths. It can be concluded that slightly different results can be obtained with different window lengths, but the results remain intact to a large extent despite changes in the window length. For

Table 9.2: Koopman modes obtained for the Case (ii) fault for different window lengths where Δl_w is the difference in number of samples from the original window length 421. The modes that closely correspond to the previously identified KMs 2-4 have been picked up among the dominant modes.

Case	‘KM-2’		‘KM-3’		‘KM-4’	
Δl_w	Freq. [Hz]	Norm	Freq. [Hz]	Norm	Freq. [Hz]	Norm
-30	1.07	0.30	1.37	0.29	1.25	0.71
-6	1.10	0.36	1.32	0.38	-	-
0	1.00	0.17	1.40	0.33	1.23	1.73
6	0.96	0.14	1.38	0.73	1.18	0.82
30	0.92	0.25	1.36	0.91	1.19	0.50

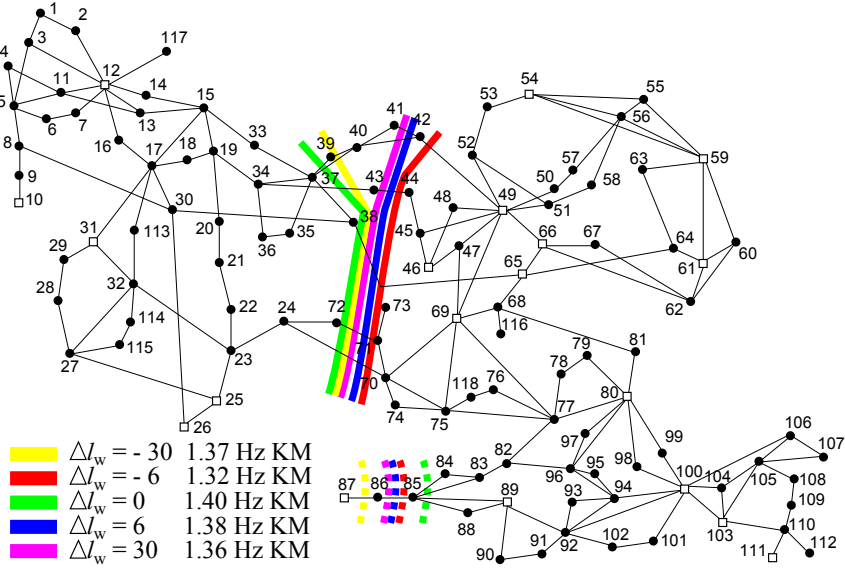


Figure 9.11: Partitioning of the IEEE 118-bus test system for all cases in Table 9.2 for ‘KM-3’.

more reliable result it could be useful to use a moving window and combine the results from a collection of runs.

9.5 Number of partitions and multiple coherency

So far, the power network was partitioned according to the method outlined in Algorithm 1. An appropriate number of groups of buses, called Number of Coherent Groups (NCG) from now on, was determined by looking at how closely groups of buses were clustered along the unit circle, and the distance between the centers of those clusters (provided by k -means). Another way to determine an appropriate number of groups is to utilize the

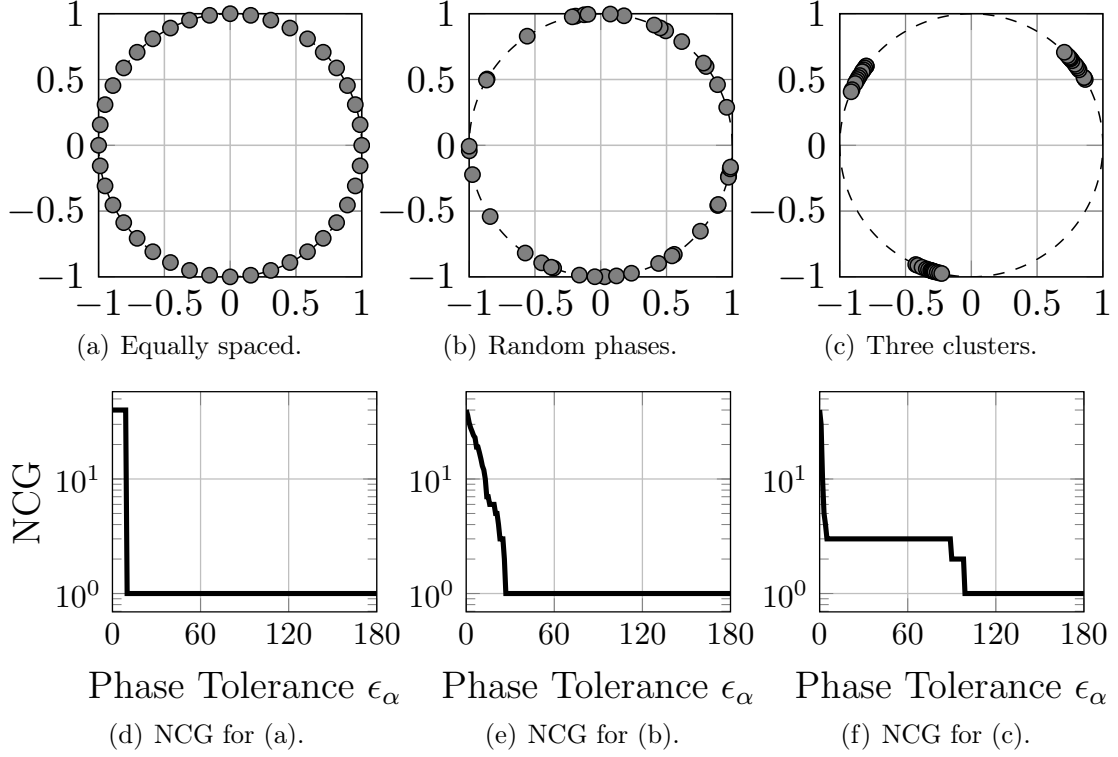


Figure 9.12: Illustrative examples of how the Number of Coherent Groups (NCG) varies as a function of phase tolerance ϵ_α . The number of phase points (and maximal NCG) is 40, (a) shows uniformly distributed angles with 9° phase difference between two consecutive points, (b) uniformly distributed random angles $[0^\circ, 360^\circ]$, and (c) three clear clusters of phase points. Plots (d)–(f) show NCG when ϵ_α is varied from 0° to 120° for the phase distribution given directly above.

coherency criterion (9.13), which provides information on how the NCG varies with the chosen tolerance ϵ_α . Figure 9.12 provides examples of identification of coherent groups based on the condition (9.13) as a function of the phase tolerance ϵ_α . If phase angles are distributed uniformly on the unit circle with a fixed spacing (9° is used in Fig. 9.12(a)), then obviously NCG decreases from the number of phase angles to 1 when $\epsilon_\alpha = 10^\circ$. Let $\epsilon_{m\alpha}$ denote the minimum angle difference between two points such that $\text{NCG}(\epsilon_{m\alpha} + \epsilon) = 1$, where ϵ is a small, positive constant ($\epsilon_{m\alpha} = 9^\circ$ for Fig. 9.12(a)). For the randomly distributed set of phase angles in Fig. 9.12(b), NCG decreases from its maximal value until $\text{NCG}(\epsilon_{m\alpha})$ in a linear manner on the logarithmic plot, which corresponds to an exponential decrease. On the other hand, for a clustered phase angle distribution (Figs. 9.12(c) and (f)), a large permissiveness to changes in ϵ_α apparent as constant plateaus is evident. Thus, the coherent groups for KMs are identified by locating the plateaus.

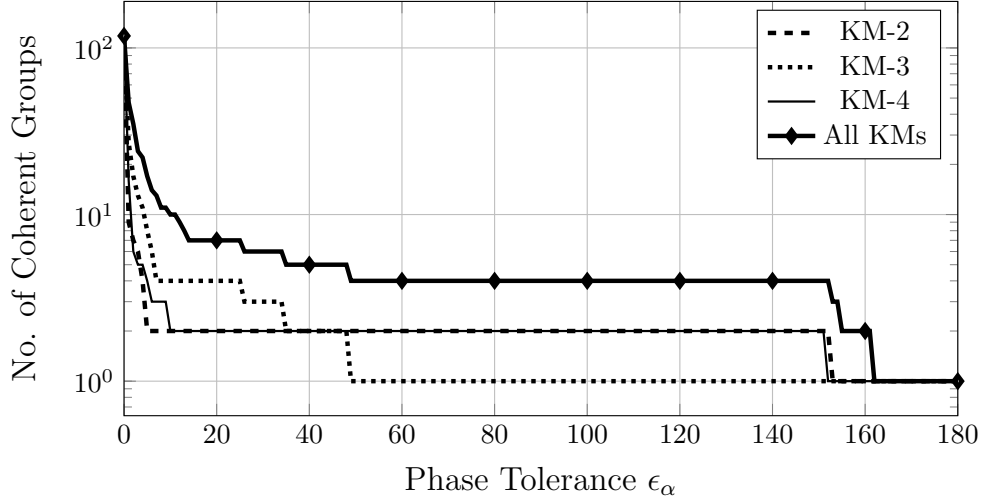


Figure 9.13: Number of Coherent Groups (NCG) vs. phase tolerance ϵ_α according to (9.13) for KMs 2-4 and for the combination of them (‘All KMs’).

9.5.1 Demonstration of multiple coherency in KMs

Now, the results of multiple coherence in KMs are demonstrated by considering the dominant KMs for Case (i) in Table 9.1 as an example. Figure 9.13 presents NCG vs. ϵ_α plots for the three KMs individually and for their combination based on the multiple coherency criterion (9.14). KM-2 displays a clear partitioning for two groups which is maintained for $\epsilon_\alpha \in [5^\circ, 152^\circ]$, and KM-4 gives a similar result. KM-3 displays clear plateaus for a NCG of 4, 3 and 2. For the plot corresponding to NCG for multiple KMs denoted ‘All KMs’, a wide plateau is visible for NCG = 4. However, for this plateau NCG decreases to 1 for KM-4 and does not provide any information. Thus, the second largest plateau occurring for NCG = 5 is used here as a multi-way partitioning for KMs. This actually corresponds to the union of cutsets for NCG = 2 for all KMs, shown in Fig. 9.9 (a). The almost identical cutsets (except for the two load-buses 71 and 73) in the left central part provided by KM-3 and 4 are treated as one which ultimately provides a 4-way partitioning of the network.

The KMD-based partitioning is shown in Fig. 9.14. For the sake of comparison, the same 4-way partitioning via slow-coherency as shown earlier is also depicted. 3 and 4-way partitioning using the Laplacian eigenvectors $\{\mathbf{V}_2, \mathbf{V}_3\}$ and $\{\mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4\}$, together with k -means are also illustrated in Fig. 9.14. Again, note that the Laplacian is constructed with all edge weights $w(i, j)$ set to 1.

9.5.2 Time-domain simulations

Now, time-domain simulations are run for the four different partitioning results in Fig. 9.14. The system remains in a steady state until $t = 1$ s when the network separation is initiated by simultaneously disconnecting all tie-lines associated with the partitioning. In Fig. 9.15,

time responses for angular frequencies ω_i of generators are given for the four cases. It is clear that the generators converge to new steady states in all cases. The magnitude of the frequency deviation depends on the balance between load and generation, and the smallest deviations are achieved for the partitioning according to spectral graph theory into three parts.

For 4-way partitioning, the proposed KMD-based method yields the overall smallest frequency deviations, implying that the mismatch between power generation and consumption within the areas is the smallest. Spectral graph partitioning demonstrated that cut-sets based on identification of weak links between tightly connected clusters of buses can provide quite good results in terms of load/generation balance (see Fig. 9.15 (c)), since the power transferred over the tie-lines between the areas is often small compared to the generated and consumed power within the large areas. The results of applying KMD demonstrates that the interface between incoherent buses in KMs falls on these tie-lines too, yielding viable cut-sets. To improve the results, it would make sense to utilize an algorithm identifying the most optimal cut (e.g. such that the total power over the lines is minimized) around the identified interfaces.

A close-up showing the angular responses of generators during the first second after separation are given in Fig. 9.16. The KMD-based partitioning displays possibly the smallest oscillation amplitudes of generators for 4-way partitioning: compare e.g. Figs. 9.16 (a) and (b). This is caused by a smaller mismatch in power compared to the other methods.

Keep in mind that these simulations do not constitute a comprehensive evaluation of the performance of the proposed method. In a controlled islanding strategy, operational constraints such as stability limits and transmission capabilities should be evaluated in addition to load-generation balance and coherency of generators (see e.g. [207]).

9.6 Conclusions

In this chapter, a data-driven method for power network partitioning based on KMD was demonstrated. It was numerically and analytically shown that the data-driven method can cover network partitions derived from both spectral graph and a slow-coherency-based method. Apart from identifying oscillatory modal structures used for partitioning, the KMD-based partitioning identifies the oscillatory frequencies, as well as information on damping and participation, thus it can be used for monitoring and control purposes, e.g. in a controlled islanding technique [187, 188].

By exploiting a new coherency condition for KMs based on a maximum allowable phase difference tolerance between two buses, an appropriate amount of partitions can be identified as a permissiveness to changes in phase difference tolerance for a certain Number of Coherent Groups (NCG), which is apparent as plateaus in a NCG vs. phase tolerance plot. This technique is valid for multiple KMs and can thus derive multiple frequency coherent partitions.

Integrating the data-driven partitioning with a controlled islanding strategy is one of the future works. Such a realization could possibly incorporate PMU-based out-of-step

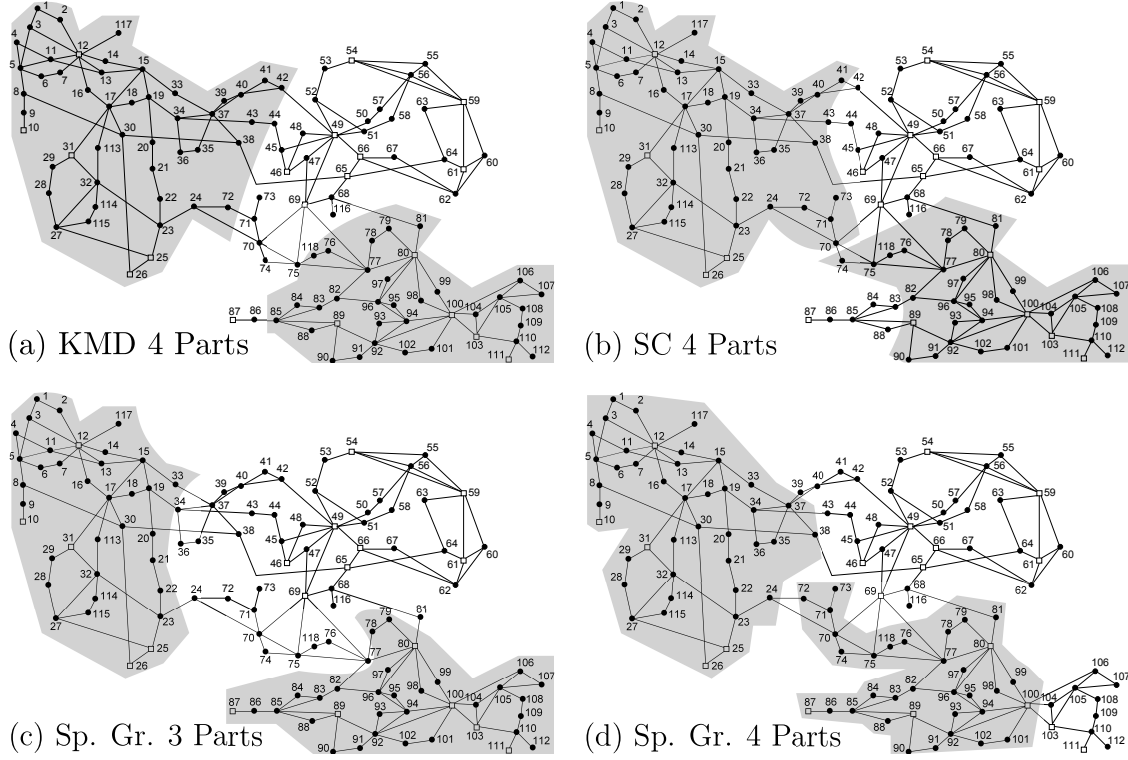


Figure 9.14: Multi-way partitioning of the IEEE 118-bus test system based on (a) phase-coherency in KMs 2–4 identified from dynamics of the three phase fault, (b) a Slow-Coherency (SC) technique including load buses, and (c),(d) for spectral graph theory for 3 and 4-way partitioning. The partitions can be distinguished based on the coloring.

protection as proposed in [208]. Further investigations for practical implementation of the proposed algorithm are of great importance: influence of network delay and measurement noise on the partitioning result, and data-acquisition methods for successful partitioning.

The advantage of the proposed method lies in that it only requires a window of sampled dynamics and implicitly takes into account versatile features in the power system dynamics that otherwise requires a substantial amount of model parameters. With the expected increase in availability of synchronized measurements in the future, it is important to investigate new tools for monitoring and control based on sampled data. Since the proposed method is intended only for application to sampled data following large faults, it is limited to the extraction of the dynamic behavior excited by disturbances, and should act as a complimentary tool to the standard power system analysis and control. For real-world applications of data-based methods, one has to be careful about noise in the data acquisition (this was discussed in Chapter 3). It has been shown that the type of algorithm applied here is sensitive to noise [83], and thus care should be taken by filtering out the noise in the measured data.

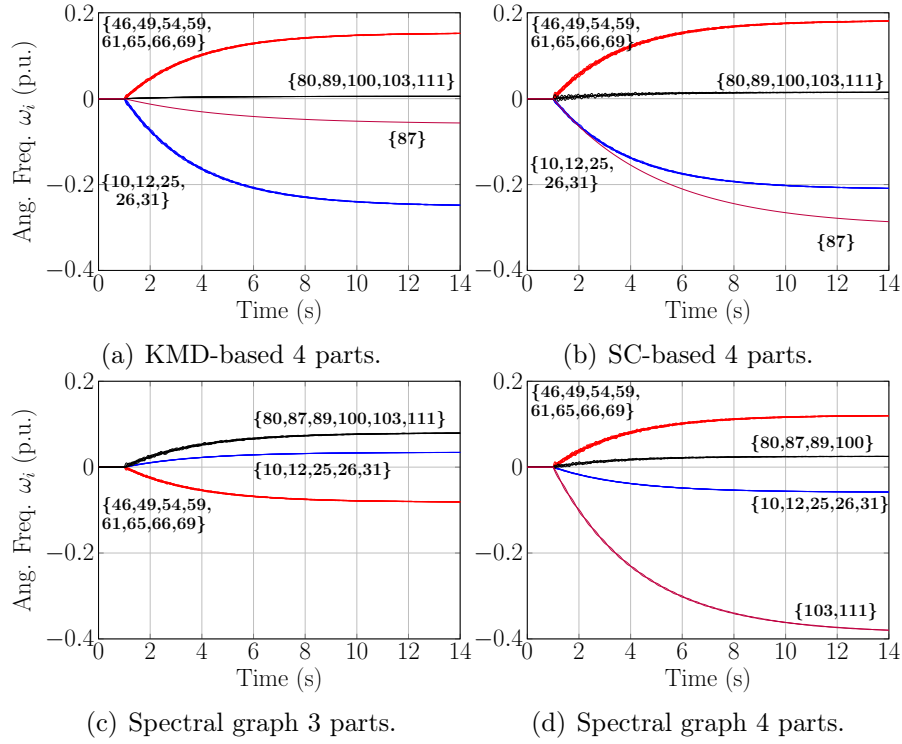


Figure 9.15: Angular frequency deviations ω_i for simulations of system islanding initiated at $t = 1$ s for the partitionings ((a)–(d)) in Fig. 9.14. ω_i for generators belonging to the same group are shown in the same color and their bus numbers are given in the brackets.

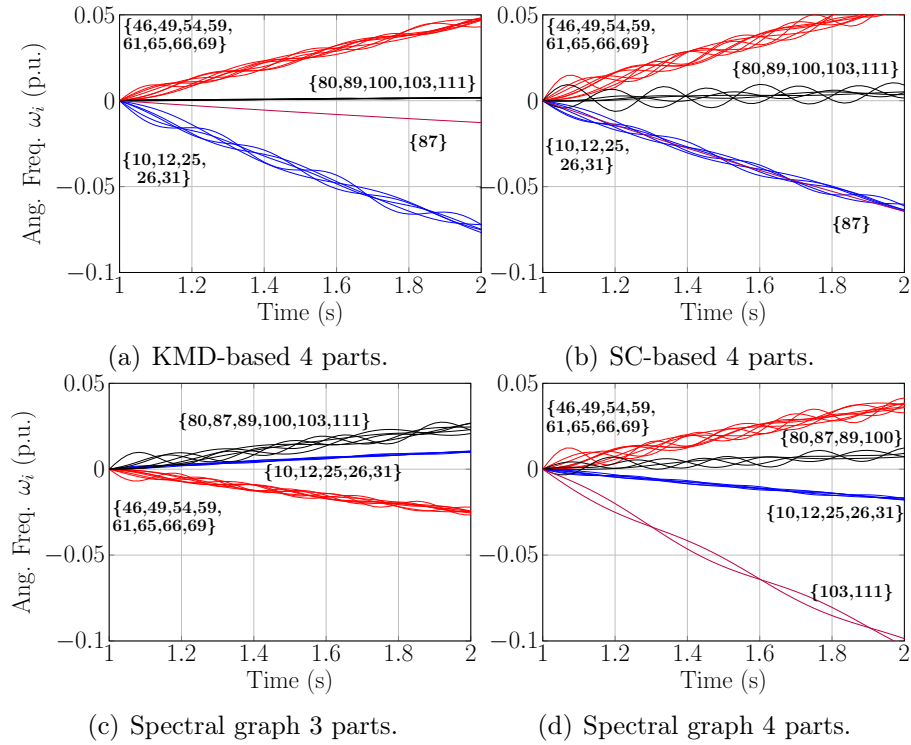


Figure 9.16: Same as Fig. 9.15 shown for the first second following the islanding operation.

Chapter 10

Conclusions and future outlook

The main focus of this dissertation was on numerical analysis of wind power fluctuations and power system oscillations by utilizing numerical techniques related to the recently developed Koopman analysis framework. In this dissertation, we called these types of techniques Koopman Mode Decompositions (KMD), in accordance with [23]. New KMD-applications were demonstrated such as partitioning of power networks, smoothing effects of wind power in Japan, and wind speed prediction. Wind power fluctuations were analyzed on spatial scales ranging from hundreds of kilometers to hundreds of meters, and on time-scales ranging from seconds to weeks. A weather simulator, Cloud Resolving Storm Simulator (CReSS) [63], not previously used in wind power analysis was utilized and was evaluated against real measurements from a Wind Farm (WF) in Japan. The results suggested CReSS's ability to capture wind power variability on an hourly scale. A statistical evaluation of short-term wind power fluctuations was conducted which showed that significant Wind Turbine (WT) output fluctuations were to a large extent not suppressed in the WF output, causing large fluctuations. This implies that care should be taken about short-term correlations when installing large-scale wind power in a region. The results from an analysis of large-scale wind power provided clues to the smoothing effects of aggregated offshore wind power in northern Japan, which were quantified by a proposed smoothing index via KMD. Furthermore, through analysis of power system oscillations, we demonstrated connections between Koopman modes and spectral graph theory. Specifically, an indication that graph properties can be inferred from dynamics on networks via the spectrum of the Koopman operator. Also, a practical data-driven method utilizing KMD to partition a power system into disjoint parts was outlined. Such a method is a vital component of a so-called controlled islanding strategy aimed to save power systems from large-scale blackouts. Through all these applications, a comprehensive theoretical background, and an evaluation of KMD-techniques, we clarified applicability of KMD to various problems and data. We believe that this will stimulate further research efforts of using KMD to improve the monitoring and security of power systems, and increase our understanding of characteristics of aggregated of wind power and renewables.

10.1 Conclusions

Chapter 2 reviewed Koopman operator theory and related numerical methods to perform KMD. KMD is closely linked to the numerical method called Dynamic Mode Decomposition (DMD) [16], an acronym sometimes used interchangeably. KMD is based on spectral analysis of the linear but infinite dimensional Koopman operator which governs the evolution of observables of nonlinear dynamical systems. The numerical approximations of KMD such as DMD all perform spectral decompositions from data by a finite number of eigenvalues and modes. The chapter reviewed the state-of-the-art of KMD by describing and comparing different algorithms, and clarified their relation by new theorems.

Chapter 3 evaluated numerical performance of common KMD/DMD algorithms by three types of experimental data: linear dynamics, power system oscillations, and wind speed predictions. For example, it is shown how the sampling frequency and temporal length of the data acquisition window affect the performances of the algorithms. Their performance is evaluated by comparing against dominant frequencies in the dynamics, or by how well they predict future values. In particular, the so-called spatial dimension, which is defined as the number of measurement points, is shown to be critical to the performance. It was shown that the so-called vector Prony analysis [62] works well for a small spatial dimension, where standard DMD is not suitable.

Chapter 4 explained the wind power fundamentals and modeling of wind power systems. Models of the most common WT configurations including standard controls and mechanical systems were described. Furthermore, an example of a dynamic simulation of a WF incorporated with wind predictions was presented.

Chapter 5 evaluated short-term wind speed and wind power fluctuations with measurements from an offshore WF located in Japan. A dynamic model incorporating data-derived characteristics of the real WTs was shown to reproduce realistic output fluctuations to a large extent. This result implies that a large part of the output fluctuations are due to the deterministic electromechanical conversion or control system and not random. This suggests the importance of a representative short-term model of the WT for understanding and controlling fluctuations in the short-term. Furthermore, we showed that aggregated WT fluctuations are not mitigated on a second scale for variations less than about 5–6 standard deviations, which indicates strong short-term correlations on the scale of hundreds of meters to 1.5 km. This implies that large WF output variations are fed into the power system which could possibly cause voltage and power fluctuations, and that care should be taken about short-term correlations and their potential impact on the power system operation.

Chapter 6 looked at larger scale wind power fluctuations, which were analyzed with data from a weather prediction model called CReSS. Both coarser 2-km simulations encompassing the whole of Japan and extremely detailed 200-m resolution simulations of a smaller domain were evaluated. Predictions were compared against real measurements. It was shown that wind power increments (e.g. [27, 64]) can be assessed on the time-scale of several hours. Improvements in the prediction of wind power increments were demon-

strated for hourly averaged power, which suggests CReSS's ability to predict wind power ramps on an hourly scale. The measured and predicted wind spectra were found similar in both the long- and short-term. The results for the detailed simulations showed that CReSS captures the so-called spectral coherence of wind on the scale of one to a few kilometers.

Chapter 7 investigated smoothing effects of large-scale wind power in Japan by again incorporating wind speed predictions from CReSS. With increasing penetration of new renewable energy such as solar and wind power, it becomes important for power system and WF operators to grasp the amount of smoothing achieved in the aggregated power, and at what temporal and spatial scales smoothing is achieved. We proposed a new smoothing index for wind power based on KMD, which can be regarded as a generalization of a previously proposed index based on power spectral densities. WFs around six regions in northern Honshu, Japan, were considered as a test case for the evaluation. Through the proposed index, it was shown how the smoothing improves by distributing WFs over different regions. The results indicate that by distributing WFs over only one to three regions, smoothing results vary considerably depending on the choice of regions. However, as the number of considered regions increases the smoothing improves, and the particular choice of regions matters less for smoothing effects at the investigated time-scales. These results highlight the importance of deliberately selecting sites for large-scale wind power production to more effectively smooth the aggregated power.

Chapter 8 discussed KMD with applications to short-term predictions of wind speeds and control of optimal power extraction of WTs. 1-Hz measurements of wind speeds and output powers from WTs in Japan were used for validation and benchmarking. The most fundamental and important task of a WT is to extract as much power as possible and feed the power into the power grid. To do so, the turbine's speed needs to be controlled as close to its optimal value as possible. The optimal rotor speed depends on the ratio between the speed at the tip of the turbine's blades and the incoming wind, called the tip speed ratio. However, the wind speed changes quickly, and a turbine cannot change its speed as fast as the wind due to its inertia. This will render the turbine unable to follow the fast changes, and consequently unable to fully utilize the wind energy. We showed here that investigated wind speed time-series do not exhibit low-dimensional, nonlinear deterministic features on the time-scale of seconds. However, they exhibit short-term linear correlations which are well described by a stochastic differential equation. We also showed that DMD/KMD provide improvements compared to a naive predictor for prediction horizons of up to 5 seconds.

Chapter 9 discussed KMD related to power system dynamics and security. KMD was used to determine the points of separation in a so-called controlled islanding strategy in power system emergency control. A practical data-driven algorithm incorporating KMD was proposed for network partitioning. Comparisons were made with two techniques previously applied for network partitioning: spectral graph theory [69], which is based on the eigenstructure of the graph Laplacian, and slow-coherency, which identifies coherent groups of generators for a specified number of low-frequency modes. The partitioning results share common features with results obtained with graph theory and

slow-coherency-based techniques. The suggested partitioning method was evaluated with two test systems, and similarities between Koopman modes and Laplacian eigenvectors were showed numerically and elaborated theoretically.

10.2 Future work and outlook

This dissertation touched on a rather large number of topics related to wind power, Koopman operator theory, time-series analysis, complex networks, and power systems. In the following, several future directions are outlined related to wind power analysis and prediction; and power systems and network analysis.

Connections between KMD and spectral graph theory were discussed in this dissertation related to the application of KMD to power system data. Recently, in [209, 210], a clear relationship between graph spectral graph theory and the spectrum of the Koopman operator was demonstrated, and methods were proposed to infer network information from only sparse measurements of dynamics on the network. Deriving a network from wind speed fluctuations via KMD-like methods for predicting and understanding large-scale wind power characteristics is an interesting new research direction. This is inspired by the recently promising approach of applying complex network theory to climate dynamics [211–214], which was demonstrated to work well for future predictions. Specifically, closely related to the proposed idea, [214] looked at prediction of large-scale wind fields. Furthermore, this dissertation only looked at wind power generation, while a comprehensive analysis of smoothing of aggregated solar and wind power would be highly interesting to investigate, particularly in Japan where solar power covers a substantial amount of the electricity demand. Such study can clarify possibilities, limitations, and requirements related to a future power system incorporating a large amount of renewables. Recently, a similar study was conducted for the Nordic countries in [61], indicating the possibility of a completely renewable power system. As a direction extension to the quantification of smoothing demonstrated in this dissertation, one should apply some of the latest DMD/KMD-methods which are directly compatible with the proposed smoothing index, to evaluate smoothing. In the prediction of wind speed, it is also worthwhile to evaluate potentially better observable functions, however, care should be taken about the computational time, which becomes critical for short-term predictions.

There are many unexplored applications of KMD/DMD in monitoring, control, and analysis of power system networks. Power systems are becoming more uncertain and stochastic in their nature as an effect of distributed generation and high penetration of renewables [215]. This poses a challenge to system operators in terms of properly assessing the stability of the system [216]. In [216], stability assessment is conducted with a power system modeled by stochastic differential equations, and it is argued that conventional stability assessment is too conservative. Thus, instead of conventional deterministic differential equations, it may be more correct to model the power systems by stochastic differential equations as in [215]. In this context, it can be speculated that one promising potential application of KMD is to assess the state of the system by extracting a reduced

model of the underlying dynamics—separating determinism from randomness.

Related to that, it is important to investigate the most suitable choice of observable functions for monitoring power system dynamics. In this dissertation, the Arnoldi-like KMD was applied directly to “raw” data. However, improvements might be achievable by using sophisticated observation functions and methods such as Extended DMD (EDMD). For example the algorithm proposed in [217] could possibly be applied to help identify suitable observables as in [76]. Moreover, through monitoring to keep track on poorly damped oscillations, one should utilize the data-derived information to automatically tune controllers to more effectively damp oscillations. Also, investigating applications of KMD to normal operating conditions should be considered—an example of a system identification method for such conditions is e.g. proposed in [59]. Lastly, data-driven network partitioning should be included in a complete controlled islanding strategy by utilizing PMUs and control of relays to disconnect lines.

Bibliography

- [1] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer-Verlag, New York, 2002.
- [2] K. Alligood, T. Sauer, and J. Yorke, *Chaos: An Introduction to Dynamical Systems*. Textbooks in Mathematical Sciences, Springer-Verlag, New York, 2000.
- [3] M. Hirsch, S. Smale, and R. Devaney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Academic Press, San Diego, CA, 2004.
- [4] M. Casdagli, “Chaos and deterministic versus stochastic non-linear modelling,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 54, no. 2, pp. 303–328, 1992.
- [5] J. P. Crutchfield, “Between order and chaos,” *Nature Physics*, vol. 8, no. 1, pp. 17–24, 2011.
- [6] H. Holttinen, P. Meibom, A. Orths, B. Lange, M. O’Malley, J. O. Tande, A. Estanqueiro, E. Gomez, L. Söder, G. Strbac, J. C. Smith, and F. van Hulle, “Impacts of large amounts of wind power on design and operation of power systems, results of IEA collaboration,” *Wind Energy*, vol. 14, no. 2, pp. 179–192, 2011.
- [7] P. Kundur, *Power System Stability and Control*. McGraw-Hill, New York, 1994.
- [8] G. Froyland, K. Padberg, M. H. England, and A. M. Treguier, “Detection of coherent oceanic structures via transfer operators,” *Physical Review Letters*, vol. 98, no. 22, 224503, 2007.
- [9] G. Froyland, G. A. Gottwald, and A. Hammerlindl, “A computational method to extract macroscopic variables and their dynamics in multiscale systems,” *SIAM Journal on Applied Dynamical Systems*, vol. 13, no. 4, pp. 1816–1846, 2014.
- [10] S. Klus, F. Nüske, P. Koltai, H. Wu, I. Kevrekidis, C. Schütte, and F. Noé, “Data-driven model reduction and transfer operator approximation,” *arXiv preprint arXiv:1703.10112*, 2017.
- [11] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*. Springer-Verlag, New York, 1994.

- [12] B. O. Koopman, "Hamiltonian systems and transformation in Hilbert space," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 17, no. 5, pp. 315–318, 1931.
- [13] M. Budišić, R. Mohr, and I. Mezić, "Applied Koopmanism," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 22, no. 4, 047510, 2012.
- [14] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," *Nonlinear Dynamics*, vol. 41, no. 1-3, pp. 309–325, 2005.
- [15] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, "Spectral analysis of nonlinear flows," *Journal of Fluid Mechanics*, vol. 641, pp. 115–127, 2009.
- [16] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of Fluid Mechanics*, vol. 656, pp. 5–28, 2010.
- [17] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, "On dynamic mode decomposition: Theory and applications," *Journal of Computational Dynamics*, vol. 1, no. 2, pp. 391–421, 2014.
- [18] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition," *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, 2015.
- [19] Y. Susuki and I. Mezić, "Nonlinear Koopman modes and coherency identification of coupled swing dynamics," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 1894–1904, 2011.
- [20] Y. Susuki and I. Mezić, "Nonlinear Koopman modes and a precursor to power system swing instabilities," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1182–1191, 2012.
- [21] Y. Susuki and I. Mezić, "Nonlinear Koopman modes and power system stability assessment without models," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 899–907, 2014.
- [22] Y. Susuki, I. Mezić, F. Raak, and T. Hikihara, "Applied Koopman operator theory for power systems technology," *Nonlinear Theory and Its Applications, IEICE*, vol. 7, no. 4, pp. 430–459, 2016.
- [23] I. Mezić, "Analysis of fluid flows via spectral properties of the Koopman operator," *Annual Review of Fluid Mechanics*, vol. 45, pp. 357–378, 2013.
- [24] S. Bagheri, "Koopman-mode decomposition of the cylinder wake," *Journal of Fluid Mechanics*, vol. 726, pp. 596–623, 2013.

- [25] X. Lu, M. B. McElroy, and J. Kiviluoma, "Global potential for wind-generated electricity," *Proceedings of the National Academy of Sciences*, vol. 106, no. 27, pp. 10933–10938, 2009.
- [26] R. Baïle and J.-F. Muzy, "Spatial intermittency of surface layer wind fluctuations at mesoscale range," *Physical Review Letters*, vol. 105, no. 25, 254501, 2010.
- [27] H. Holttinen, "Hourly wind power variations in the Nordic countries," *Wind Energy*, vol. 8, no. 2, pp. 173–195, 2005.
- [28] A. Malvaldi, S. Weiss, D. Infield, J. Browell, P. Leahy, and A. M. Foley, "A spatial and temporal correlation analysis of aggregate wind power in an ideally interconnected Europe," *Wind Energy*, vol. 20, no. 8, pp. 1315–1329, 2017.
- [29] M. Ilić and J. Zaborszky, *Dynamics and Control of Large Electric Power Systems*. John Wiley & Sons, New York, 2000.
- [30] Y. V. Makarov, V. I. Reshetov, A. Stroeve, and I. Voropai, "Blackout prevention in the United States, Europe, and Russia," *Proceedings of the IEEE*, vol. 93, no. 11, pp. 1942–1955, 2005.
- [31] G. Andersson, P. Donalek, R. Farmer, N. Hatziaargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, J. Sanchez-Gasca, R. Schulz, A. Stankovic, C. Taylor, and V. Vittal, "Causes of the 2003 major grid blackouts in North America Europe, and recommended means to improve system dynamic performance," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1922–1928, 2005.
- [32] N. W. Miller, "Keeping it together: Transient stability in a world of wind and solar generation," *IEEE Power and Energy Magazine*, vol. 13, no. 6, pp. 31–39, 2015.
- [33] M. Milligan, B. Frew, B. Kirby, M. Schuerger, K. Clark, D. Lew, P. Denholm, B. Zavadil, M. O'Malley, and B. Tsuchida, "Alternatives no more: Wind and solar power are mainstays of a clean, reliable, affordable grid," *IEEE Power and Energy Magazine*, vol. 13, no. 6, pp. 78–87, 2015.
- [34] J. W. Simpson-Porco, F. Dörfler, and F. Bullo, "Voltage collapse in complex power grids," *Nature Communications*, vol. 7, 10790, 2016.
- [35] Global Wind Energy Council (GWEC), *Global Wind Report Annual Market Update*, 2016. [Online]. Available at <http://www.gwec.net/global-figures/graphs/> (Last accessed on 24 July 2017).
- [36] X. Sun, D. Huang, and G. Wu, "The current state of offshore wind energy technology development," *Energy*, vol. 41, no. 1, pp. 298–312, 2012.
- [37] T. Ackermann, "Historical development and current status of wind power," in *Wind Power in Power Systems* (T. Ackermann, ed.), pp. 21–24, John Wiley & Sons, Ltd, Chichester, U.K, 2012.

- [38] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation," *Renewable Energy*, vol. 37, no. 1, pp. 1–8, 2012.
- [39] Renewable Energy Policy Network for the 21st Century (REN21), *Renewables 2016 Global Status Report (Paris: REN21 Secretariat)*, 2016. [Online]. Available at http://www.ren21.net/wp-content/uploads/2016/06/GSR_2016_Full_Report.pdf (Last accessed on 24 July 2017).
- [40] D. Flynn, Z. Rather, A. Ardal, S. D'Arco, A. Hansen, N. Cutululis, P. Sorensen, A. Estanquero, E. Gómez, N. Menemenlis, C. Smith, and Y. Wang, "Technical impacts of high penetration levels of wind power on power system stability," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 6, no. 2, e216, 2017.
- [41] M. Bilgili, A. Yasar, and E. Simsek, "Offshore wind power development in Europe and its comparison with onshore counterpart," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 2, pp. 905–915, 2011.
- [42] B. Snyder and M. Kaiser, "Ecological and economic cost-benefit analysis of offshore wind energy," *Renewable Energy*, vol. 34, no. 6, pp. 1567–1578, 2009.
- [43] R. Komiyama and Y. Fujii, "Analysis of Japan's long-term energy outlook considering massive deployment of variable renewable energy under nuclear energy scenario," *Electrical Engineering in Japan (English translation of Denki Gakkai Ronbunshi)*, vol. 190, no. 2, pp. 24–40, 2015.
- [44] A. Yamaguchi, T. Ishihara, K. Sakai, T. Ogawa, and Y. Fujino, "A physical-statistical approach for the regional wind power forecasting," in *Proceedings of the European Wind Energy Conference EWEK*, (Milan, Italy), 7–10 May 2007.
- [45] R. Doherty, E. Denny, and M. O'Malley, "System operation with a significant wind power penetration," in *2004 IEEE Power Engineering Society General Meeting*, (Denver, Colorado, USA), June 2004.
- [46] D. Gautam, V. Vittal, and T. Harbour, "Impact of increased penetration of DFIG-based wind turbine generators on transient and small signal stability of power systems," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1426–1434, 2009.
- [47] G. Shafiullah, A. M. Oo, A. S. Ali, and P. Wolfs, "Potential challenges of integrating large-scale wind energy into the power grid—a review," *Renewable and Sustainable Energy Reviews*, vol. 20, pp. 306–321, 2013.
- [48] G. Lalor, A. Mullane, and M. O'Malley, "Frequency control and wind turbine technologies," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1905–1913, 2005.

- [49] Z. Chen, “Issues of connecting wind farms into power systems,” in *2005 IEEE/PES Transmission Distribution Conference Exposition: Asia and Pacific*, (Dalian, China), pp. 1–6, 2005.
- [50] J. MacDowell, S. Dutta, M. Richwine, S. Achilles, and N. Miller, “Serving the future: Advanced wind generation technology supports ancillary services,” *IEEE Power and Energy Magazine*, vol. 13, no. 6, pp. 22–30, 2015.
- [51] F. Santos-Alamillos, D. Pozo-Vázquez, J. Ruiz-Arias, V. Lara-Fanego, and J. Tovar-Pescador, “Analysis of spatiotemporal balancing between wind and solar energy resources in the southern Iberian Peninsula,” *Journal of Applied Meteorology and Climatology*, vol. 51, no. 11, pp. 2005–2024, 2012.
- [52] K. Klima and J. Apt, “Geographic smoothing of solar PV: Results from Gujarat,” *Environmental Research Letters*, vol. 10, no. 10, 104001, 2015.
- [53] A. D. Mills and R. H. Wiser, “Implications of geographic diversity for short-term variability and predictability of solar power,” in *2011 IEEE Power and Energy Society General Meeting*, (Detroit, Michigan, USA), pp. 1–9, July 2011.
- [54] A. Tuohy and M. O’Malley, “Wind power and storage,” in *Wind Power in Power Systems* (T. Ackermann, ed.), pp. 465–487, John Wiley & Sons Ltd, Chichester, U.K, 2012.
- [55] D. Lindley, “The energy storage problem,” *Nature*, vol. 463, no. 7277, pp. 18–20, 2010.
- [56] J. Hauer, D. Trudnowski, G. Rogers, B. Mittelstadt, W. Litzenberger, and J. Johnson, “Keeping an eye on power system dynamics,” *IEEE Computer Applications in Power*, vol. 10, no. 4, pp. 50–54, 1997.
- [57] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. Van Cutsem, and V. Vittal, “Definition and classification of power system stability,” *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387–1401, 2004.
- [58] J. De La Ree, V. Centeno, J. S. Thorp, and A. G. Phadke, “Synchronized phasor measurement applications in power systems,” *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 20–27, 2010.
- [59] H. Ghasemi, C. Canizares, and A. Moshref, “Oscillatory stability limit prediction using stochastic subspace identification,” *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 736–745, 2006.
- [60] T. Browne, V. Vittal, G. Heydt, and A. Messina, “A comparative assessment of two techniques for modal identification from power system measurements,” *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1408–1415, 2008.

- [61] J. Olauson, M. N. Ayob, M. Bergkvist, N. Carpman, V. Castellucci, A. Goude, D. Lingfors, R. Waters, and J. Widén, “Net load variability in Nordic countries with a highly or fully renewable power system,” *Nature Energy*, vol. 1, 16175, 2016.
- [62] Y. Susuki and I. Mezić, “A prony approximation of Koopman mode decomposition,” in *2015 IEEE 54th Conference on Decision and Control (CDC)*, (Osaka, Japan), pp. 7022–7027, December 2015.
- [63] K. Tsuboki, “High-resolution simulations of high-impact weather systems using the cloud-resolving model on the earth simulator,” in *High Resolution Numerical Modelling of the Atmosphere and Ocean*, pp. 141–156, Springer, 2008.
- [64] P. Milan, M. Wächter, and J. Peinke, “Turbulent character of wind energy,” *Physical Review Letters*, vol. 110, no. 13, 138701, 2013.
- [65] E. Koutroulis and K. Kalaitzakis, “Design of a maximum power tracking system for wind-energy-conversion applications,” *IEEE Transactions on Industrial Electronics*, vol. 53, no. 2, pp. 486–494, 2006.
- [66] J. Manwell, J. McGowan, and A. Rogers, *Wind Energy Explained: Theory, Design and Application*. John Wiley & Sons, Chichester, U.K, 2009.
- [67] R. Sanchez-Garcia, M. Fennelly, S. Norris, N. Wright, G. Niblo, J. Brodzki, and J. Bialek, “Hierarchical spectral clustering of power grids,” *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2229–2237, 2014.
- [68] M. Bolla, *Spectral Clustering and Biclustering: Learning Large Graphs and Contingency Tables*. John Wiley & Sons, Chichester, U.K, 2013.
- [69] T. Bıyıkoglu, J. Leydold, and P. Stadler, *Laplacian Eigenvectors of Graphs: Perron-Frobenius and Faber-Krahn Type Theorems*. Lecture notes in mathematics, Springer Berlin Heidelberg New York, 2007.
- [70] J. H. Chow, R. Galarza, P. Accari, and W. W. Price, “Inertial and slow coherency aggregation algorithms for power system dynamic model reduction,” *IEEE Transactions on Power Systems*, vol. 10, no. 2, pp. 680–685, 1995.
- [71] M. Georgescu and I. Mezić, “Building energy modeling: A systematic approach to zoning and model reduction using Koopman mode analysis,” *Energy and Buildings*, vol. 86, pp. 794–802, 2015.
- [72] B. W. Brunton, L. A. Johnson, J. G. Ojemann, and J. N. Kutz, “Extracting spatial-temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition,” *Journal of Neuroscience Methods*, vol. 258, pp. 1–15, 2016.
- [73] H. Wu, F. Nüske, F. Paul, S. Klus, P. Koltai, and F. Noé, “Variational Koopman models: Slow collective variables and molecular kinetics from short off-equilibrium simulations,” *The Journal of Chemical Physics*, vol. 146, no. 15, 154104, 2017.

- [74] T.-Y. Li, “Finite approximation for the Frobenius-Perron operator. A solution to Ulam’s conjecture,” *Journal of Approximation Theory*, vol. 17, no. 2, pp. 177–186, 1976.
- [75] S. Klus, P. Koltai, and C. Schütte, “On the numerical approximation of the Perron-Frobenius and Koopman operator,” *Journal of Computational Dynamics*, vol. 3, no. 1, pp. 51–79, 2016.
- [76] S. L. Brunton, B. W. Brunton, J. L. Proctor, and J. N. Kutz, “Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control,” *PLOS ONE*, vol. 11, no. 2, e0150171, 2016.
- [77] J. L. Proctor, S. L. Brunton, and J. N. Kutz, “Dynamic mode decomposition with control,” *SIAM Journal on Applied Dynamical Systems*, vol. 15, no. 1, pp. 142–161, 2016.
- [78] K. K. Chen, J. H. Tu, and C. W. Rowley, “Variants of dynamic mode decomposition: Boundary condition, Koopman, and Fourier analyses,” *Journal of Nonlinear Science*, vol. 22, no. 6, pp. 887–915, 2012.
- [79] J. N. Kutz, *Data-Driven Modeling & Scientific Computation: Methods for Complex Systems & Big Data*. Oxford University Press, Oxford, U.K., 2013.
- [80] M. Gavish and D. L. Donoho, “The optimal hard threshold for singular values is $4/\sqrt{3}$,” *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 5040–5053, 2014.
- [81] M. O. Williams, C. W. Rowley, and I. G. Kevrekidis, “A kernel-based method for data-driven Koopman spectral analysis,” *Journal of Computational Dynamics*, vol. 2, no. 2, pp. 247–265, 2015.
- [82] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, “Chaos as an intermittently forced linear system,” *Nature Communications*, vol. 8, 2017.
- [83] D. Duke, J. Soria, and D. Honnery, “An error analysis of the dynamic mode decomposition,” *Experiments in Fluids*, vol. 52, no. 2, pp. 529–542, 2012.
- [84] S. Bagheri, “Effects of weak noise on oscillating flows: Linking quality factor, Floquet modes, and Koopman spectrum,” *Physics of Fluids*, vol. 26, no. 9, 094104, 2014.
- [85] S. T. Dawson, M. S. Hemati, M. O. Williams, and C. W. Rowley, “Characterizing and correcting for the effect of sensor noise in the dynamic mode decomposition,” *Experiments in Fluids*, vol. 57, no. 3, pp. 1–19, 2016.
- [86] M. S. Hemati, C. W. Rowley, E. A. Deem, and L. N. Cattafesta, “De-biasing the dynamic mode decomposition for applied Koopman spectral analysis of noisy datasets,” *Theoretical and Computational Fluid Dynamics*, pp. 1–20, 2017.

- [87] H. Knudsen and J. N. Nielsen, "Introduction to the modelling of wind turbines," in *Wind Power in Power Systems* (T. Ackermann, ed.), pp. 767–797, John Wiley & Sons, Ltd, Chichester, U.K, 2012.
- [88] O. Wasynczuk, D. Man, and J. Sullivan, "Dynamic behavior of a class of wind turbine generators during random wind fluctuations," *IEEE Transactions on Power Apparatus and Systems*, vol. PER-1, no. 6, pp. 2837–2845, 1981.
- [89] A. Ellis, Y. Kazachkov, J. Sanchez-Gasca, P. Pourbeik, E. Muljadi, M. Behnke, J. Fortmann, and S. Seman, "Reduced-order modelling of wind turbines," in *Wind Power in Power Systems* (T. Ackermann, ed.), pp. 799–820, John Wiley & Sons Ltd, Chichester, U.K, 2012.
- [90] E. Bossanyi, "Wind turbine control for load reduction," *Wind Energy*, vol. 6, no. 3, pp. 229–244, 2003.
- [91] S. Müller, M. Deicke, and R. W. De Doncker, "Doubly fed induction generator systems for wind turbines," *IEEE Industry Applications Magazine*, vol. 8, no. 3, pp. 26–33, 2002.
- [92] K. Elkington, J. G. H. Slootweg, M. Ghandhari, and W. L. Kling, "Reduced-order modelling of wind turbines," in *Wind Power in Power Systems* (T. Ackermann, ed.), pp. 821–847, John Wiley & Sons, Ltd, Chichester, U.K, 2012.
- [93] V. Yaramasu, B. Wu, P. C. Sen, S. Kouro, and M. Narimani, "High-power wind energy conversion systems: State-of-the-art and emerging technologies," *Proceedings of the IEEE*, vol. 103, no. 5, pp. 740–788, 2015.
- [94] T. Ackermann and L. Söder, "Wind power in power systems: An introduction," in *Wind Power in Power Systems* (T. Ackermann, ed.), pp. 25–51, John Wiley & Sons, Ltd, Chichester, U.K, 2012.
- [95] J. G. Slootweg, S. W. H. de Haan, H. Polinder, and W. L. Kling, "General model for representing variable speed wind turbines in power system dynamics simulations," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 144–151, 2003.
- [96] F. Mei and B. Pal, "Modal analysis of grid-connected doubly fed induction generators," *IEEE Transactions on Energy Conversion*, vol. 22, no. 3, pp. 728–736, 2007.
- [97] F. Milano, *Power System Modelling and Scripting*. London, U.K.: Springer, 2010.
- [98] F. Milano, "An open source power system analysis toolbox," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1199–1206, 2005.
- [99] H.-W. Kim, S.-S. Kim, and H.-S. Ko, "Modeling and control of PMSG-based variable-speed wind turbine," *Electric Power Systems Research*, vol. 80, no. 1, pp. 46–52, 2010.

- [100] J. G. Slootweg, *Wind Power Modelling and Impact on Power System Dynamics*. PhD thesis, Delft University of Technology, 2003.
- [101] K. Tsuboki and A. Sakakibara, “Large-scale parallel computing of cloud resolving storm simulator,” in *High Performance Computing* (H. P. Zima, K. Joe, M. Sato, Y. Seo, and M. Shimasaki, eds.), vol. 2327 of *Lecture Notes in Computer Science*, pp. 243–259, Springer Berlin Heidelberg, 2002.
- [102] On the status of wind power introduction, Agency for Natural Resources and Energy, 2015. (in Japanese) [Online]. Available at http://www.meti.go.jp/committee/sougouenergy/shoene_shinene/shin_ene/pdf/009_04_00.pdf (Last accessed on 24 July 2017).
- [103] P. Anderson and A. Bose, “Stability simulation of wind turbine systems,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-102, no. 12, pp. 3791–3795, 1983.
- [104] J. G. Slootweg, H. Polinder, and W. Kling, “Initialization of wind turbine models in power system dynamics simulations,” in *2001 IEEE Porto Power Tech Proceedings*, vol. 4, (Porto, Portugal), pp. 162–167, 2001.
- [105] V. Akhmatov, H. Knudsen, A. Hejde Nielsen, J. Kaas Pedersen, and N. Kjølstad Poulsen, “Modelling and transient stability of large wind farms,” *International Journal of Electrical Power and Energy Systems*, vol. 25, no. 2, pp. 123–144, 2003.
- [106] Global Wind Energy Council (GWEC), *Global Wind Report Annual Market Update*, 2015. [Online]. Available at http://www.gwec.net/wp-content/uploads/vip/GWEC-Global-Wind-2015-Report-April-2016_22_04.pdf (Last accessed on 24 July 2017).
- [107] M. Anvari, G. Lohmann, M. Wächter, P. Milan, E. Lorenz, D. Heinemann, M. R. R. Tabar, and J. Peinke, “Short term fluctuations of wind and solar power systems,” *New Journal of Physics*, vol. 18, no. 6, 063027, 2016.
- [108] O. Kamps, “Characterizing the fluctuations of wind power production by multi-time statistics,” in *Wind Energy - Impact of Turbulence* (M. Hölling, J. Peinke, and S. Ivanell, eds.), pp. 67–72, Springer Berlin Heidelberg, 2014.
- [109] S. Weitemeyer, D. Kleinhans, T. Vogt, and C. Agert, “Integration of renewable energy sources in future power systems: The role of storage,” *Renewable Energy*, vol. 75, pp. 14–20, 2015.
- [110] P. Milan, M. Wächter, and J. Peinke, “Stochastic modeling and performance monitoring of wind farm power production,” *Journal of Renewable and Sustainable Energy*, vol. 6, no. 3, 033119, 2014.

- [111] L. Ljung, *System Identification: Theory for the User*. Prentice-Hall information and system sciences series, Prentice-Hall, Upper Saddle River, NJ, 1999.
- [112] A. Morales, M. Wächter, and J. Peinke, “Characterization of wind turbulence by higher-order statistics,” *Wind Energy*, vol. 15, no. 3, pp. 391–406, 2012.
- [113] International Electrotechnical Commission (IEC), *IEC 61400-1 Wind turbine generator systems - Part 1: Design requirements*, 2005.
- [114] F. K. Davis and H. Newstein, “The variation of gust factors with mean wind speed and with height,” *Journal of Applied Meteorology*, vol. 7, no. 3, pp. 372–378, 1968.
- [115] A. Bossavy, R. Girard, and G. Kariniotakis, “Forecasting ramps of wind power production with numerical weather prediction ensembles,” *Wind Energy*, vol. 16, no. 1, pp. 51–63, 2013.
- [116] D. B. Wright and J. A. Herrington, “Problematic standard errors and confidence intervals for skewness and kurtosis,” *Behavior Research Methods*, vol. 43, no. 1, pp. 8–17, 2011.
- [117] M. M. Bandi, “Spectrum of wind power fluctuations,” *Physical Review Letters*, vol. 118, no. 2, 028301, 2017.
- [118] M. Courtney, R. Wagner, and P. Lindelöw, “Commercial lidar profilers for wind energy: A comparative guide,” in *Proc. European Wind Energy Conference, Brussels, Belgium*, 2008.
- [119] T. Mikkelsen, “Lidar-based research and innovation at DTU wind energy – a review,” *Journal of Physics: Conference Series*, vol. 524, no. 1, 012007, 2014.
- [120] D. McQueen and S. Watson, “Validation of wind speed prediction methods at off-shore sites,” *Wind Energy*, vol. 9, no. 1-2, pp. 75–85, 2006.
- [121] D. Carvalho, A. Rocha, C. S. Santos, and R. Pereira, “Wind resource modelling in complex terrain using different mesoscale-microscale coupling techniques,” *Applied Energy*, vol. 108, pp. 493–504, 2013.
- [122] F. Castro, C. Silva Santos, and J. Lopes da Costa, “One-way mesoscale-microscale coupling for the simulation of atmospheric flows over complex terrain,” *Wind Energy*, vol. 18, no. 7, pp. 1251–1272, 2015.
- [123] M. Bilal, K. Solbakken, and Y. Birkelund, “Wind speed and direction predictions by WRF and WindSim coupling over Nygårdsfjell,” *Journal of Physics: Conference Series*, vol. 753, no. 8, 082018, 2016.

- [124] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, “The state-of-the-art in short-term prediction of wind power: A literature overview,” tech. rep., ANEMOS. plus, 2011. Available at http://www.anemos-plus.eu/images/pubs/deliverables/aplus.deliverable_d1.2.stp_sota_v1.1.pdf (Last accessed on 24 July 2017).
- [125] J. Jung and R. P. Broadwater, “Current status and future advances for wind speed and power forecasting,” *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 762–777, 2014.
- [126] D. Carvalho, A. Rocha, M. Gómez-Gesteira, and C. S. Santos, “WRF wind simulation and wind energy production estimates forced by different reanalyses: Comparison with observed data for Portugal,” *Applied Energy*, vol. 117, pp. 116–126, 2014.
- [127] O. Krogsæter and J. Reuder, “Validation of boundary layer parameterization schemes in the weather research and forecasting model under the aspect of offshore wind energy applications—Part I: Average wind speed and wind shear,” *Wind Energy*, vol. 18, no. 5, pp. 769–782, 2015.
- [128] D. Carvalho, A. Rocha, M. Gómez-Gesteira, and C. Santos, “A sensitivity study of the WRF model in wind simulation for an area of high wind energy,” *Environmental Modelling & Software*, vol. 33, pp. 23–34, 2012.
- [129] N. J. Cutler, H. R. Outhred, and I. F. MacGill, “Using nacelle-based wind speed observations to improve power curve modeling for wind power forecasting,” *Wind Energy*, vol. 15, no. 2, pp. 245–258, 2012.
- [130] Y. Che, X. Peng, L. Delle Monache, T. Kawaguchi, and F. Xiao, “A wind power forecasting system based on the weather research and forecasting model and Kalman filtering over a wind-farm in Japan,” *Journal of Renewable and Sustainable Energy*, vol. 8, no. 1, 013302, 2016.
- [131] X.-M. Hu, P. M. Klein, and M. Xue, “Evaluation of the updated YSU planetary boundary layer scheme within WRF for wind resource and air quality assessments,” *Journal of Geophysical Research: Atmospheres*, vol. 118, no. 18, pp. 10,490–10,505, 2013.
- [132] C. Ferreira, J. Gama, L. Matias, A. Botterud, and J. Wang, “A survey on wind power ramp forecasting,” tech. rep., Argonne National Laboratory, Argonne, Illinois, USA, 2010. Available at <http://www.anl.gov/energy-systems/publication/survey-wind-power-ramp-forecasting> (Last accessed on 24 July 2017).
- [133] M. Negnevitsky and P. Johnson, “Very short term wind power prediction: A data mining approach,” in *2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, (Pittsburgh, Pennsylvania, USA), pp. 1–3, July 2008.

- [134] B. Greaves, J. Collins, J. Parkes, and A. Tindal, “Temporal forecast uncertainty for ramp events,” *Wind Engineering*, vol. 33, no. 4, pp. 309–319, 2009.
- [135] T. Nanahara, M. Asari, T. Sato, K. Yamaguchi, M. Shibata, and T. Maejima, “Smoothing effects of distributed wind turbines. Part 1. Coherence and smoothing effects at a wind farm,” *Wind Energy*, vol. 7, no. 2, pp. 61–74, 2004.
- [136] A. Viguera-Rodríguez, P. Sørensen, A. Viedma, M. Donovan, and E. G. Lázaro, “Spectral coherence model for power fluctuations in a wind farm,” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 102, pp. 14–21, 2012.
- [137] A. Smaïli and C. Masson, “On the rotor effects upon nacelle anemometry for wind turbines,” *Wind Engineering*, vol. 28, no. 6, pp. 695–713, 2004.
- [138] Y.-H. Wan, E. Ela, and K. Orwig, “Development of an equivalent wind plant power-curve,” in *Proc. WindPower*, (Dallas, Texas, USA), pp. 1–20, 2010.
- [139] B. Smith, H. Link, G. Randall, and T. McCoy, “Applicability of nacelle anemometer measurements for use in turbine power performance tests,” in *Proceedings of the American Wind Energy Association (AWEA) WINDPOWER 2002*, (Portland, Oregon, USA), 2002. Available at <http://www.nrel.gov/docs/fy02osti/32494.pdf> (Last accessed on 24 July 2017).
- [140] I. Antoniou and T. Friis Pedersen, “Nacelle anemometry on a 1MW wind turbine: Comparing the power performance results by use of the nacelle or mast anemometer. Risø National Laboratory, Risø-R-941(EN), Roskilde, Denmark,” 1997. Available at http://orbit.dtu.dk/files/7753492/RIS_R_941.pdf (Last accessed on 24 July 2017).
- [141] P. Berens, “Circstat: A MATLAB toolbox for circular statistics,” *Journal of Statistical Software*, vol. 31, no. 1, pp. 1–21, 2009.
- [142] K. Tsuboki and A. Sakakibara, *Numerical Prediction of High-Impact Weather Systems: The textbook for seventeenth IHP Training Course in 2007*, 2007. Available at http://www.rain.hyarc.nagoya-u.ac.jp/~tsuboki/cress_html/src_cress/CReSS2223_users_guide_eng.pdf (Last accessed on 24 July 2017).
- [143] A. Arakawa and V. R. Lamb, “Computational design of the basic dynamical processes of the UCLA general circulation model,” in *General Circulation Models of the Atmosphere* (J. Chang, ed.), vol. 17 of *Methods in Computational Physics: Advances in Research and Applications*, pp. 173–265, Elsevier, 1977.
- [144] N. Salvação, M. Bernardino, and C. G. Soares, “Assessing mesoscale wind simulations in different environments,” *Computers & Geosciences*, vol. 71, pp. 28–36, 2014.

- [145] J. B. Klemp and R. B. Wilhelmson, “The simulation of three-dimensional convective storm dynamics,” *Journal of the Atmospheric Sciences*, vol. 35, no. 6, pp. 1070–1096, 1978.
- [146] M. Lange, “On the uncertainty of wind power predictions—analysis of the forecast accuracy and statistical distribution of errors,” *Journal of Solar Energy Engineering*, vol. 127, pp. 177–184, 2005.
- [147] D. Carvalho, A. Rocha, and M. Gómez-Gesteira, “Ocean surface wind simulation forced by different reanalyses: Comparison with observed data along the Iberian Peninsula coast,” *Ocean Modelling*, vol. 56, pp. 31–42, 2012.
- [148] Q. Yang, L. K. Berg, M. Pekour, J. D. Fast, R. K. Newsom, M. Stoelinga, and C. Finley, “Evaluation of WRF-predicted near-hub-height winds and ramp events over a Pacific Northwest site with complex terrain,” *Journal of Applied Meteorology and Climatology*, vol. 52, no. 8, pp. 1753–1763, 2013.
- [149] T. Nanahara, M. Asari, T. Sato, K. Yamaguchi, M. Shibata, and T. Maejima, “Smoothing effects of distributed wind turbines. Part 2. Coherence among power output of distant wind turbines,” *Wind Energy*, vol. 7, no. 2, pp. 75–85, 2004.
- [150] J. Apt, “The spectrum of power from wind turbines,” *Journal of Power Sources*, vol. 169, no. 2, pp. 369–374, 2007.
- [151] M. R. R. Tabar, M. Anvari, G. Lohmann, D. Heinemann, M. Wächter, P. Milan, E. Lorenz, and J. Peinke, “Kolmogorov spectrum of renewable wind and solar power fluctuations,” *The European Physical Journal Special Topics*, vol. 223, no. 12, pp. 2637–2644, 2014.
- [152] J. Kaimal, J. Wyngaard, Y. Izumi, and O. Coté, “Spectral characteristics of surface-layer turbulence,” *Quarterly Journal of the Royal Meteorological Society*, vol. 98, no. 417, pp. 563–589, 1972.
- [153] F. Fernandez-Bernal and J. Alonso-Alonso, “Wind speed generation for dynamic analysis,” *Wind Energy*, vol. 20, no. 6, pp. 1049–1068, 2017.
- [154] M. Ragheb and A. M. Ragheb, “Wind turbines theory - the Betz equation and optimal rotor tip speed ratio,” in *Fundamental and Advanced Topics in Wind Power* (R. Cariveau, ed.), InTech, 2011.
- [155] J. Seguro and T. Lambert, “Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis,” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 85, no. 1, pp. 75–84, 2000.
- [156] L. Mahrt, “Surface wind direction variability,” *Journal of Applied Meteorology and Climatology*, vol. 50, no. 1, pp. 144–152, 2011.

- [157] H. Louie, “Correlation and statistical characteristics of aggregate wind power in large transcontinental systems,” *Wind Energy*, vol. 17, no. 6, pp. 793–810, 2014.
- [158] G. McNerney and R. Richardson, “The statistical smoothing of power delivered to utilities by multiple wind turbines,” *IEEE Transactions on Energy Conversion*, vol. 7, no. 4, pp. 644–647, 1992.
- [159] H. G. Beyer, J. Luther, and R. Steinberger-Willms, “Power fluctuations in spatially dispersed wind turbine systems,” *Solar Energy*, vol. 50, no. 4, pp. 297–305, 1993.
- [160] G. Giebel, *On the benefits of distributed generation of wind energy in Europe*. PhD thesis, University of Oldenburg, 2000.
- [161] B. Tarroja, F. Mueller, J. D. Eichman, J. Brouwer, and S. Samuelson, “Spatial and temporal analysis of electric wind generation intermittency and dynamics,” *Renewable Energy*, vol. 36, no. 12, pp. 3424–3432, 2011.
- [162] W. Katzenstein, E. Fertig, and J. Apt, “The variability of interconnected wind plants,” *Energy Policy*, vol. 38, no. 8, pp. 4400–4410, 2010.
- [163] F. Raak, Y. Susuki, K. Tsuboki, M. Kato, and T. Hikiyara, “On smoothing effects of wind power via Koopman mode decomposition,” in *2016 International Symposium on Nonlinear Theory and its Applications (NOLTA2016)*, (Yugawara, Japan), pp. 366–369, November 2016.
- [164] Y.-H. Wan, M. Milligan, and B. Parsons, “Output power correlation between adjacent wind power plants,” *Journal of Solar Energy Engineering*, vol. 125, no. 4, pp. 551–555, 2003.
- [165] C. St. Martin, J. Lundquist, and M. Handschy, “Variability of interconnected wind plants: Correlation length and its dependence on variability time scale,” *Environmental Research Letters*, vol. 10, no. 4, 044004, 2015.
- [166] X. Zhu and M. G. Genton, “Short-term wind speed forecasting for power system operations,” *International Statistical Review*, vol. 80, no. 1, pp. 2–23, 2012.
- [167] L. Cavalcante, R. J. Bessa, M. Reis, and J. Browell, “LASSO vector autoregression structures for very short-term wind power forecasting,” *Wind Energy*, vol. 20, no. 4, pp. 657–675, 2017.
- [168] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge University Press, Cambridge, U.K., 2004.
- [169] Y. Hirata, D. P. Mandic, H. Suzuki, and K. Aihara, “Wind direction modelling using multiple observation points,” *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 366, no. 1865, pp. 591–607, 2008.

- [170] S. A. Pourmousavi Kani and M. M. Ardehali, "Very short-term wind speed prediction: A new artificial neural network–Markov chain model," *Energy Conversion and Management*, vol. 52, no. 1, pp. 738–745, 2011.
- [171] M. Ragwitz and H. Kantz, "Detecting non-linear structure and predicting turbulent gusts in surface wind velocities," *Europhysics Letters*, vol. 51, no. 6, pp. 595–601, 2000.
- [172] R. Wayland, D. Bromley, D. Pickett, and A. Passamante, "Recognizing determinism in a time series," *Physical Review Letters*, vol. 70, no. 5, pp. 580–582, 1993.
- [173] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D: Nonlinear Phenomena*, vol. 110, no. 1–2, pp. 43–50, 1997.
- [174] M. Ragwitz and H. Kantz, "Markov models from data by simple nonlinear time series predictors in delay embedding spaces," *Physical Review E*, vol. 65, no. 5, 056201, 2002.
- [175] T. Sauer, J. A. Yorke, and M. Casdagli, "Embedology," *Journal of Statistical Physics*, vol. 65, no. 3, pp. 579–616, 1991.
- [176] M. Abdullah, A. Yatim, C. Tan, and R. Saidur, "A review of maximum power point tracking algorithms for wind energy systems," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 5, pp. 3220–3227, 2012.
- [177] E. Welfonder, R. Neifer, and M. Spanner, "Development and experimental identification of dynamic models for wind turbines," *Control Engineering Practice*, vol. 5, no. 1, pp. 63–73, 1997.
- [178] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical Review A*, vol. 33, no. 2, pp. 1134–1140, 1986.
- [179] H. Abarbanel, *Analysis of Observed Chaotic Data*. Springer-Verlag, New York, 1996.
- [180] F. Böttcher, C. Renner, H.-P. Waldl, and J. Peinke, "On the statistics of wind gusts," *Boundary-Layer Meteorology*, vol. 108, no. 1, pp. 163–173, 2003.
- [181] A. G. Roy, P. M. Biron, and M. F. Lapointe, "Implications of low-pass filtering on power spectra and autocorrelation functions of turbulent velocity signals," *Mathematical Geology*, vol. 29, no. 5, pp. 653–668, 1997.
- [182] G. Sugihara and R. M. May, "Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series," *Nature*, vol. 344, pp. 734–741, 1990.
- [183] I. Mezić and A. Banaszuk, "Comparison of systems with complex behavior," *Physica D: Nonlinear Phenomena*, vol. 197, no. 1, pp. 101–133, 2004.

- [184] J. Hauer, C. Demeure, and L. Scharf, "Initial results in Prony analysis of power system response signals," *IEEE Transactions on Power Systems*, vol. 5, no. 1, pp. 80–89, 1990.
- [185] J. D. Farmer and J. J. Sidorowich, "Predicting chaotic time series," *Physical Review Letters*, vol. 59, no. 8, pp. 845–848, 1987.
- [186] F. Hamilton, T. Berry, and T. Sauer, "Ensemble Kalman filtering without a model," *Physical Review X*, vol. 6, no. 1, 011021, 2016.
- [187] H. You, V. Vittal, and Z. Yang, "Self-healing in power systems: An approach using islanding and rate of frequency decline-based load shedding," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 174–181, 2003.
- [188] H. You, V. Vittal, and X. Wang, "Slow coherency-based islanding," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 483–491, 2004.
- [189] B. Yang, V. Vittal, and G. T. Heydt, "Slow-coherency-based controlled islanding - A demonstration of the approach on the august 14, 2003 blackout scenario," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1840–1847, 2006.
- [190] P. V. Kokotovic, B. Avramovic, J. H. Chow, and J. R. Winkelman, "Coherency based decomposition and aggregation," *Automatica*, vol. 18, no. 1, pp. 47–56, 1982.
- [191] S. Yusof, G. Rogers, and R. Alden, "Slow coherency based network partitioning including load buses," *IEEE Transactions on Power Systems*, vol. 8, no. 3, pp. 1375–1381, 1993.
- [192] K. Sun, D. . Zheng, and Q. Lu, "Splitting strategies for islanding operation of large-scale power systems using OBDD-based methods," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 912–923, 2003.
- [193] K. Sun, D. . Zheng, and Q. Lu, "A simulation study of OBDD-based proper splitting strategies for power systems under consideration of transient stability," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 389–399, 2005.
- [194] K. Sun, D.-Z. Zheng, and Q. Lu, "Searching for feasible splitting strategies of controlled system islanding," *IEE Proceedings: Generation, Transmission and Distribution*, vol. 153, no. 1, pp. 89–98, 2006.
- [195] L. Ding, F. Gonzalez-Longatt, P. Wall, and V. Terzija, "Two-step spectral clustering controlled islanding algorithm," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 75–84, 2013.
- [196] B. Avramovic, P. V. Kokotovic, J. R. Winkelman, and J. H. Chow, "Area decomposition for electromechanical models of power systems.," *Automatica*, vol. 16, no. 6, pp. 637–648, 1980.

- [197] J. Winkelman, J. Chow, J. Allemong, and P. Kokotovic, "Multi-time-scale analysis of a power system," *Automatica*, vol. 16, no. 1, pp. 35–43, 1980.
- [198] J. R. Winkelman, J. Chow, B. C. Bowler, B. Avramovic, and P. Kokotovic, "An analysis of interarea dynamics of multi-machine systems," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 2, pp. 754–763, 1981.
- [199] S. Koch, S. Chatzivasileiadis, M. Vrakopoulou, and G. Andersson, "Mitigation of cascading failures by real-time controlled islanding and graceful load shedding," in *2010 IREP Symposium Bulk Power System Dynamics and Control - VIII (IREP)*, (Buzios, RJ, Brazil), pp. 1–19, 2010.
- [200] P. V. Kokotovic, J. J. Allemong, J. R. Winkelman, and J. H. Chow, "Singular perturbation and iterative separation of time scales," *Automatica*, vol. 16, no. 1, pp. 23–33, 1980.
- [201] M. Fiedler, "Algebraic connectivity of graphs," *Czechoslovak Mathematical Journal*, vol. 23, no. 98, pp. 298–305, 1973.
- [202] M. A. Riolo and M. E. J. Newman, "First-principles multiway spectral partitioning of graphs," *Journal of Complex Networks*, vol. 2, no. 2, pp. 121–140, 2014.
- [203] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithms: Analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, 2002.
- [204] Y.-F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, "State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 33–43, 2012.
- [205] Power Systems Test Case Archive, College of Engineering, University of Washington. [Online]. Available at <http://www.ee.washington.edu/research/pstca/> (Last accessed on 24 July 2017).
- [206] H.-D. Chiang and L. Fekih-Ahmed, "On the direct method for transient stability analysis of power system structure preserving models," in *[Proceedings] 1992 IEEE International Symposium on Circuits and Systems*, vol. 5, (San Diego, California, USA), pp. 2545–2548 vol.5, 1992.
- [207] Q. Zhao, K. Sun, D.-Z. Zheng, J. Ma, and Q. Lu, "A study of system splitting strategies for island operation of power system: A two-phase method based on OBDDs," *IEEE Transactions on Power Systems*, vol. 18, no. 4, pp. 1556–1565, 2003.

- [208] Y. Deng, H. Lin, S. Shukla, J. Thorp, and L. Mili, “Co-simulating power systems and communication network for accurate modeling and simulation of pmu based wide area measurement systems using a global event scheduling technique,” in *2013 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, (Berkeley, California, USA), pp. 1–6, May 2013.
- [209] A. Mauroy and J. Goncalves, “Linear identification of nonlinear systems: A lifting technique based on the Koopman operator,” in *2016 IEEE 55th Conference on Decision and Control (CDC)*, (Las Vegas, Nevada, USA), pp. 6500–6505, December 2016.
- [210] A. Mauroy and J. Hendrickx, “Spectral identification of networks using sparse measurements,” *SIAM Journal on Applied Dynamical Systems*, vol. 16, no. 1, pp. 479–513, 2017.
- [211] A. Tsonis and P. Roebber, “The architecture of the climate network,” *Physica A: Statistical Mechanics and its Applications*, vol. 333, pp. 497–504, 2004.
- [212] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths, “Complex networks in climate dynamics,” *The European Physical Journal-Special Topics*, vol. 174, no. 1, pp. 157–179, 2009.
- [213] N. Boers, B. Bookhagen, H. M. Barbosa, N. Marwan, J. Kurths, and J. Marengo, “Prediction of extreme floods in the eastern central andes based on a complex networks approach,” *Nature Communications*, vol. 5, 5199, 2014.
- [214] M. Gelbrecht, N. Boers, and J. Kurths, “A complex network representation of wind flows,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 27, no. 3, 035808, 2017.
- [215] F. Milano and R. Zárate-Miñano, “A systematic method to model power systems as stochastic differential algebraic equations,” *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4537–4544, 2013.
- [216] Z. Y. Dong, J. H. Zhao, and D. J. Hill, “Numerical simulation for stochastic transient stability assessment,” *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1741–1749, 2012.
- [217] S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Discovering governing equations from data by sparse identification of nonlinear dynamical systems,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–3937, 2016.

List of Publications

Many of the results, discussions, and figures in this dissertation are based on the following papers and presentations (I do not claim copyright for any of them).

Journal Papers

First Author

Published

- [1] F. Raak, Y. Susuki, and T. Hikiyara, “Data-driven partitioning of power networks via Koopman mode analysis,” *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2799–2808, 2016.

Accepted

- [2] F. Raak, Y. Susuki, K. Tsuboki, M. Kato, and T. Hikiyara, “Quantifying smoothing effects of wind power via Koopman mode decomposition: A numerical test with wind speed predictions in Japan,” *Nonlinear Theory and Its Applications, IEICE*, vol. 8, no. 4, 2017.

Submitted

- [3] F. Raak, Y. Susuki, K. Tsuboki, M. Kato, S. Eguchi, and T. Hikiyara, “Assessment of offshore wind farm characteristics with the cloud resolving storm simulator: A case study in Japan,” 2017. In revision.

In Preparation

- [4] F. Raak, Y. Susuki, S. Eguchi, and T. Hikiyara, “Statistics of short-term wind power fluctuations in Japan via measurements and a dynamic wind turbine model,” 2017. Tentative.
- [5] F. Raak, Y. Susuki, and T. Hikiyara, “Time-series prediction via Koopman mode decomposition,” 2017. Tentative.

Co-Author

Published

- [6] Y. Susuki, I. Mezić, F. Raak, and T. Hikiyara, “Applied Koopman operator theory for power systems technology,” *Nonlinear Theory and Its Applications, IEICE*, vol. 7, no. 4, pp. 430–459, 2016.

Peer-Reviewed International Conference Proceedings

- [7] F. Raak, Y. Susuki, I. Mezić, and T. Hikiyara, “On Koopman and dynamic mode decompositions for application to dynamic data with low spatial dimension,” in *2016 IEEE 55th Conference on Decision and Control (CDC)*, (Las Vegas, Nevada, USA), pp. 6485–6491, December, 2016.
- [8] F. Raak, Y. Susuki, K. Tsuboki, M. Kato, and T. Hikiyara, “On smoothing effects of wind power via Koopman mode decomposition,” in *2016 International Symposium on Nonlinear Theory and its Applications (NOLTA2016)*, (Yugawara, Japan), pp. 366–369, November 2016.
- [9] F. Raak, Y. Susuki, and T. Hikiyara, “Multi-way partitioning of power networks via koopman mode analysis,” *IFAC-PapersOnLine*, vol. 48, no. 30, pp. 421–426, 2015. 9th IFAC Symposium on Control of Power and Energy Systems CPES 2015, (New Delhi, India).
- [10] F. Raak, Y. Susuki, R. Morita, T. Wada, K. Tsuboki, H. Uyeda, Y. Fujisaki, and T. Hikiyara, “Dynamic simulation of a wind farm incorporated with highly-resolved wind prediction data,” in *Industrial Electronics Society, IECON 2015 - 41st Annual Conference of the IEEE*, (Yokohama, Japan), pp. 004406–004411, November 2015.
- [11] F. Raak, Y. Susuki, T. Hikiyara, H. R. Chamorro, and M. Ghandhari, “Partitioning power grids via nonlinear Koopman mode analysis,” in *Innovative Smart Grid Technologies Conference (ISGT), 2014 IEEE PES*, (Washington D.C., USA), pp. 1–5, IEEE, February 2014.

Other Conference Proceedings and Presentations

First Author

- [12] F. Raak, Y. Susuki, I. Mezić, and T. Hikiyara, “Comparison of dynamic mode decomposition, Koopman mode decomposition, and vector Prony analysis.” SIAM Conference on Applications of Dynamical Systems (SIAM DS17), Snowbird, Utah, USA, May 21–25, 2017.

- [13] F. Raak, Y. Susuki, K. Tsuboki, M. Kato, S. Eguchi, and T. Hikihara, “Wind power simulation and analysis incorporating highly-resolved weather prediction and measurement data of Japan,” in *The 15th World Wind Energy Conference and Exhibition (WWEC2016)*, (Tokyo, Japan), November 2016.
- [14] F. Raak, Y. Susuki, S. Eguchi, and T. Hikihara, “Statistical test of dynamic model of a wind turbine incorporated with measured data in Japan,” in *60th Annual Conference of the Institute of Systems, Control and Information Engineers (ISCIE)*, 311-2, (Kyoto, Japan), May 2016.
- [15] F. Raak, Y. Susuki, K. Tsuboki, M. Kato, and T. Hikihara, “Quantifying smoothing effects of wind power via Koopman modes,” in *Proceedings of The 2016 IEEE Annual Meeting*, 7-080, (Sendai, Japan), March 2016.
- [16] F. Raak, Y. Susuki, R. Morita, T. Wada, K. Tsuboki, H. Uyeda, Y. Fujisaki, and T. Hikihara, “Simulation study of a wind farm incorporated with highly-resolved wind prediction data around Japan,” in *37th National Symposium on Wind Energy*, (Tokyo, Japan), pp. 173–176, November 2015.
- [17] F. Raak, Y. Susuki, R. Morita, T. Wada, K. Tsuboki, H. Uyeda, Y. Fujisaki, and T. Hikihara, “Dynamic simulation of a wind turbine incorporated with highly-resolved wind prediction data,” in *59th Annual Conference of the Institute of Systems, Control and Information Engineers (ISCIE)*, 126-3, (Osaka, Japan), May 2015.
- [18] F. Raak, Y. Susuki, T. Hikihara, H. Chamorro, and M. Ghandhari, “Connecting graph and dynamic properties of power grids with spectrum of Koopman operator,” in *58th Annual Conference of the Institute of Systems, Control and Information Engineers (ISCIE)*, 112-5, (Kyoto, Japan), May 2014.
- [19] F. Raak, Y. Susuki, T. Hikihara, and H. Chamorro, “Investigation of power grid islanding based on nonlinear Koopman modes,” in *IEICE Technical Report, Nonlinear Problems*, vol. 113, (Miyakojima, Japan), pp. 75–80, July 2013.
- [20] F. Raak, Y. Susuki, and T. Hikihara, “Koopman mode analysis of inter-area oscillations in the Nordic 32-bus test system,” in *Proceedings of The 2013 IEEE Annual Meeting*, 6-084, (Nagoya, Japan), March 2013.

Co-Author

- [21] R. Hamasaki, Y. Susuki, F. Raak, and A. Ishigame, “On identification of power system inertia via Koopman mode decomposition,” in *IEEE Power and Energy Society Conference*, (Tokyo, Japan), September 2017. (In Japanese).
- [22] Y. Susuki, K. Sako, F. Raak, and T. Hikihara, “Assessment of voltage collapse phenomena in power grids based on continuous spectrum of the Koopman operator.”

SIAM Conference on Applications of Dynamical Systems (SIAM DS17), Snowbird, Utah, USA, May 21–25, 2017.

[23] K. Sako, Y. Susuki, F. Raak, and T. Hikihara, “Time-series analysis of voltage collapse via continuous spectrum of the Koopman operator,” in *Kansai-section Joint Convention of Institutes of Electrical Engineering*, (Osaka, Japan), pp. 196–197, November 2016. (In Japanese).

[24] Y. Ota, Y. Susuki, F. Raak, and I. Mezić, “Koopman mode analysis of synchrophasor data in Japanese 60-Hz power system,” in *Proceedings of The 2015 IEEJ Annual Meeting*, 6-107, (Tokyo, Japan), March 2015.