Informatics

Data Driven Detection of Misconfigurations in Power Distribution Systems

DISSERTATION

zur Erlangung des akademischen Grades

Doktor der Technischen Wissenschaften

eingereicht von

Dipl.-Ing. David Fellner, B.Sc.
Matrikelnummer 01226059

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ.Prof. Dipl.-Ing. Dr.techn. Wolfgang Kastner
Zweitbetreuung: Priv.-Doz. Dipl.-Ing. Dr.techn. Thomas I. Strasser

Diese Dissertation haben begutachtet:

Ferdinanda Ponci

Geert Deconinck

Wien, 27. Juli 2023

David Fellner

TU WIEN Informatics

Data Driven Detection of Misconfigurations in Power Distribution Systems

DISSERTATION

submitted in partial fulfillment of the requirements for the degree of

Doktor der Technischen Wissenschaften

by

Dipl.-Ing. David Fellner, B.Sc.
Registration Number 01226059

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Prof. Dipl.-Ing. Dr.techn. Wolfgang Kastner
Second advisor: Priv.-Doz. Dipl.-Ing. Dr.techn. Thomas I. Strasser

The dissertation has been reviewed by:

Ferdinanda Ponci	Geert Deconinck

Vienna, 27th July, 2023

David Fellner

Erklärung zur Verfassung der Arbeit

Dipl.-Ing. David Fellner, B.Sc.

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 27. Juli 2023

David Fellner

Danksagung

An erster Stelle möchte ich meinen Betreuern Thomas I. Strasser und Wolfgang Kastner für ihre tatkräftige Unterstützung danken. Ihre wertvollen Anmerkungen, profunde Expertise und die von ihnen geteilten Erfahrungen haben wesentlich zum Entstehen dieser Arbeit beigetragen.

Weiters gilt mein Dank meinen Kolleginnen und Kollegen am AIT Austrian Institute of Technology. Ganz besonders Helfried Brunner und Mark Stefan, welche mir die Chance gegeben haben, an dieser Dissertation zu arbeiten. Paul Zehetbauer und Michael Spiegel, die mir mit vielen praktischen Tipps weitergeholfen haben und von denen ich vieles lernen durfte. Schließlich hat mich die Zusammenarbeit mit Sarah Reisenbauer bestärkt und motiviert, an der Vervollständigung der Dissertation zu arbeiten.

Meinen Freunden gebührt großer Dank, denn sie haben stets meine Tage erhellt und mich auch durch die langen Winter der COVID-19 Pandemie zuversichtlich begleitet. Ganz besonders möchte ich mich bei meiner wunderbaren Partnerin Lara bedanken, die mich immer aufgefangen hat und somit einen großen Beitrag zur guten Vollendung der Arbeit geleistet hat. Selbstverständlich will ich mich bei meiner Familie bedanken. Bei meinem Bruder Klaus, der mir auch in diesem Lebensabschnitt als Orientierungshilfe gedient hat. Am meisten aber natürlich bei meinen Eltern, ohne die keine meiner Leistungen jemals möglich gewesen wäre, und die mich auch während meiner Dissertation stets begleitet und unterstützt haben.

Acknowledgements

First and foremost, I would like to thank my supervisors Thomas I. Strasser and Wolfgang Kastner for their energetic support. Their valuable comments, profound expertise, and shared experiences have significantly contributed to the creation of this work.

I would also like to thank my colleagues at the AIT Austrian Institute of Technology. In particular, Helfried Brunner and Mark Stefan, who gave me the opportunity to work on this dissertation, further Paul Zehetbauer and Michael Spiegel, who helped me with many practical tips and from whom I learned a lot. Finally, the cooperation with Sarah Reisenbauer encouraged and motivated me to work on the completion of the dissertation.

My friends also deserve special recognition; they have always brightened my days and accompanied me confidently through the long winters of the COVID-19 pandemic. I would also like to thank my wonderful partner Lara who always caught me and thus made a great contribution to the completion of the work. Of course, I want to thank my family. My brother Klaus also served as a guide in this phase of life. But most of all, of course, my parents, without whom none of my achievements would ever have been possible, and who accompanied and supported me throughout my dissertation.

Kurzfassung

Das elektrische Energiesystem unterliegt Veränderungen durch die nachhaltige Bereitstellung von Energie und deren Verbrauch. Diese Veränderungen wirken sich insbesondere auf die Verteilnetze aus, da diese für den Anschluss sehr gut vorhersehbarer Haushaltslasten ausgelegt sind. Heutzutage müssen diese Netze in zunehmendem Maße dezentrale erneuerbare Erzeuger mit schwankenden Einspeisecharakteristiken sowie neuartige elektrifizierte Lasten wie Elektrofahrzeuge oder Heizsysteme aufnehmen. Die Integration dieser neuartigen Geräte führt zu Problemen wie Spannungsbandverletzungen oder Überlastungen des Netzes und stellt somit eine Herausforderung für den sicheren und zuverlässigen Betrieb des Stromnetzes dar. Um solchen Schwierigkeiten entgegenzuwirken, beinhalten diese neuen Geräte oft netzstützende Regelungsfunktionen wie Blindleistungsregelkurven oder Ladestromregelungen. Aus historischen Gründen fehlt es dem Verteilnetz jedoch an Messkapazitäten. Die Netzbetreiber haben daher keine Möglichkeit sicherzustellen, dass die notwendigen netzstützenden Funktionalitäten wie vorgesehen funktionieren. Folglich werden neue Überwachungsmöglichkeiten benötigt.

In dieser Arbeit wird ein datengetriebener Überwachungsansatz vorgestellt, der in Verteilnetzen funktionsfähig ist. Es werden zunächst Bewertungen der verfügbaren Daten und ihrer Eigenschaften, wie z. B. Abtastraten oder Einschränkungen aufgrund von Datenschutzbestimmungen, durchgeführt. Anschließend werden Methoden zur Erkennung von Fehlkonfigurationen der Steuerfunktionen von netzgekoppelten Geräten entwickelt. Die entdeckten Anomalien werden auch durch Klassifizierungsmethoden kategorisiert. Die Methoden sind an die verfügbaren Daten angepasst: am Verteiltransformator sind dies traditionelle Machine Learning (ML) Methoden, auf der Geräteebene werden Deep Learning (DL) Methoden eingesetzt. Außerdem werden Data-Mining-Methoden entwickelt und getestet, um trotz des Mangels an Sensoren Informationen über das Niederspannungsnetz (LV) zu gewinnen. Diese Methoden werden anhand von im Labor gesammelten Daten sowie durch Simulationen validiert. Außerdem wird die Entwicklungsumgebung beschrieben, die für die Analyse der Methoden und die Erstellung der Simulationsdaten verwendet wurde.

Die Ergebnisse der Arbeit stellen die Ansätze zur Überwachung netzunterstützender Funktionalitäten in elektrischen Verteilnetzen dar. Sie zeigen, dass diese unter Verwendung der aktuellen Mess- und Zählerinfrastruktur integriert werden können und eine gute Erkennungs- und Klassifizierungsgüte aufweisen, die die Implementierung eines nützlichen Entscheidungsunterstützungswerkzeugs für Verteilnetzbetreiber (DSO) ermöglicht.

Abstract

The modern power system is undergoing fundamental changes in order to adapt to new requirements, caused by the need of sustainable provision and consumption of energy. These changes especially affect power distribution systems, as they have been designed to host very predictable and similar household loads. Nowadays, they are increasingly required to connect decentralized renewable generation which shows volatile power infeed characteristics as well as novel electrified loads such as electric vehicles or heating systems. The integration of these novel devices can create problems of voltage band violations or overloading to the grid and therefore be a challenge for the safe and reliable operation of the power grid. To counter such difficulties, the already mentioned new devices often implement grid-supporting control functionalities as reactive power control curves or charging current controls. However, due to the former sole purpose of passing on energy in a very foreseeable and one-directional manner, the distribution grid lacks sensory capacities. Therefore, grid operators have no way of ensuring that the necessary grid-supporting functionalities work as intended. Thus, new monitoring capabilities are needed.

As an outcome, the thesis provides a layout of a data-driven monitoring approach which works under the mentioned circumstances found in distribution grids. Assessments of the available data and their properties such as sampling rates or restrictions due to privacy issues were conducted. Following, methods were developed to detect misconfigurations of grid-connected devices' control functionalities. These detected abnormalities are also categorized by classification methods. These methods are adjusted to the data available: at substation level these are traditional Machine Learning (ML) methods, at the device level Deep Learning (DL) methods are employed. Also, data mining methods are developed and assessed to gain information about the Low Voltage (LV) grid despite the lack of sensors. The methods are tested and validated using data collected in laboratory setup as well as through simulations, which are in turn also validated using the laboratory data. In addition, the development framework used for developing and assessing the methods and creating the simulation data is described.

The results of the thesis present the best suited approaches for monitoring grid-supporting functionalities in electrical distribution grids. They show that these functionalities can be integrated with the current sensing and metering infrastructure and show a good detection and classification performance which enables the implementation of a meaningful decision-support tool for Distribution Grid Operators (DSO).

Contents

Preface

The present dissertation is carried out in a cumulative manner. A compilation of four peer-reviewed journal and conference papers is presented and forms the integral part of this thesis. The publications contributing to the thesis are listed below:

1. D. Fellner, T. I. Strasser, and W. Kastner, "Applying deep learning-based concepts for the detection of device misconfigurations in power systems," Sustainable Energy, Grids and Networks, vol. 32, p. 100851, 2022.
   Available: https://doi.org/10.1016/j.segan.2022.100851

2. D. Fellner, T. I. Strasser, W. Kastner, B. Feizifar and I. F. Abdulhadi, "Data Driven Transformer Level Misconfiguration Detection in Power Distribution Grids," 2022 IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC), Prague, Czech Republic, 2022, pp. 1840-1847. doi: 10.1109/SMC53654.2022.9945534

3. D. Fellner, T. I. Strasser and W. Kastner, "Data-Driven Misconfiguration Detection in Power Systems with Transformer Profile Disaggregation," IEEE Access, vol. 11, pp. 80123-80136, 2023. doi: 10.1109/ACCESS.2023.3300236

4. D. Fellner, T. I. Strasser and W. Kastner, "The DeMaDs Open Source Modeling Framework for Power System Malfunction Detection," 2023 Open Source Modelling and Simulation of Energy Systems (OSMSES), Aachen, Germany, 2023, pp. 1-6. doi: 10.1109/OSMSES58477.2023.10089746

The published articles are presented in Chapters 2 - 5. The order of these chapters is not strictly chronological and chosen to facilitate reading. The layout of the papers is aligned to maintain readability and unified referencing, but the content of each chapter is identical to the original publication. A reference to the underlying publication is given at the beginning of each chapter.

Doctoral School Smart Industrial Concept! (NextGeneration SIC!)

NextGeneration SIC![1] is a cooperative doctoral college of the TU Wien that complements industrial expertise with scientific excellence to form a well-rounded overall picture. Science is at the center of the program. The close cooperation with our partners from other research institutions – the AIT Austrian Institute of Technology (AIT), Montanuniversität Leoben (MUL) – and from industry – A1, EVN, evon and Fundermax – is of central importance.

As part of the doctoral school, the thesis contributes mainly to the thematic area of data processing and data driven methods. It provides methods and concepts for data handling as well as systems integration, which can be used to build a monitoring solution for electric energy distribution systems. The presented methods and concepts in this thesis can be seen as a foundation for further research and development work in the other areas of the SIC! consortium.

Industrial PhD - FFG Grant

This work received funding from the Austrian Research Promotion Agency (FFG) under the "Research Partnerships – Industrial Ph.D. Program" in Data Driven Detection of Malfunctioning Devices in Power Distribution Systems (DeMaDs) (FFG No. 879017).

[1]https://www.tuwien.at/en/doc/sic

CHAPTER 1

Introduction

1.1 Motivation and Problem Statement

Electricity grid operators are facing multiple difficulties due to the massive changes in the energy system, which are necessary for the shift to a sustainable energy system. These challenges can range from regulatory barriers to environmental issues and technical problems in the field of energy storage and transmission. Especially a high density of distributed generation can have a significant impact on the grid. The replacement of traditional generation from central power plants by inverter-interfaced generation, such as photovoltaic (PV) power plants, is one example. This leads to problems with frequency stability as inertia provided by the rotational masses of the traditional generators needs to be replaced [1]. As the system frequency is a global variable in the power grid, this poses a problem for all parts of it. This is also true for reliability issues, which can occur at a high penetration of volatile renewable energy sources such as wind and solar in the grid. At levels exceeding 20-30%, the N-1 reliability criterion of the grid can not be ensured in the traditional manner [2]. However, major challenges also occur especially with regard to distribution grids as they will bear the brunt of Distributed Generation (DG) integration of renewable energy sources. Moreover, the operation of distribution grids will also be affected severely by mobility and heating electrification. Both lead to higher volatility in grid operation through unknown generation and demand profiles and to higher demand, in general [3].

Other impacts of the decentralization of power generation and electrification of loads are local voltage limit violations, congested distribution lines as well as overloaded transformers. This can for example be caused by electric vehicle (EV) charging which has an impact on the bus voltage or increases demand. Therefore, the voltage may fall beneath admissible limits. To handle these problems, costly grid reinforcements may be necessary [4]. The opposite is the case for the integration of the widespread penetration of DGs. The distribution network was not conceived to handle the previously unseen power infeed in decentral locations. As a result, local congestions and overvoltages can occur which are violating the upper voltage limits [5].

To prevent these unwelcome effects without limiting renewable energy generation or investing heavily in costly grid reinforcement, certain relief strategies can be employed. One of them is the On-Load Tap Changer (OLTC) which can be used to control and adjust the voltage centrally at the substation by altering the winding ratio between the primary high voltage side and the secondary low voltage side [6]. However, OLTCs only pose a solution in cases in which the voltage profile is either too high or too low throughout the entire grid segment. Furthermore, they also only provide relief in case of voltage problems, leaving congestion or overloading unattended. One way of facing all of these problems is by making loads more flexible allowing for so-called Demand Side Management (DSM). This can for example come in the form of operating Thermostatically Controlled Loads (TLC) in a grid supporting manner [7]. These can be for example the already mentioned electrified heating systems, such as heat pumps. Their thermal inertia and large power consumption allow to use them to counteract the high production or consumption of other grid participants. The scheduling of EV charging can be employed in a similar manner [8]. Local renewable generation, such as building integrated wind energy, can for example be used to charge EVs by scheduling the two to coincide [9]. This can help to mitigate both overvoltages or congestions due to overproduction as well as undervoltages due to too high local consumption. There are numerous strategies in literature to satisfy these goals both using local renewable generation as well as thermal loads or charging schedules [10]. Also, Real-Time Charging Control (RTC) can help to regulate EV charging power directly and instantly in order not to violate constraints and avoid overloading of the grid [11]. The same control can be used to feed energy back from the electric vehicles to the grid (V2G) in order to bolster the local voltage or support the grid in general. Nevertheless, voltage regulation is considered as the most important aspect when it comes to integrating decentral generation in distribution networks, which is implemented by grid supporting functionalities provided by the generation units [12].

These functionalities can range from limiting the active power dispatched to controlling the reactive power injection of the generation units or Battery Energy Storage Systems (BESS) with inverters [13]. Combined PV and BESS systems can also be used for both frequency control by providing Frequency Containment Reserve (FCR), as well as voltage control through reactive power provided by the inverter [14]. Grid supporting functionalities may also come in the form of peak reduction and increased self-consumption, which can also be motivated by flexible capacity tariffs [15].

At the moment, voltage control is done using local reactive power control, for example in the form of a droop control using local measurements as inputs [16]. However, this control does not allow for optimal operation of DGs and does also not guarantee that voltages are actually kept within limits. Therefore, coordinated voltage controls are superior for this application [17].

The depiction in Figure 1.1 shows a local power factor control curve of an inverter-connected generation unit in order to better illustrate the principle behind such grid

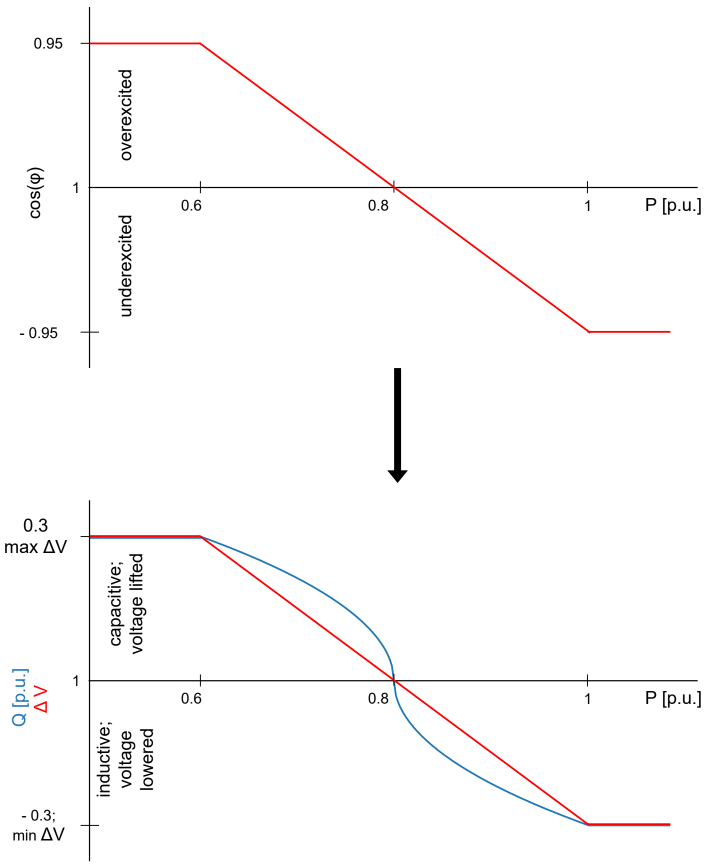


Figure 1.1: cosϕ(p) power factor control (top) and resulting reactive power p as well as voltage lift (bottom)

supporting functionalities. The top part of the figure shows the alteration of the power factor depending on the active power fed in by, for example, a PV. At low active power infeed the power factor is overexcited, which leads, as depicted in the bottom part of the figure, to a capacitive reactive power dispatch. This in turn helps to bolster the voltage by lifting it. At high active power dispatch, the power factor is underexcited, which has the opposite effect of lowering the voltage by feeding in inductive reactive power.

Grid connected generation units are required to implement certain grid supporting functionalities as defined by national regulations [18]. A common form of grid supporting functionalities are the already mentioned reactive power controls in order to control the local voltage. The regulations usually define certain capability curves. The left side of Figure 1.2 depicts such a capability curve as it is mandated by the grid code of a country. This specification defines the amount of reactive power that, for example, an inverter connected generation unit must provide to the grid at a certain infeed of active power. The right side of the figure shows a possible implementation of this specification. The configuration depicted is the already elaborated cosϕ(p) power factor control curve.

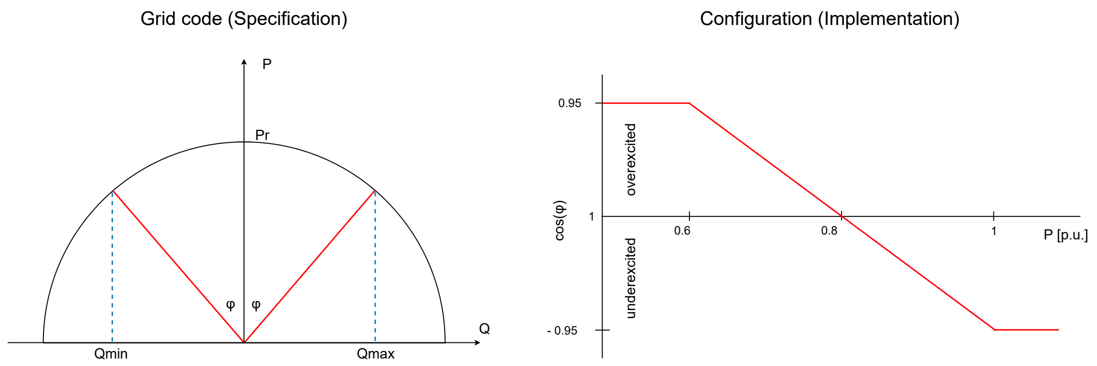


Figure 1.2: Specification as defined by the grid code (left) and possible implementation as configuration (right)

As illustrated earlier, these controls can take very different configurations, as they can be dependent on central coordination or local variables and react to different control variables in general. As these controls and the support functionalities they provide are critical to a safe and reliable operation of the grid Distribution System Operators (DSO) need to be able to ensure that they are actually performed correctly. Currently, DSOs are lacking the possibility to properly monitor their grids, especially decentrally installed devices. However, as the control can vary, it can be difficult to determine whether a certain configuration is the intended one, as more than one can satisfy a certain specification. Moreover, a certain configuration can change due to malfunctions or software glitches. This may result in a misconfiguration which in turn disables the desired control and grid supporting functionality. Therefore, misconfigurations need to be detected as they pose a threat to grid operation, especially if new generation and load devices are rolled out at a large scale.

However, the circumstances for the detection of such misconfigurations can be difficult. Grid data is not up to date or incorrect, apart from the fact that configurations can differ, as already elaborated. To adjust to various as well as changing circumstances and settings, a data driven solution is preferable. Additionally, an approach using operational data is therefore advisable, because the required knowledge about grid component characteristics can be kept to a minimum. Because of legal restrictions concerning data privacy or the general lack of measurements the data necessary is often not available, especially at connection points [19]. Preferably, monitoring is to be conducted remotely, even if Smart Meters (SM) are installed. One reliable source of data that is currently available to DSOs are substation measurement data [20]. At the transformer level, various channels such as currents, voltages, or power flows are captured at a high rate. If SM data are to be used, DSOs often need to obtain the permission of customers. Furthermore, these data are only sampled every 15 minutes [21]. Both factors hamper the applicability of local SM data for monitoring purposes. However, certain properties of the monitoring problem at hand can also be exploited when crafting a relevant solution. A misconfiguration is not a transient event and dynamics are therefore not of interest [22]. This means lower data sampling rates are satisfactory as only operational changes on a slower time scale are to be supervised. Therefore, data of a resolution of 15 minutes might be sufficient for the task in certain applications, as the time ranges of the effects on the grid are similar. Detection is less time critical as a not shifted or incorrectly scheduled load or a non functional control might only have an impact after a few hours or even whenoccurring in greater numbers. This means avoiding or fixing them one by one at a reasonable pace might be sufficient, even if that means they are not tackled instantaneously. For this reason, it is to be determined which measurements are required such as active and reactive power flows or voltage and current magnitudes, as well as at what sampling rate these data are needed for a suitable and applicable monitoring solution. Furthermore, it is to be explored which methods can be employed and applied to these data to actually detect misconfigurations. In the following the identified sub-problems and the related works are further elaborated.

1.1.1 Anomaly Detection

As the detection of misconfigurations can be regarded as an anomaly detection problem, Kernel Principle Component Analysis (kPCA) [23], offers an interesting approach. It can be used to extract key features of the operational data. Furthermore, a statistical model of the regular state of a system can be built using it [24]. Due to the unsupervised character of this learning method, it allows capturing a predominant stable nominal state along with numerous possible abnormal states, about which only little information is available. Both grid participants at the low voltage level as well as the substation transformer are expected to show such behavior. Hidden structures in the operational data collected can be revealed using Partially Hidden Structured Support Vector Machine (pSVM) [25]. The pSVM allows to incorporate labeled, partially labeled as well as unlabeled data and the information stored therein. The same principle is used in [24] to map relations between 'different' and 'stable' events. Another approach to anomaly detection for Heating, Ventilation, and Air Conditioning (HVAC) using a one-class support vector machine is explored in [26]. Most variability in the data occurs during the regular operation of the system and not due to anomalies, as the results show. This also applies to households or PV systems. Therefore, a combination of Primary Component Analysis (PCA) and Support Vector Machine (SVM) could be applied to the data sources at the medium voltage level, as the data there are of high dimensionality. On the low voltage level, a Deep Learning (DL) approach could be more promising. As the data is of a lower time resolution as well as dimensionality, filtering out relevant features can become much harder. Additionally, the grid environment is different for every single grid participant due to their location in the grid, which further complicates the detection of abnormal states. DL can help in this regard, as it is well suited to extract detection features autonomously and condense the special traits of certain anomalies. It could then be used decentrally to detect certain anomalies [27]. The non-functional requirements of integrability as well as scalability are at the focus of these methods: It is not necessary to manually fit the solution to a specific device characteristic or transformer setup. A feasible and rapid rollout of such a solution is therefore possible. Another non functional requirement targeted is adaptability. After changes in configurations or the setup, the solution does not have to be adapted manually. The collection of data merely needs to be restarted.

1.1.2 Classification of Abnormal Behaviour

The already mentioned pSVM [24] could be implemented for the classification of unusual behavior found. The first two steps of detecting and classifying anomalies could also be combined using this approach. This might not be feasible on the low voltage level, since, as mentioned above, the biggest variability in operation, and therefore the operational data, occurs during the regular operation of households or PV systems. Artificial Neural Networks (ANN) could be used for this purpose [28]. Fast Fourier transformed voltage waveforms of normal and abnormal data are required here to train the ANN. As already elaborated DL could be employed to conduct this preprocessing step in a data driven way [29]. A Multiclass Relevance Vector Machine (mRVM) [30] uses auxiliary variables as intermediate regression targets and provides probabilities of class membership and therefore a classification. Usually, only binary class membership classification is conducted, whereas the multiclass classification can point grid operators in a specific direction concerning incorrect behavior of devices, and due to this the approach could be helpful. This could also be provided by a clustering or SVM approach combined with a feature selection using Decision Trees [31]. The non functional requirement of usability is targeted by these approaches: Misconfigured grid connected devices need to be detectable for DSOs in order to gain knowledge about the current state of the grid. Information about the origin of the abnormal behavior detected is necessary for a useful tool that enhances the system operator's understanding. Non-functional requirements concerning third-party components need to be considered here in addition. However, grid connected devices providing the above-mentioned grid support functionalities are mostly not installed nor operated by the DSO, but by other actors or entities. It is a requirement to monitor these lacking full information about their specifications.

1.1.3 Disaggregation of Medium Voltage Profiles

In order to make information about decentral grid participants available while respecting limitations set by either privacy issues or lacking sensing capacities, the load profile at the medium voltage substation can be split up into its contributions. This is usually referred to as disaggregation, which [32] uses an ANN for. SM data of households as well as of generation units or other grid connected devices can be used to create an appliance signature database. Here, every grid participant could be treated as an appliance. The signatures could either be regarded as load profiles during entire days or weeks or treated as routine load or dispatch patterns at specific times, as a PVinverter dispatch profile during the morning for example. The former would allow for the detection of gradual alterations in operation which should be sufficient for monitoring. The neural network is trained on the so-defined signatures. Using the so gained information, the neural network could distinguish contributions to the aggregated substation profile, as to yield the partial profiles timely. As an alternative, a hybrid Support Vector Machine/Gaussian Mixture Model (SVM-GMM) [33] could be used. This method has the advantage of creating a power feature model for appliances on its own when these devices are turned on, not needing SM data. Nevertheless, it has to be determined if this approach can be used on grid participants, as it is probably not easy to determine when to start collecting the feature models as such a clear point of turning on is not given. The central non functional requirement here is data retention. Data privacy laws are protecting data and are inhibiting its usage. This means access to the data is limited or not possible at all. One of the main aspects under this point is designing the function for data mining so as to work in accordance with this requirement.

1.1.4 Software, Data Acquisition and Validation

To start implementing the concepts and approaches presented and to develop a monitoring solution, a software framework is to be constructed. This framework must be used to synthesize operational data employing grid simulation software in the first step. The data generated in this manner ought to include operational data from distribution substations as well as from grid participants at the low voltage level. The channels to be captured should include currents, voltages as well as power flows at applicable time resolutions, such as 15 minutes down to a few Hertz frequencies. The necessary sampling rates are to be determined for the individual measurement points. The scenarios to be simulated here should comprise regular operation as well as malfunctions, such as misconfigured PVinverters or Electrical Vehicle Supply Equipment (EVSE) dispatching or consuming energy in a manner that is not intended to support grid operation. The substation data can either be captured at the medium or low voltage side. The low voltage distribution grid participant data can also be collected using simulation. Data are to be gathered in a 15-minute resolution to mimic SM data in order to evaluate disaggregation as well as for the general purposes of monitoring method evaluation such as anomaly detection and classification. These data ought to be collected at the corresponding connection point of SMs in order to allow for as life-like conditions as possible. Furthermore, various grid setups should be used in order to address the non-functional requirement of robustness. As data from different grid setups have to be handled equally well, this aims at ensuring robustness. The evaluation and testing of algorithms are to be conducted individually but also in combination with other parts of the complete solution in order to evaluate the entire 'pipeline' of methods. When the algorithms are tested individually, the other parts of the method are assumed to work with perfect accuracy. This also points to the most rewarding and biggest potential for improvement of the entire solution. Key Performance Indicators (KPI) are to be defined. These can be measures such as precision which is the misclassification rate, or recall, which is listed in order to be able to assess how many misconfigurations of the ones present actually get detected. The former is of interest in order to prevent false alarms, and the latter in order to be sure the solution provides a reliable monitoring solution. The non-functional requirement of quality is to be satisfied here. Comparing the performance of the found solution against a benchmark can help ensure that this requirement is met.

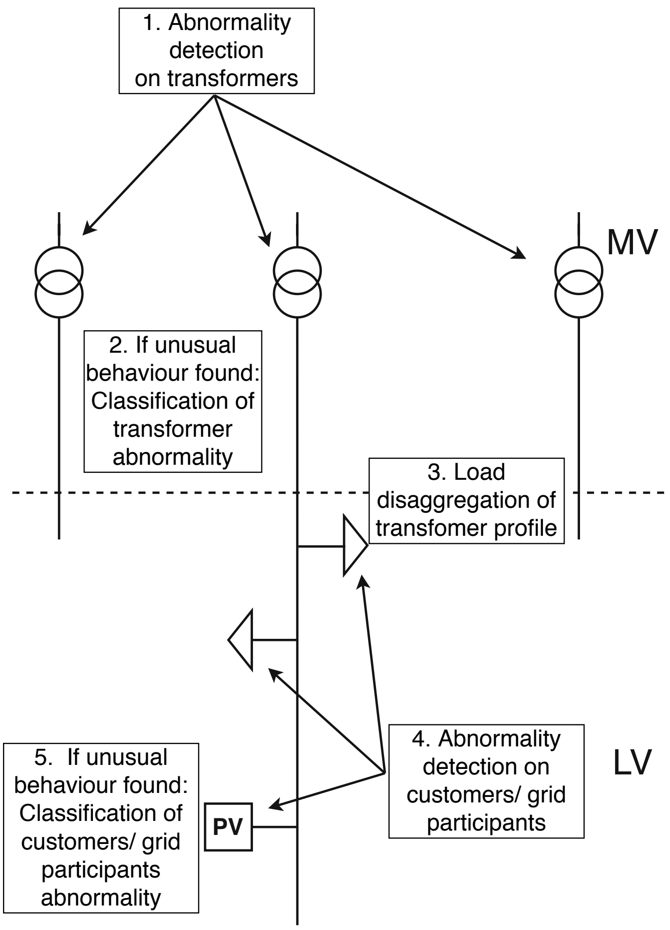


Figure 1.3: General concept [34]

The challenges and approaches identified above lead to the concept depicted in Figure 1.3, which shows how the said 'pipeline' can be assembled in order to create an architecture that allows monitoring of grid connected devices and their supporting functionalities. The depiction shows transformers, the two overlapping rings, connecting the medium voltage MV grid to the low voltage LV grid. Furthermore, it shows a PV unit as well as loads, symbolized by the triangles, on the low voltage level.

The machine learning algorithms described before are to be applied to data spanning from operational data of generation units or novel loads integrated with household load profiles to transformers experiencing typical load profiles. This aims at developing a control misconfiguration detection which is different from fault condition monitoring approaches which can be found in literature [35]. Data of abnormal behavior as well as of regular operation are necessary, as in the case of regular or incorrectly parameterized control curves. At least party-labeled data is going to be needed to enable the training of algorithms for classification [36]. In addition to the generated data obtained through simulation, real-world data are provided by DSOs or collected in laboratory settings that emulate actual grid conditions. These data can then be used for validation of the methods but also of the simulations generating data. Ultimately, the monitoring approach is to be tested on a test site, which should be constituted by an actual part of the grid.

1.2 Aims and Research Questions

1.2.1 Research Questions

The research opportunity identified concerning monitoring decentralized grid connected devices and detection of malfunctions which influence the provided grid supporting functionalities was formulated in one main research question and refined by splitting it into sub-research questions:

> *Main Research Question: Which approach, applying data driven algorithms and methods on operational data at the medium and low voltage level is best suited to detect misconfigurations of grid connected devices' grid supporting functions in a low voltage power distribution grid?*

This question can be answered by taking several consecutive steps to treat the operational data available. Methods for anomaly detection, disaggregation of load profiles, or classification already exist. Combining and developing them to adapt to the task, however, is yet to be studied. An architecture is therefore proposed which uses data collected at medium voltage transformers as well as information about the low voltage grid which is underlying as inputs. The architecture then applies various data driven approaches on these data. The employed data encompasses measurements of currents, voltages as well as power flows. Doing so should enable the detection and identification of grid participants whose grid supporting functionalities show abnormal behavior, meaning not the behavior they should display according to the configuration parameterized in accordance with the specifications. Therefore, prior to a malfunction that changes this configuration, the device is assumed to show behavior corresponding to the implementation intended. Providing such a monitoring solution should allow, for example, to determine automatically and remotely whether control schemes of distributed generation units installed at the low voltage level are executed correctly.

The exploration and development of this approach entail several sub-research questions (SRQ):

- *SRQ1: How can abnormal transformer load profiles or other unusual sensor readings be detected using solely aggregated measurement data recorded at medium voltage substations?* The first sub question is aimed at identifying these transformers and their underlying grid segments amongst the great number of transformers in the domain of a grid operator. The main point of interest is how a stable state model using operational data for each transformer can be built, as well as how deviations from this model can be detected. Furthermore, investigations into which sorts of operational data, such as voltage, currents, or power flows are of use for this task.

- *SRQ2: How can abnormal transformer behavior be distinguished depending on the fundamental cause?* The detected anomalous transformer behavior should be classified, given it is identified to be unusual. This means, finding the device causing the detected abnormal behavior on the low voltage level responsible for the anomaly on the medium voltage level. To do this, deviations of controllable loads or distributed generation units from their behavior observed up to that point could be detected. This could be implemented by employing the disaggregated load profiles.
- *SRQ3: How is it possible, based on the medium voltage data, to obtain information about the behavior of grid connected devices on the low voltage level?* Further insights into the operation of the underlying low voltage grid are necessary, which can be gathered based on the data employed to identify and classify the unusual medium voltage transformer behavior in SRQ1 and SRQ2. This should help to make statements about its exact cause. Exploring the disaggregation of the medium voltage transformer load profile on the low voltage side can be used here to obtain the load and generation profiles of individual customers or households and other grid connected devices, such as Battery Energy Storage Systems (BESS), Heat Pumps (HP), Electric vehicles EV, PV inverters and households as depicted in Figure 1.4. It remains to be explored how this can be implemented through a data driven approach, aided by smart meter data [37] to build load signatures along with grid data and topology to improve performance.
- *SRQ4: Which quality of data is necessary and from which sources to deliver a certain detection performance?* The needed type, such as power flow readings, currents or voltages, and resolution of data is to be determined. What properties of the data, such as generality in the data meaning, for example, data are collected from different grid types, is necessary to create a robust solution is to be investigated. Moreover, the desired performance is to be determined and evaluated by developing or using certain indicators as Key Performance Indicators (KPI)s.

1.2.2 Aims

The overall aim of the thesis is to provide concepts and methods for a monitoring solution of grid connected devices in power distribution grids. Therefore, the following four goals are defined and put into relation to the research questions mentioned above.

- *G1: Develop a data driven method to detect abnormal substation transformer readings.* A general approach to anomaly detection on the medium voltage substation

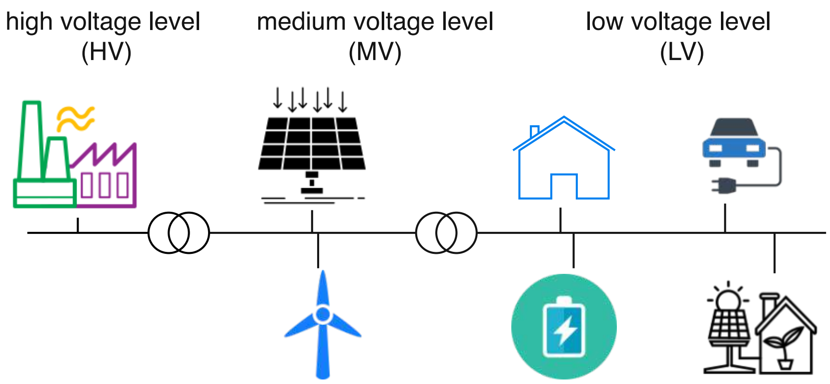


Figure 1.4: Voltage levels and components in power systems [34]

level is missing for implementing monitoring approaches for grid connected devices in the electric power distribution grids. This approach should use operational data and respect constraints on its availability. This goal is targeted by sub-research question SRQ1.

- *G2: Provide a method to relate abnormal behavior to its underlying cause.* Having insights into the background of anomalies is very important for a monitoring solution to actually be a helpful tool. These insights can be provided by the classification of detected anomalies into predefined types of misconfigurations, which are to be looked out for. Thus, sub-research question SRQ2 is targeting this need to point to the cause of an anomaly detected.
- *G3: Design an approach to mine information about the low voltage distribution grid given centrally aggregated data.* As the cause of an anomaly rooted on the low voltage level might not be easily determinable given just readings at a substation level, some form of information extraction is necessary. This is to be done through disaggregation of the aggregated medium voltage substation power profiles in order to gain information about its contributors. This goal is addressed by sub-research question SRQ3.
- *G4: Assess the data needed and the performance yielded by a potential monitoring solution.* A monitoring solution is only viable if it works with the data at hand and provides monitoring at a certain level of confidence. Therefore, an assessment of data properties, quality, and origin as well as monitoring performance is necessary. This goal is tackled by sub-research question SRQ4.

1.3 Methodology

In this section, the methodology is described which is applied to give answers to the above-stated research questions and reach the aims of the thesis. As the research questions are partly related to each other and sometimes depend on each other, approaches are sometimes applied and evaluated independently and combined later. Also, the combined solution is then evaluated. Thus, the works and publications were not conducted in a strictly chronological order. Neither do they fit necessarily to the same order as the stated research questions. Especially SRQ4 spans all publications as it targets the data foundations as well as performance evaluation. In general, the research questions were answered using explorative research which was qualitatively evaluated by developing proof-of-concept implementations, a method which is applied to certain case studies. In addition, it is to be mentioned that a comparison of other methods found in literature was already presented earlier. Following, a more elaborate description of the applied methods for answering each of the research questions is given. Additionally, Figure 1.5 gives another overview of where the respective SRQs and Goals are to be localized and which parts of the grid they can be assigned to. The symbols are the same ones for the transformer, loads, and PV as elaborated above.

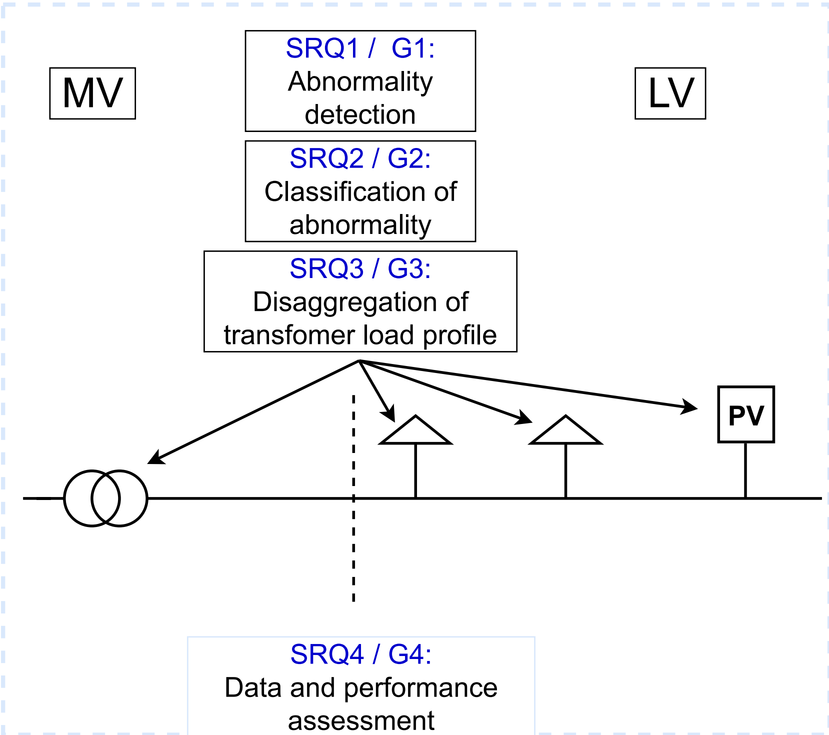


Figure 1.5: SRQs and Goals localized and assigned to grid parts

Applied Methods for SRQ1 - Detection of abnormal behavior: As done initially a literature review was conducted to identify suitable approaches to the problem of detection of abnormal behavior. In general, two opportunities to detect abnormal behavior were found: either detection performed on aggregated data at the substation level on the one hand, or detection at the device level using data at the connection point of the device. Both approaches were explored also while bearing the thought of combining the two and their monitoring capabilities in mind. For the transformer-level detection approach traditional machine-learning methods were applied. The rationale for this was that the properties of the substation data are grid specific since every substation connects a different low-voltage grid to the medium-voltage distribution grid. Furthermore, the data available is collected at a higher sampling rate and a higher dimensionality as at the device level, since the sampling rate is higher as well as the number of channels, and therefore, variable, recorded. Lastly, also changes in the underlying grid have to be accommodated, and therefore, frequent retraining of the detection model is necessary. This is only feasible with the more lightweight models provided by traditional learning. These included Primary Component Analysis (PCA) [38] for dimensionality reduction and feature extraction as well as the k Nearest Neighbours (kNN) [39], Decision Tree (DT) [40] and Support Vector Machine (SVM) [41] classifiers to determine whether a state is abnormal or not. At the device level, DL approaches were employed and compared to one another and to a benchmark provided by traditional Machine Learning (ML) methods. These were chosen for the opposite reasons as mentioned for the substation level detection: data are only available at lower temporal resolutions as well as in a lower dimensionality, as only for example voltage magnitudes might be available. Additionally, the circumstances under which these data come into existence are mainly static, as readings at the device level and the impacts of misconfigurations on them are rather grid independent. Therefore, DL approaches that can extract general and universal features of traces of misconfigurations in the data were chosen as the method to be explored here. As the data available are time series data and temporal relations are crucial, mainly Recurrent Neural Network (RNN) [42] approaches were under scrutiny. All of these were compared and also hyperparameters were tuned putting in a lot of effort.

Applied Methods for SRQ2 - Classification of abnormal behavior: Also here a review of the literature related to the classification of abnormal behavior was carried out beforehand. As for the task of detecting unusual behavior, classifying it in order to point to an underlying cause can be done both at the low voltage level device connection point as well as at the medium voltage substation transformer level. At the device level, again for the reasons of greater universal applicability as mentioned above, DL methods were employed. As the data at hand were still time series, also here forms RNNs were mostly employed. These included the basic RNN, the Long-Short-Term-Memory (LSTM) [43] and Gated-Recurrent-Unit (GRU) [44] architectures of RNNs. Furthermore, Neural Networks (NN) with attention mechanisms, as the Transformer architecture [45] as the only non-recurrent architecture, as well as the R-Transformer [46], an architecture combining recurrent properties with the attention mechanism, were applied. For each type of misconfiguration to be detected a separate Neural Network was trained in order to be able to use them to distinguish between causes of unusual behavior. This network was trained using data from different grids as well as from different points in this grid, as the so-trained models are to be used universally. A similar approach was taken on the task of determining the kind of unusual behavior detected at the substation level. The already mentioned PCA was used to extract features of the time series data, which is of the same importance here as it is in the detection case due to the high dimensionality of the data available here. Due to the individual properties of the underlying grids elaborated above, a separate classifier was needed for each grid, which is also supposed to be adaptable to changes in the said grid. Moreover, either a separate classifier was built for each misconfiguration which could potentially be the cause of the detected unusual behavior or the misconfiguration was added as a faulty class. Meaning there are either multiple binary classifiers per grid or a multi-class classifier. Also here, the kNN, DT, and SVM classifiers were compared in order to determine the best performing but also the best set of hyperparameters for each of the algorithms.

Applied Methods for SRQ3 - Disaggregation of substation load profiles: Based on a literature review the available approaches were analyzed as presented initially. The predominant form of disaggregation is generally known as Non-intrusive Load Monitoring (NILM) [47] which refers to disaggregating household load profiles into its contributing application load profiles, often using appliance signatures that are highly specific to a certain appliance type, such as a dishwasher. Here the problem to be tackled is similar, as the aggregated medium voltage substation transformer profile is to be disaggregated into the individual load and generation profiles attributed to the devices or households on the low voltage level. These profiles might be similar to each other in some cases, however, they are never alike as the consumption pattern of the same household appliance. This obliterates the possibility of using appliance signatures in a similar manner as they are used in NILM. However, disaggregation is an important capability in order to enable some of the other questions to be answered and goals to be fulfilled. Therefore, another approach was chosen; For each grid setup, an array of simulations was conducted calculating load flows and recreating a great number of operational situations. Load and generation values were uniformly distributed in the profiles, ranging from the minimum to the maximum value seen during regular operation, and assigned to the respective grid participants. This created a dataset covering all expected states the grid could be in and captures the voltage, current, and load flow values for these. Due to this, the grid properties are engrained in this dataset. This dataset was then used to train an Artificial Neural Networks in the form of a simple Feedforward Neural Network (FNN) [48]. The so-trained model was then used to estimate the loads at the low voltage level given the aggregated load at the medium voltage substation along with voltage measurements at neuralgic points of the low voltage grid. Additionally, the same data was used to build a Linear Regression (LR) [49] regressor, which was used as a benchmark for disaggregation.

Applied Methods for SRQ4 - Evaluation of data needs and performance: To answer all the above-posed questions, data is necessary. Through literature research, certain typical parameters and constraints were defined for the data. These data differ in origin as well as in properties: they can stem from the medium voltage substation level as well as from the meters at the connection points of devices on the low voltage level. The former are of higher dimensionality and frequency, for the latter only voltage data might be available and only at a much lower time resolution. The explorations, assessments, and validations were conducted with data reflecting properties that are typical for measurements available in the distribution grid at the moment. The substation transformer level data was collected and used at a sampling rate of 4 Hertz and spanning

PV Control Side Load Control Side

Figure 1.6: Laboratory setups used for data collection [50]

various readings such as currents, voltages as well as reactive and active power flows. The data collected at the device level are sampled at a 15-minute rate as they are supposed to be close to SM data. Also, only a few channels are recorded, as voltage data alone is realistically available and also sufficient to build a monitoring solution. Data were collected in both a laboratory environment as well as through the usage of simulation. The measurements taken in the laboratory were recorded phase-wise, which is done to ensure the measurements are as close to a real-life setup where single-phase or three-phase installation of devices might have an impact. Additionally, all misconfigurations were enacted in different laboratory grid topologies and the measurements were recreated by simulation in order to be able to validate the simulations in general. Figure 1.6 shows parts of the laboratory and the setup of load banks and the PV controls which were used to create misconfigurations.

The assessment and validation of the detection and classification methods elaborated earlier were done on various use cases. One of these use cases was the already mentioned PV inverter use case on which both the transformer level as well as the device level approach was tested. A power factor control curve was the grid supporting functionality to be monitored. The reactive power infeed of the inverter is regulated depending on the active power. The local voltage is controlled through this functionality. Figure 1.7 depicts the control curve as well as its modeled misconfigurations. Two misconfiguration examples, the dotted lines, as well as the correct control curve, which is the fully drawn line. The one misconfiguration is an inversion of the curve whereas the other is simply a flat curve, called 'wrong' here. The latter does not provide reactive power at all. The former has the inversed effect on the voltage as the original control by feeding in reactive power of the opposite sign.

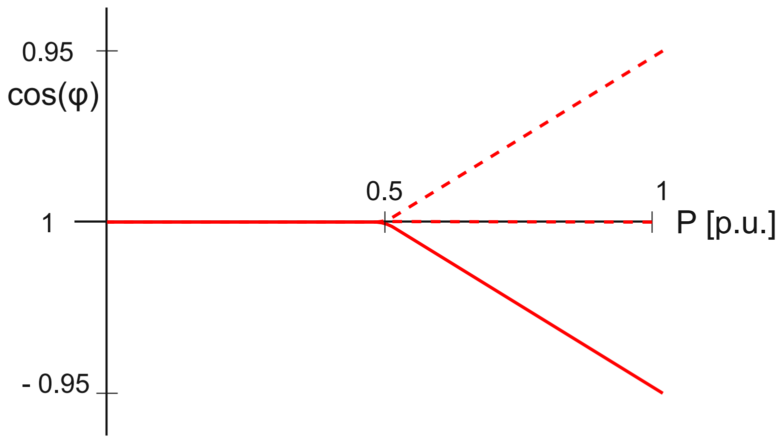


Figure 1.7: Inverter configurations to be detected [50]

Data were collected both in cases in which misconfigurations like the misconfigured curve were presented, as well as in regular operation. Entire days of data were collected in this manner. Furthermore, data from these cases were synthesized. this allowed for the generation of entire years worth of data. This allowed for the variation of the length of the timeseries used for training and evaluation as well as taking into account a wide range of grid setups. Grid models were mainly used in the form of synthetic grids provided by open-source research projects such as the SIMBENCH project [51]. These reflect common properties of electrical grids found in Germany for various kinds of grid types and settings such as urban or rural and with or without PV proliferation.

In order to determine which of the methods is the best performing for certain tasks and under certain circumstances, result metrics had to be chosen. As already elaborated, the individual methods were tested separately, assuming perfect performance of the rest of the methods, as well as in combination to assess the overall performance of the monitoring solution found. An important result metric for the disaggregation of the substation profiles into their contributors is the Mean Squared Error (MSE) [52] to assess if the estimation of loads is accurate:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{1.1}$$

whereas n is the number of samples, Y_i is the target value for sample i and $\hat{Y}_i$ is the estimate value for sample i. For the detection and classification tasks, metrics like Precision [53] which measures how many of the found anomalies or misconfigurations of a certain class are actually misconfigurations, for which the number of true positives is the most important count:

$$Precision = \frac{tp}{tp + fp} \tag{1.2}$$

tp being the number of true positives, while *fp* is the number of false positives. Another important metric is Recall [53], which makes a statement about how many anomalies or misconfigurations of a certain category that are present were actually found, also taking into account the false negatives:

$$Recall = \frac{tp}{tp + fn} \tag{1.3}$$

tp being the number of true positives, while *fn* is the number of false negatives, as in misconfigurations that went unnoticed. The last two metrics of Precision and Recall can be condensed into the very expressive F-score [54] metric:

$$F - score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{1.4}$$

where obviously Precision and Recall are being balanced off. therefore, the F-score reflects quite well the practical usability of the method and solution for a user. For once, the solution is obviously meant to detect as many misconfigurations as possible. However, no false alarms are to be set off, as this can either entail costly and resource-intense interventions or at least diminish the monitoring solution's reliability. A high F-score, as a result, implies that the found solution is applicable.

1.4 Outline of Publications

1.4.1 Applying Deep Learning-based Concepts for the Detection of Device Misconfigurations in Power Systems

The first paper (cf. Chapter 2) describes the development and assessment of approaches to detection and classification at the low voltage device level. Data was generated using simulations for which synthetic grids were used [51]. In this manner datasets of up to 100k samples, whereas each sample is a timeseries of either a day or a week worth of voltage data. This data is collected at 15 minute intervals at the connection point of households with attached PV generation or EVSE. These PVinverters or EVSEs are then either working correctly or experience a misconfiguration, altering their respective control curves which impairs their grid supporting behavior. Data is collected in both cases. The framework and implementations to set up simulations and model misconfigurations to assemble such datasets are described in detail. Furthermore, various methods for detection and classification of these misconfigurations among the samples are presented, all of them DL methods as these are suited to extract fundamental features in this data and therefore best suited as the initial assessments and benchmarking with traditional ML methods showed. Also, hyperparameter tuning was conducted in order to determine which method is best suited. All of this is implemented in the framework, and elaborations on how to conduct such assessments are integrated in the paper. This exploration of DL methods for the application on the device level showed that the R Transformer

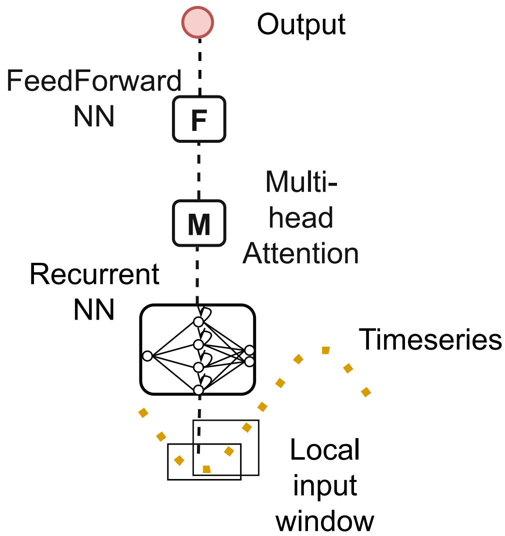


Figure 1.8: R Transformer architecture [50]

architecture showed the best performance. The R Transformer, sketched in Figure 1.8, captures time dependencies in the time-series data locally using a RNN. Furthermore, attention mechanisms are employed to capture these dependencies on a global scale over the entire time-series. So-called keys are used by the attention mechanism encoding the source data feature, this can be the features of some part of the entire time-series considered. Moreover, queries are employed, which for instance hold the hidden states related to the last output. A score function which defines the relationship between keys and queries and is used to decide the next output. This score function is called the energy score. This score marks the impact of the queries in the output, and therefore also the impact of the matched inputs which are encoded in the key.

The R Transformer, as the only architecture or method in general under scrutiny, was able to detect and classify misconfigurations also when fed with data collected in multiple grids. This means the approach is applicable grid unspecifically. Other, traditional, ML or DL approaches did not manage to fulfill this requirement. Therefore, the R Transformer architecture found was selected as the best-fitting method for the device-level monitoring.

1.4.2 Data Driven Transformer Level Misconfiguration Detection in Power Distribution Grids

This second paper (cf. Chapter 3) focuses on the methods and data used for the development and assessment of detection and classification approaches on the medium voltage substation transformer level, but also on the general assessment of data properties. For this reason, data were collected in a laboratory environment and later recreated using simulation. The data was analyzed and the simulations were verified in this manner. The data collected comprised of 15 highly dimensional time-series for each operational scenario such as regular operation or some type of misconfiguration present. The data comprised, amongst others, voltage, current, and power flow data. Each time-series is the equivalent of one day of grid operation data collected both at the substation transformer as well as at the connection points of households with and without attached PV generation. The data were collected at a sampling rate of 4 Hertz and in two different gird setups. The position of the PV generation was altered between the two in order to assess the impacts of renewable generation on the data as well as on the detection performance. Figure 1.9 shows laboratory data on the left side and its recreation using simulations on the right. Further clustering analyses were conducted to show that the simulation data shows the same fundamental properties regarding the similarity between time-series samples.

These timeseries data were then condensed into single sample using PCA, each sample representing one day of data. The literature review and an assessment of requirements suggested using traditional machine learning algorithms for detection and classification in this case. The kNN, DT, and SVM classifier were compared, also under variation of hyperparameters such as the number of neighbors or the kernels. As a result, the SVM was identified as the best-suited method for the transformer-level monitoring. These results were consistent for both the data collected in the laboratory as well as the simulation data, giving further validity to the use of simulation data when developing or operating monitoring solutions.

1.4.3 Data Driven Misconfiguration Detection in Power Systems with Transformer Profile Disaggregation

This paper (cf. Chapter 4) sketches a complete monitoring solution by linking the medium voltage substation transformer level data to the low voltage device level data through disaggregation. This disaggregation helps gain information on loads on the low voltage level. In addition to the PVinverter use case already analyzed in 4 a DSM use case is added. In the DSM use case, a load has either functioning DSM control which shifts its demand to match the generation of an attached PV better, or simply no control. Also for this use case, data was collected in a laboratory environment and recreated through simulation. Two new grid setups were used, also here the PV position was varied, and therefore the load which is DSM controlled. Using data of both use cases the disaggregation method was developed: an ANN approach was compared to LR, finding that they mostly perform equally well. However, the LR had a slight edge over the ANN in grids with few loads and many extensive lines. As necessary inputs for the simulation, the substation data along with voltage measurements at crucial points at the low voltage grid along with an estimate of PV production in the grid were determined. Given these, the disaggregation showed a very good performance in estimating the loads in the grid. Having this disaggregation at ones disposal, the overall monitoring functionality could be completed: a certain period of time's data are used for calibrating the monitoring solution. In the case evaluated this were 14 days' worth of measurements. These measurements are assumed to be taken while all grid connected devices' grid supporting functionalities work as intended. The measurements are, again,

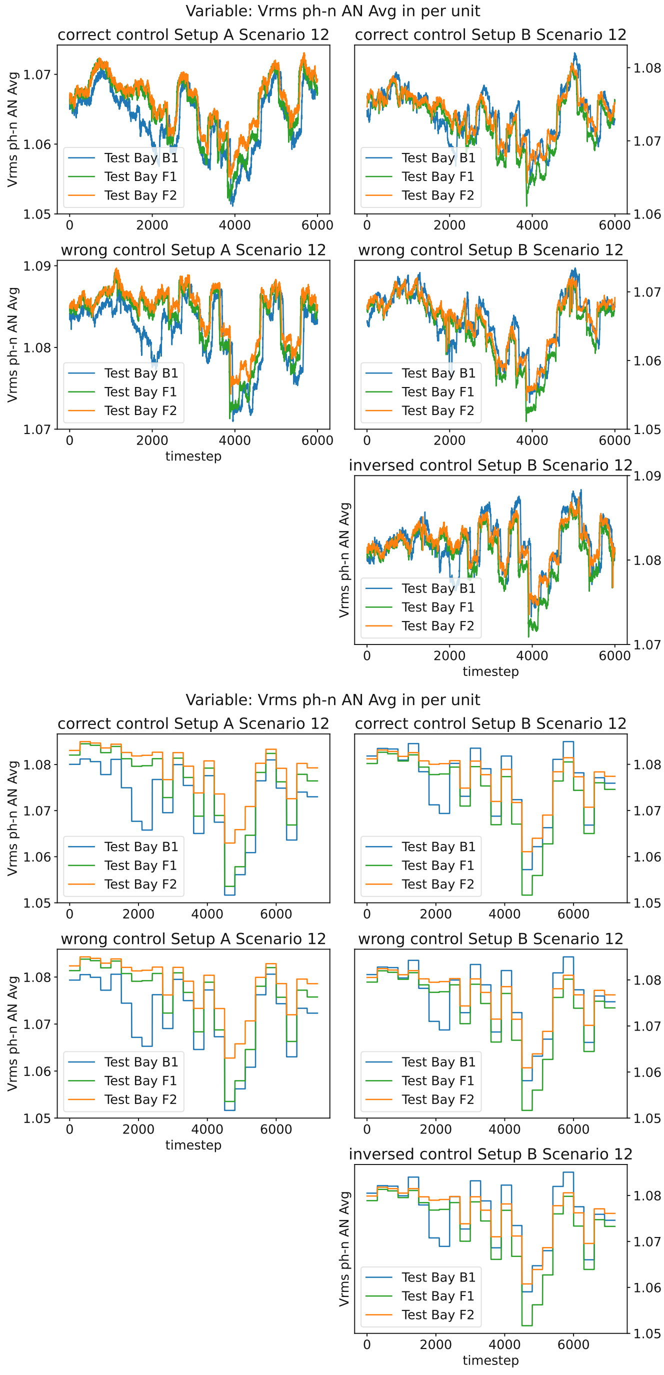


Figure 1.9: Laboratory (top) and simulation data (bottom) by measurement point (note that the measurements in Setup A with an inversed control curve are not available due to lab access time limitations).

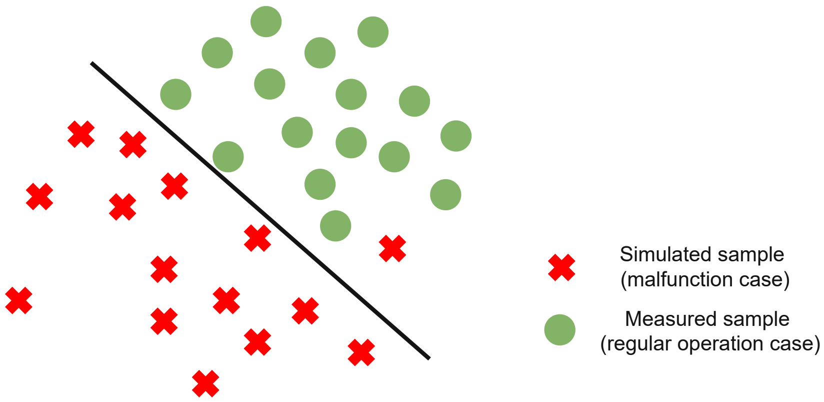


Figure 1.10: Sketch of SVM classifier [50]

condensed into a single sample for each day using PCA. For each of these measured, correct, samples, an 'incorrect' sample is simulated. Simulations are employed here as there is no other way of knowing how a misconfiguration of a certain type would have impacted the grid data on this particular day. The simulation uses the load values gained through the disaggregation and incorporates a certain misconfiguration in order to yield this sample of transformer data of a malfunctioning case. Also here a SVM was found to be the best-performing method. Figure 1.10 illustrates the method in general and in particular the working principle of the SVM; the classifier is built by finding a decision boundary that separates the classes and maximized the margins as in the distance to the samples. Here a simple example of a binary classification is shown, also multi-class detection and classification for more than one misconfiguration is possible.

The simulations, the necessary data mining through disaggregation, the dataset creation, and updates on the SVM model are done once a day and can be done centrally. This allows for a reliable monitoring solution which regularly scans for misconfiguration without manual intervention while still keeping computational costs low.

1.4.4 The DeMaDs Open Source Modeling Framework for Power System Malfunction Detection

The last paper (cf. Chapter 5) presents the software framework designed to enable filling the identified gap in monitoring capabilities for smart grids. The framework provides modeling and data generation, processing, and analysis capabilities to develop and assess functional monitoring solutions. It aims to detect malfunctions during the regular operation of grid connected devices and can be used in arbitrary grid setups, and malfunctions can be modeled freely with a variety of detection methods employed. The framework is entirely written in Python and allows for the integration of different data sources such as collected real-world or laboratory data or data generated using grid simulation tools. The work elaborates in detail on the structure of the framework as well as on the Python libraries used and the resulting dependencies. The framework provides three main classes: Deep Learning and Transformer Detection, which serve as isolated test beds for methods, and Detection Application, which allows for the integration of the individually assessed methods into a practically applicable detection application. The framework's impact is mainly threefold: it allows for the development of detection methods on a device level, at the transformer level, as well as it provides development and testing capabilities for a full detection application. The framework is designed in a flexible manner, allowing users to exchange parts of it and use whichever means of grid simulation or data mining technique they prefer. An example configuration file is presented as well in order to illustrate the usage of the framework. Using this configuration file, an illustratory example of how the framework can be employed is presented. The example provided shows how grid simulation can be used to generate data on electric vehicle charging equipment misconfigurations. The misconfiguration is an inverted charging power control curve dependent on the local voltage. The example also shows how Deep learning can then be employed and tested on this use case as well as which performance metrics such as the F-score are provided. The metrics can on the one hand be used to evaluate the method, and on the other hand to tune hyperparameters in order to enhance the performance of the found solution.

1.5 Contributions and Conclusions

This work contributes to the state of the art in different areas regarding the stated problems and research questions presented in Section 1.2. In the following, the main contributions of the published articles are summarized and the fulfillment of the goals derived from the research question is discussed. Finally, conclusions are drawn and an outlook on future work is given.

1.5.1 Scientific Contributions to the State of the Art

G1: Develop a data driven method to detect abnormal substation transformer readings.

As the monitoring capabilities in distribution grids are limited to readings at medium voltage substations as well as to SM readings, if usable, this goal was tackled on these two levels. Therefore, suited detection approaches were developed and tested for both data sources: traditional ML solutions were employed at the substation using transformer-level data sampled at higher frequencies and at more channels. Furthermore, frequent retraining is necessary as the solution has to be fitted to the individual grids. Due to this, a reasonable computational cost of the method used is beneficial. The best fitting solution was determined, also by tuning its hyperparameters. At the device level, DL methods were used, as they were deemed better fit to the constraints of data sampled at lower frequency and fewer channels, as well as universal applicability regardless of the grid environment. Also here the best performing approach was singled out, as well as hyperparameters tuned and architecture exploration conducted. In this way, a novel architecture using a RNN and attention mechanisms was created. The approaches on both levels yield a daily diagnosis of the regularity of the state of grid operation.

G2: Provide a method to relate abnormal behavior to its underlying cause.

After having detected an abnormal sensor reading at a substation, knowing what caused this reading is very valuable. Again, both data sources, meaning the medium voltage substation measurement data as well the data at the connection point of devices at the low voltage level are utilized in order to gain knowledge about the cause. The same constraints and specifics mentioned above about these data with regard to sampling frequency and dimensionality as well as assumptions about grid-specificity of the solutions apply. Therefore, DL approaches were chosen for the device level classification task here as well. Also traditional ML methods were used for the classification of abnormal behavior at the transformer level. This classification was enabled by modeling and applying misconfigurations of grid connected devices during grid operation and capturing the produced grid operational data. This also allowed to build classifiers for either groups of misconfigurations, as PV inverter misconfigurations, or specific misconfiguration, such as an inverted PV inverter reactive control curve. On both available levels, transformer and device level, a daily statement about the type of the abnormality is made.

G3: Design an approach to mine information about the low voltage distribution grid given centrally aggregated data.

As device level measurement data might not always be available, or only to a certain extent, a data mining solution is necessary to ensure the necessary insights into the operation of the low voltage level grid are gained, given aggregated medium voltage substation data. This solution was found in disaggregation of the medium voltage substation load profiles into its contributing low voltage grid load profiles. This disaggregation was implemented through load estimation performed by either an ANN or LR. The inputs determined for a well performing solution are for once the medium voltage load profiles, but also voltage measurements at neuralgic points in the low voltage grid as well as power generation values. Both methods are able to determine the distribution of loads well when provided with a dataset containing typical grid operational data of a certain grid along with an estimate of, for example, PV power generation. This dataset can be obtained through grid simulations requiring only historic minimum and maximum load values of the low voltage level devices in order to be able to cover the operational states of interest of the grid. In this manner, information can be mined to enable the daily monitoring of the power distribution grid.

G4: Assess the data needed and the performance yielded by a potential monitoring solution.

For all three tasks, detection and classification of abnormal measurements as well as data mining, data was collected and assessed. The data collection was both done in a laboratory environment and through simulation. The gathered laboratory data was recreated using simulation, validating the simulations and its results also for further use in development and assessment of solutions. Based on the laboratory data, but also on data generated using synthetic representative grid models along with grid simulations, the approaches for the respective tasks were developed and the solutions found were assessed. The necessary input data quality and properties were evaluated, as well as the necessary quality of the results of the disaggregation in the form of load estimation used for data mining. These evaluations were done by defining KPIs such as the MSE for the disaggregation performance or the F-score for the tasks of detecting and classifying an abnormality. The solutions found were also benchmarked against simpler approaches in order to justify their applicability. In this manner, the prerequisites for the methods chosen and the necessary quality of the solutions found was determined.

1.5.2 Conclusions

The changes the energy system is undergoing triggered by the need for sustainable energy generation and consumption also raise novel challenges for grid operation. These new generators and loads are often installed decentrally in places where they used to be uncommon. However, the grid is, due to legacy issues, built to only statically distribute energy decentrally. New capabilities are therefore introduced to control these new load and generation devices in order to support the grid operation. For the same historic reasons as already mentioned, the distribution grid lacks monitoring capabilities, making it impossible for the grid operator to ensure these grid supporting functionalities are actually delivered. Therefore, a great need for novel distribution grid monitoring solutions is present and a research gap concerning approaches to them was identified. The developed complete monitoring solution fulfilling all the goals elaborated above is sketched in Figure 1.11.

At first, data at the medium voltage substation are measured along with data at the connection points of grid participants at the low voltage level. The data collection can happen through Smart Meters at the low voltage level and through substation measurement points at the transformer level. The operational data from all these sources are then fed to the monitoring solution, which was developed under the DeMaDs (Data Driven Detection of Malfunctioning Devices in Power Distribution Systems) framework. DeMaDs then employs the aforementioned data driven methods to deliver its monitoring capabilities. The data is generally used to build models of the grid state in regular operation, as well as of states in which grid participants' grid supporting functionalities experience misconfigurations. These models are then used to determine whether the grid is in an irregular operational state, therefore detecting anomalous grid behavior. The misconfiguration specific models are also used to determine the cause and origin of such anomalous grid data. On the medium voltage substation level, this works as follows: transformer-level operational data is used for a calibration period of about two weeks. This data is considered to stem from regular grid operation and serves as daily samples of the same. In order to obtain the respective irregular grid operational data, data mining and simulations are necessary. The aggregated transformer level data is disaggregated into the contributing load profiles at the low voltage level using a load estimation model trained on generic grid operational data stemming from simulations of the specific grid monitored. This load data then allows for the simulation of the same grid operation scenario, just with a misconfiguration present somewhere in the grid. This yields another set of operational data at the transformer which is used as a sample of anomalous grid

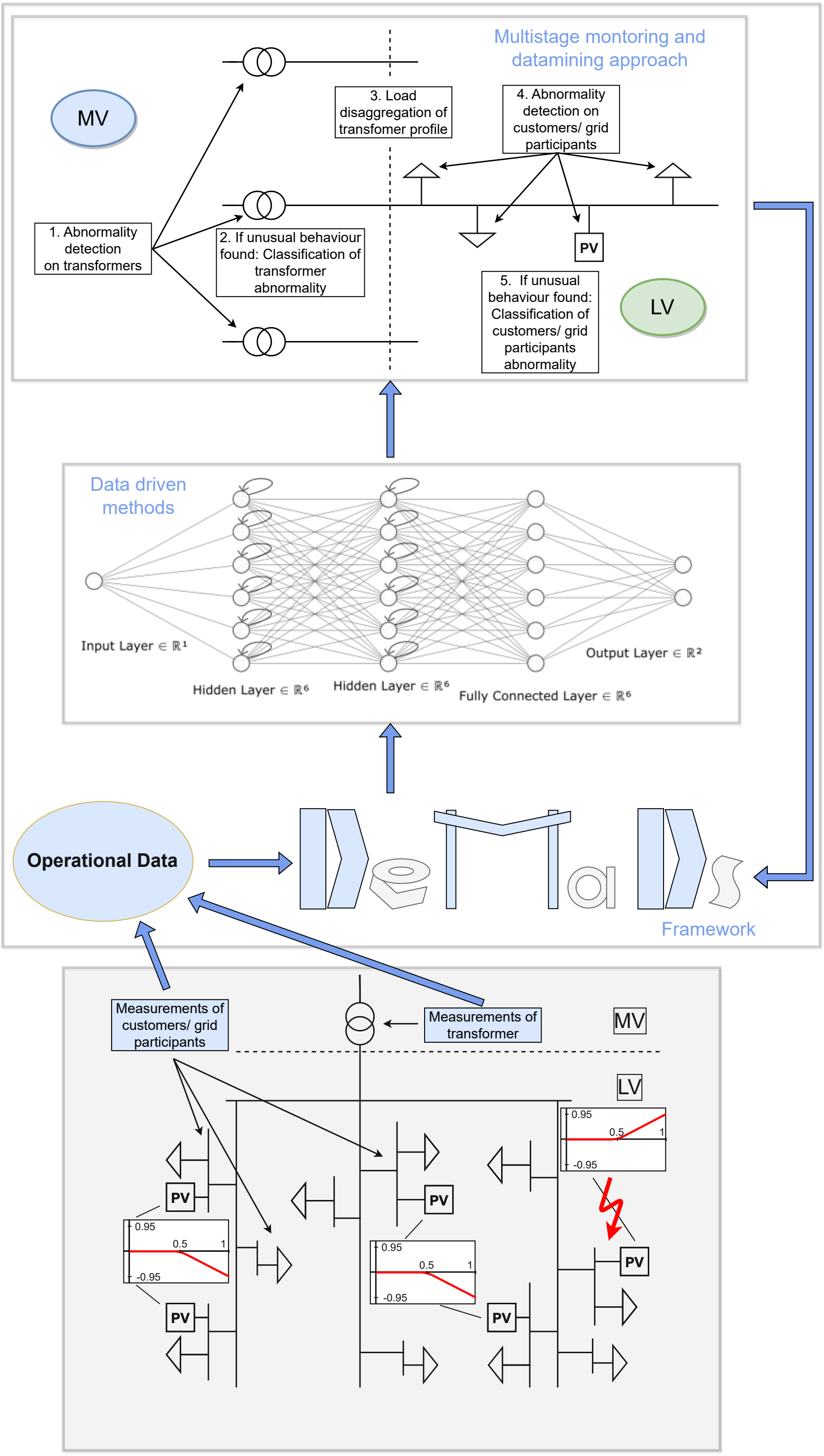


Figure 1.11: Final DeMaDs framework [50]

behavior. The so-built model is used to determine whether a misconfiguration is present, and if yes, of which type it is. The algorithms used here are PCA for data preprocessing, SVM to build the detection and classification models, and ANN for disaggregation. This approach on the substation level is complimented by detection and classification on the device level, which the R Transformer DL method is employed for. Here, the approach is fundamentally the same as grid simulations are used to create grid operational data of regular and irregular grid states on which the DL model is trained. This model is then used to detect anomalous grid behavior and make a statement from which type of misconfiguration it stems. As the models are to be universally usable, data is simulated using various different grid models. The trained models can either be used centrally by feeding SM data to them or rolled out in an edge-computing manner so as to process SM data locally. The former can be difficult at times due to data privacy or communication infrastructure issues, whereas the latter avoids these difficulties, however at the price of greater installation efforts. In the case of local SM data processing, only a flag containing the type and location of misconfiguration detected needs to be sent to the central DeMaDs monitoring application. Generally, due to the detection and classification of anomalous behavior on the local level, a more reliable statement in general but specifically about the location of a misconfiguration is possible.

The chosen approach of a combined grid-specific medium voltage substation transformer-level approach using traditional ML with a universally applicable, local DL method allows for the envisioned and need monitoring capabilities. The DeMaDs framework provides this monitoring without much cumbersome customization of the solution as well as with as little need for training for users, such as a DSO's control room crew, as possible. This is due to the fact, that the transformer level detection only requires a certain calibration period along with historical minimum and maximum load values. The simulations and load estimations that are needed to build the final classifier can also be conducted automatically during that time, after which the solution is ready to use. If changes in the grid occur, a retraining of the model can easily be conducted since the required computational costs for this are modest. Comparably effortless are the installation and operation of the local monitoring approach: as the models used for classification and detection are pre-trained, they solely need to be rolled out, as described, either centrally or locally on edge devices. These models require high computational effort in training, however, only limited computational resources are needed when they are fed with data in order to make statements about a certain grid connected devices state. For both monitoring approaches, the integration of new types of anomalies to be detected only requires the modeling of the respective cause, as in the modeling of a misconfiguration of a grid supporting the functionality of a device in the form of a control curve. Then this new anomaly model is added to the framework and new detection models can be created and rolled out without much development effort, allowing for high adaptability and scalability of the solution presented. Furthermore, no information about specific grid codes or intended configurations is needed as the DeMaDs monitoring framework only uses operational data. The availability of operational data is, therefore, the only real requirement, which is not an issue on the substation level, whereas it can be guaranteed by using SM data only locally on edge devices and transmitting flags. The respective sensors are already widely spread today. This makes major upgrades on the communication infrastructure or changes to the metering capabilities unnecessary for a practical and useful application of the monitoring framework.

1.5.3 Outlook

The work presented offers the basic layout of a monitoring solution, along with tests and validation on laboratory data of the same. A variety of misconfigurations as sources of anomalous grid behavior are covered. These include misconfigurations of PVinverter reactive power control curves, load DSM controls, or EVSE charging power controls. However, further use cases are to be included in order to be covered by the monitoring framework. These might encompass the monitoring of heat pumps or battery energy storage systems. Furthermore, periodical revisions of the algorithms and methods employed are to be conducted in order to keep up with recent developments in the field of ML and DL. Also, regulations aimed at data privacy and the resulting availability and usability of data ought to be under constant evaluation as the legal constraints have quite an impact on how the monitoring can be conducted. Concluding, a field test in various grids under real-life conditions is to be conducted. This is to test and further develop the method and especially its robustness under different circumstances and when confronted with different data qualities. Also, the influence of distortions or faulty measurement readings along with fitting countermeasures could be assessed by conducting extensive field testing. Ultimately, the monitoring solution developed is to be rolled out and integrated with DSOs' existing Supervisory Control and Data Acquisition (SCADA) systems as a decision support tool for control room staff. This would effectively facilitate grid operations by allowing DSOs to react to misconfigurations more quickly which otherwise might go unnoticed and harm the safe and reliable operation of the grid. This will be true, especially in the future, where decentralized renewable generation and novel loads will be even more widespread, and manual checks on them will become completely unfeasible for a lack of operational as well as human resources.

1.6 References

[1] D. Nouti, F. Ponci, and A. Monti, "Heterogeneous inertia estimation for power systems with high penetration of converter-interfaced generation," *Energies*, vol. 14, no. 16, 2021.

[2] E. Heylen, G. Deconinck, and D. Van Hertem, "Review and classification of reliability indicators for power systems with a high share of renewable energy sources," *Renewable and Sustainable Energy Reviews*, vol. 97, pp. 554–568, 2018.

[3] M. Pau, M. Mirz, J. Dinkelbach, P. Mckeever, F. Ponci, and A. Monti, "A service oriented architecture for the digitalization and automation of distribution grids," *IEEE Access*, vol. 10, pp. 37 050–37 063, 2022.

[4] C. Joglekar, B. Mortimer, F. Ponci, A. Monti, and R. W. De Doncker, “Sst-based grid reinforcement for electromobility integration in distribution grids,” *Energies*, vol. 15, no. 9, 2022.

[5] E. De Din, M. Josevski, M. Pau, F. Ponci, and A. Monti, “Distributed model predictive voltage control for distribution grid based on relaxation and successive distributed decomposition,” *IEEE Access*, vol. 10, pp. 50 508–50 522, 2022.

[6] F. Geth, G. Deconinck, and S. Claeys, “Optimization of Gang-operated On-load Tap Changers in Multi-conductor Radial Networks: Formulation and Convex Relaxation,” *TechRxiv*, 1 2023.

[7] L. Brouyaux, S. Iacovella, and G. Deconinck, “Practical comparison of aggregate control algorithms for demand response with residential thermostatically controlled loads,” in *2020 6th IEEE International Energy Conference (ENERGYCon)*, 2020, pp. 870–875.

[8] E. Gümrükcü, F. Ponci, A. Monti, G. Guidi, S. D’Arco, and J. A. Suul, “Optimal load management strategy for large electric vehicle charging stations with undersized charger clusters,” *IET Electrical Systems in Transportation*, vol. 12, no. 1, pp. 49–64, 2022.

[9] Y. Yang, Q.-S. Jia, X. Guan, X. Zhang, Z. Qiu, and G. Deconinck, “Decentralized ev-based charging optimization with building integrated wind energy,” *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 3, pp. 1002–1017, 2019.

[10] G. Deconinck and K. Thoelen, “Lessons from 10 years of demand response research: Smart energy for customers?” *IEEE Systems, Man, and Cybernetics Magazine*, vol. 5, no. 3, pp. 21–30, 2019.

[11] E. Gümrükcü, J. R. A. Klemets, J. A. Suul, F. Ponci, and A. Monti, “Decentralized energy management concept for urban charging hubs with multiple v2g aggregators,” *IEEE Transactions on Transportation Electrification*, pp. 1–1, 2022.

[12] E. De Din, M. Pitz, F. Ponci, and A. Monti, “Implementation of the online distributed voltage control based on containers,” in *2022 International Conference on Smart Energy Systems and Technologies (SEST)*, 2022, pp. 1–6.

[13] E. De Din, F. Bigalke, M. Pau, F. Ponci, and A. Monti, “Analysis of a multi-timescale framework for the voltage control of active distribution grids,” *Energies*, vol. 14, no. 7, 2021.

[14] H. Almasalma and G. Deconinck, “Simultaneous provision of voltage and frequency control by pv-battery systems,” *IEEE Access*, vol. 8, pp. 152 820–152 836, 2020.

[15] T. Peirelinck, C. Hermans, F. Spiessens, and G. Deconinck, "Combined peak reduction and self-consumption using proximal policy optimization," *eprint arXiv:2211.14831*, 2022.

[16] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Local control of reactive power by distributed photovoltaic generators," in *2010 First IEEE International Conference on Smart Grid Communications*. IEEE, Oct 2010, pp. 79–84.

[17] E. De Din, M. Pau, F. Ponci, and A. Monti, "A coordinated voltage control for overvoltage mitigation in lv distribution grids," *Energies*, vol. 13, no. 8, 2020.

[18] E-Control, "Technische und organisatorische regeln für betreiber und benutzer von netzen," E-Control, 2019.

[19] M. Pau, J. Dinkelbach, F. Ponci, and A. Monti, "A state estimation algorithm for the monitoring of distribution grids in absence of pseudo-measurements," in *NEIS 2020; Conference on Sustainable Energy Supply and Energy Storage Systems*, 2020, pp. 1–6.

[20] C. G. C. Carducci, M. Pau, F. Ponci, and A. Monti, "Towards the virtualization of measurements: architecture, solutions and challenges," in *2021 IEEE 11th International Workshop on Applied Measurements for Power Systems (AMPS)*, 2021, pp. 1–6.

[21] BMDW, "Datenformat- und Verbrauchsinformationsdarstellungs VO 2012 – DAVID-VO 2012," Bundesministerium für Digitalisierung und Wirtschaftsstandort, Jul. 2012.

[22] M. Paulus and F. Borggrefe, "The potential of demand-side management in energy-intensive industries for electricity markets in germany," *Applied Energy*, vol. 88, no. 2, pp. 432–441, 2011.

[23] H. Hoffmann, "Kernel PCA for novelty detection," *Pattern Recognition*, vol. 40, pp. 863–874, Mar. 2007.

[24] Y. Zhou *et al.*, "Abnormal event detection with high resolution micro-pmu data," in *Power Systems Computation Conference (PSCC)*, Jun. 2016, pp. 1–7.

[25] Y. Zhou *et al.*, "Veto-consensus multiple kernel learning," in *Thirtieth AAAI Conf. on Artificial Intelligence*, Feb. 2016.

[26] A. Beghi *et al.*, "A one-class svm based tool for machine learning novelty detection in hvac chiller systems," *IFAC Proceedings Volumes (IFAC-PapersOnline)*, vol. 19, pp. 1953–1958, Jan. 2014.

[27] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, "Deep learning for anomaly detection: A review," *ACM Comput. Surv.*, vol. 54, no. 2, Mar 2021.

[28] S. Khomfoi and L. M. Tolbert, “Fault diagnostic system for a multilevel inverter using a neural network,” *IEEE Transactions on Power Electronics*, vol. 22, no. 3, pp. 1062–1069, May 2007.

[29] S. Nedelkoski, J. Cardoso, and O. Kao, “Anomaly detection and classification using distributed tracing and deep learning,” in *2019 19th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)*, 2019, pp. 241–250.

[30] T. Wang *et al.*, “Cascaded h-bridge multilevel inverter system fault diagnosis using a pca and multiclass relevance vector machine approach,” *IEEE Transactions on Power Electronics*, vol. 30, no. 12, pp. 7006–7018, Dec. 2015.

[31] M. Saimurugan *et al.*, “Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine,” *Expert Systems with Applications*, vol. 38, no. 4, pp. 3819–3826, 2011.

[32] A. Ruzzelli *et al.*, “Real-time recognition and profiling of appliances through a single electricity sensor,” *IEEE International Conference on Sensing, Communication, and Networking*, pp. 1–9, Jul. 2010.

[33] Y.-H. Lai *et al.*, “Multi-appliance recognition system with hybrid svm/gmm classifier in ubiquitous smart home,” *Information Sciences*, vol. 230, pp. 39–55, May 2013.

[34] D. Fellner, “Data driven detection of malfunctions in power systems,” in *Proceedings 9th DACH+ Conference on Energy Informatics*, 2020.

[35] A. Y. Jaen-Cuellar, D. A. Elvira-Ortiz, R. A. Osornio-Rios, and J. A. Antonino-Daviu, “Advances in fault condition monitoring for solar photovoltaic and wind turbine energy generation: A review,” *Energies*, vol. 15, no. 15, 2022.

[36] P. Tamilselvan and P. Wang, “Failure diagnosis using deep belief learning based health state classification,” *Reliability Engineering & System Safety*, vol. 115, pp. 124–135, 2013.

[37] N. Lu *et al.*, “Smart meter data analysis,” in *IEEE PES Transmission and Distribution Conference 2012*, May 2012, pp. 1–6.

[38] T. Kurita, *Principal Component Analysis (PCA)*. Cham: Springer International Publishing, 2019, pp. 1–4.

[39] S. Zhang, X. Li, M. Zong, X. Zhu, and R. Wang, “Efficient knn classification with different numbers of nearest neighbors,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 5, pp. 1774–1785, 2018.

[40] B. de Ville, “Decision trees,” *WIREs Computational Statistics*, vol. 5, no. 6, pp. 448–455, 2013.

[41] M. A. Chandra and S. S. Bedi, “Survey on SVM and their application in imageclassification,” *International Journal of Information Technology*, vol. 13, no. 5, pp. 1–11, Oct. 2021.

[42] A. Sherstinsky, “Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network,” *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.

[43] Y. Yu, X. Si, C. Hu, and J. Zhang, “A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures,” *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, 07 2019.

[44] M. Xia, H. Shao, X. Ma, and C. W. de Silva, “A stacked gru-rnn-based approach for predicting renewable energy and electricity load for smart grid operation,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 10, pp. 7050–7059, 2021.

[45] K. Han, A. Xiao, E. Wu, J. Guo, C. XU, and Y. Wang, “Transformer in transformer,” in *Advances in Neural Information Processing Systems*, M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, Eds., vol. 34. Curran Associates, Inc., 2021, pp. 15 908–15 919.

[46] S. Dian, X. Zhong, and Y. Zhong, “Faster r-transformer: An efficient method for insulator detection in complex aerial environments,” *Measurement*, vol. 199, p. 111238, 2022.

[47] G.-F. Angelis, C. Timplalexis, S. Krinidis, D. Ioannidis, and D. Tzovaras, “Nilm applications: Literature review of learning approaches, recent developments and challenges,” *Energy and Buildings*, vol. 261, p. 111951, 2022.

[48] J. Lee and J.-I. Ha, “Temperature estimation of pmsm using a difference-estimating feedforward neural network,” *IEEE Access*, vol. 8, pp. 130 855–130 865, 2020.

[49] C. Yu and W. Yao, “Robust linear regression: A review and comparison,” *Communications in Statistics - Simulation and Computation*, vol. 46, no. 8, pp. 6261–6282, 2017.

[50] D. Fellner, T. I. Strasser, and W. Kastner, “An operational data-driven malfunction detection framework for enhanced power distribution system monitoring – the demads approach,” in *CIRED 2023 - The 27th International Conference and Exhibition on Electricity Distribution*. Institution of Engineering and Technology, 2023.

[51] S. Meinecke and et al., “Simbench—a benchmark dataset of electric power systems to compare innovative solutions based on power flow analysis.” *Energies*, vol. 13.12:3290, 2020.

[52] D. Chicco, M. J. Warrens, and G. Jurman, “The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation,” *PeerJ Computer Science*, vol. 7, p. e623, 2021.

[53] T. Kynkäänniemi, T. Karras, S. Laine, J. Lehtinen, and T. Aila, “Improved precision and recall metric for assessing generative models,” *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[54] Q. Song, H. Jiang, and J. Liu, “Feature selection based on FDA and F-score for multi-class classification,” *Expert Systems with Applications*, vol. 81, pp. 22–27, 2017.

CHAPTER 2

Applying Deep Learning-based Concepts for the Detection of Device Misconfigurations in Power Systems

Publication: D. Fellner, T. I. Strasser, and W. Kastner, "Applying deep learning-based concepts for the detection of device misconfigurations in power systems," Sustainable Energy, Grids and Networks, vol. 32, p. 100851, 2022.

Abstract: The electrical energy system is undergoing major changes due to the necessity for more sustainable energy generation and the following increased integration of novel grid-connected devices, such as inverters or electric vehicle supply equipment. To operate reliably in novel circumstances, as created by the decentralization of generation, power systems usually need grid supportive functions provided by these devices. These functions include control mechanisms such as reactive power dispatch used for voltage control or active power reduction depending on the voltage. As the main contribution of this work, an approach for the development of the detection of misconfigured (e.g., wrongly parameterized control curve) grid devices using solely operational data is proposed. By generating and analyzing operational data of power distribution grids, a Deep Learning -based approach is applied to the detection problem given. An end-to-end framework is used to synthesize and process the data as well as to apply machine learning techniques to it. The results offer insights into the applicability and possible ways to improve the proposed solution and how it could be employed by grid operators. The findings show that DL methods, in contrast to traditional machine learning, can be used for the problem at hand and that the framework developed offers the necessary tools to fine-tune and scale the solution for broader usage.

Keywords: Power distribution system; artificial intelligence; deep learning; device malfunctions; operational data; malfunction detection

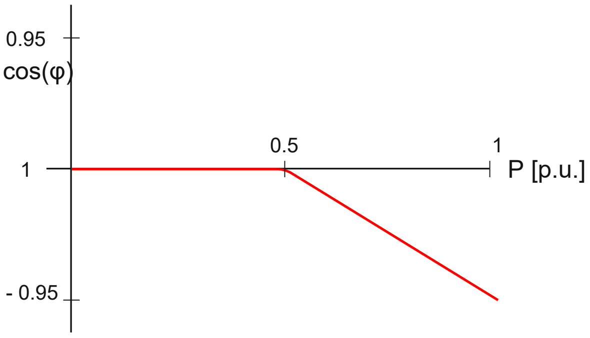
2.1 Introduction

Today, especially power distribution system operators (DSO) have to cope with new challenges arising due to the transformation of the energy system. A major shift in paradigm is the increasing penetration of decentralized power generation [1], which leads to technical challenges in the transmission and storage of power. Standing out is the impact of high photovoltaics (PV) proliferation, but also of other grid-connected devices such as electric vehicle supply equipment (EVSE) [2]. In case of generation outdoing demand locally, bidirectional power flows on different voltage levels as well as voltage rises are the consequences [3]. If the voltage is lifted too much this can lead to voltage band violations, which consist of voltages above or below the admissible limits. Control mechanisms are employed to allow for a reasonable decentralized generation of renewable energy without creating said violations. For this purpose, voltage regulation is the preferred strategy [4], which is made possible by generation units implementing grid supporting functions. These approaches target the frequency as well as the voltage amongst others. Apart from limiting the dispatch of active power, which is a possible solution in the EVSE case [5] of undervoltage, one of the most common ways to influence the voltage is via the power factor and followingly the reactive power exchanged with the network, usually controlled by a local droop control [6].

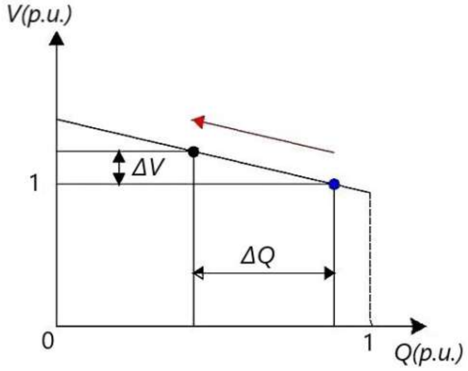
Such controls are configured, as the grid codes demand, in controllable decentralized grid-connected devices. However, they are configured once at installment and subsequently not monitored. As a result shifts in configuration, such as a reset of a control curve, can go unnoticed given the current layout of the grid's metering infrastructure and the DSO' overall metering capacities.

Figure 2.1 illustrates the functions of these reactive power controls; on the left the power factor (cosϕ) is varied depending on the active power (p) dispatch, allowing for reactive power (p) infeed, whereas the right side shows the impact of Q on the voltage (V) [7]. The active power control depending on the voltage applied to EVSEs is described in more detail later in the work.

To ensure that these grid supporting functions are actually delivered, DSO need to monitor the operation of grid-connected devices, for instance, PV inverters or EVSE, as to be sure that the network works in a stable manner. As the available information about grid components' characteristics is often limited, a data-driven approach is a favorable option [9] for a monitoring solution that is actually feasible and therefore useful to DSO. Such a solution can be crafted in a way as to only use operational data of the grid-connected devices, in order to detect misconfigurations of the same. These deviations of configurations from the specifications – as defined by grid codes – can have two reasons; firstly, a different configuration than the normative one can be purposely implemented. Secondly, the configuration can change due to malfunctions or faults. Here



(a)



(b)

Figure 2.1: Control schemes: a) p(p), b) voltage droop [8].

misconfiguration stands for the latter meaning a deviation from previous implementation of a control curve that is assumed to be initially correct. Figure 2.2 depicts how these terms are linked and what is needed to detect anomalies with respect to the type of anomaly. It becomes obvious that for the detection of involuntary misconfigurations only detection of the execution of functionalities is necessary, which does not require knowledge about an implementation code or the fundamental specification and, thus, follows a black-box approach. Therefore, only operational data is used for this purpose.

This detection of misconfigurations while only having operational data, meaning no topology information or information about the configuration other than the previous

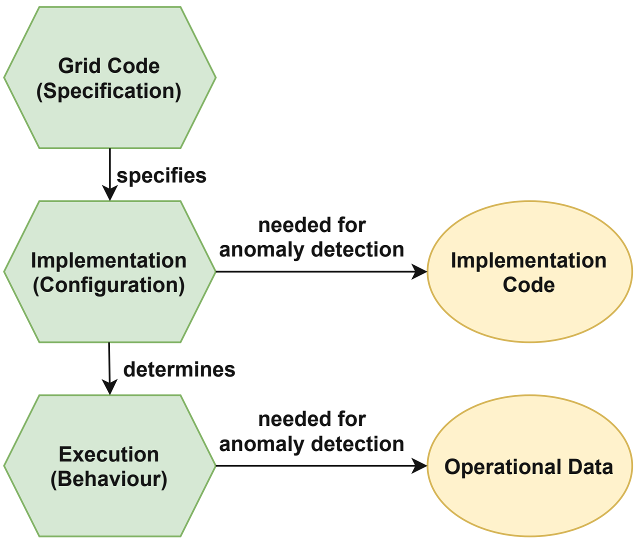


Figure 2.2: Definitions of terms and requirements for the detection of wrong implementations (code needed) respective misconfigurations (data needed).

one ingrained in the data, is becoming more and more relevant as the transformation of the energy system paces on with the installment of PVs and EVSE. At the same time, more and more data at the connection points of these devices, meaning at the Smart Meters, are becoming available and could be used. However, there is no approach to this particular problem and therefore the means of developing and assessing novel ones are needed. This leads to the formulation of the condensed question to be answered:

What approach, applying data-driven methods and algorithms solely on operational data at Smart Meter level is suited best to detect misconfigurations of functions of grid supporting devices in a low voltage distribution grid?

To answer this question a number of objectives have to be fulfilled. These are:

- Obtain data that reflects cases of relevant misconfigurations in operational grid data.
- Assess and process this data to make it usable for the development of detection methods
- Select and apply detection methods to the data
- Pick and refine the best-suited method found

Therefore, the main contribution of this work, which is an invited, revised, and extended version of [10], is the detailed description of an end-to-end framework that can be used to handle grid operational data and to detect misconfigurations. First, this framework is employed to either select or generate, clean and label data for further use. This is necessary since grid data in the form needed is almost impossible to obtain. DSO have no metering in place that would yield data indicating whether a misconfiguration is present or not. The so-created datasets are then preprocessed by, for example, scaling, in order to make it fit for usage by, and training of the detection methods. Subsequently, various detection mechanisms can be applied to the data, which lastly are evaluated and compared against each other. In this work, Deep Learning (DL) approaches are under scrutiny, in addition to being benchmarked against traditional Machine Learning (ML) approaches. DL approaches are chosen for investigation because of voltage curves being highly non-linear and, therefore, features can not be easily derived from them at a low sampling rate as the one of Smart Meter data. However, our previous work indicates a detectable impact of misconfigurations on the voltage [11]. This makes DL an interesting approach [12]. The extension over the conference version [10] consists of an extra use case concerning EVSE misconfigurations, under investigation along with a benchmarking against traditional ML methods as well as sensitivity analysis concerning the parameters of the DL models and their training.

The remaining part of this work can be summarised as follows: In Section 2.1 a detailed discussion of monitoring needs and issues in power distribution grids is conducted. Section 2.2 describes the state-of-the-art related to malfunctions in power systems as well as the usage of artificial intelligence for detecting them. In Section 2.3, the functionality and implementation details of the detection framework are lined out and in Section 2.4 a description and results of the approaches explored using the framework are presented. Finally, Section 2.5 provides the discussion, conclusions, and an outlook about potential further work.

2.2 Related Work

2.2.1 Classical Data Analysis

In the work of [13], electricity consumption data is modeled using a combination of polynomial regression and Gaussian distribution. This is done to detect anomalies in the electricity demand of several schools. This approach could be used for anomaly detection of grid-connected devices, however, the models have to be fitted individually for each device making the application less suitable for broad usage.

In [14], consumption patterns of medium voltage transformers at substations are clustered using algorithms, such as k-means and fuzzy c-means. Abnormal consumption is then identified by employing the Local Outlier Factor (LOF) of hourly load data as a measure. Indicators such as irregular peak unusual consumption, broadest peak demand, sudden large gain, and nearly zero demand unusual consumption are used as features here. Even if not applicable to this particular problem, this shows that there are features present that allow for general detection of anomalous behavior from operational data.

[15] proposes a fault detection in microgrid using traditional machine learning approaches such as Support Vector Machine (SVM), k Nearest Neighbor (kNN), or Decision Trees (DT) in the form of Random Forests. Data of high resolution is used as well as the grid topology known. However, as only Smart Meter measurements are to be used which are only available in a low time resolution, also topology detection is not feasible [16] for the problem at hand. The low resolution and lack of topology knowledge make this approach impracticable here since features could probably not be extracted. Nevertheless, the traditional ML methods of SVM, kNN, and DT are to be tried out and used as a benchmark for other approaches.

2.2.2 Feature Identification using Artificial Intelligence Approaches

Recurrent Deep Learning Architectures

This can be exploited by using DL. As elaborated in [17], Recurrent Neural Networks (RNN) can be used to classify time series data; an Elman network structure is applied to classify a time series. This includes a feed-forward part and a memory part which feeds network activation's from a previous time step as inputs to the network to influence predictions at the current time step. This is achieved through back propagation through time (BPTT); here the gradient of the cost function is propagated with regard to the parameters of the network, like weight matrices, for every time point of the sequence and each layer by unfolding the recurrent connections through time [18]. The parameters are updated using the gradient in a way that minimizes the cost function. The cost function is selected according to the task, such as classification or regression [19]. For classification a cost function as the cross-entropy loss is a common choice since it yields a linear gradient structure, as does the mean squared error used for regression. This is of particular importance to avoid a vanishing gradient while back-propagating it through time [20]. Processing the input as a sequence adds a temporal dimension to the information gained and allows a more flexible window of information to be used in contrast to a feed-forward network. Here, the most frequent classification result yielded by the output neurons is used as a classification result. This might be feasible for grammar checking but might need alteration for the problem addressed in the work here. Especially because RNNs are mostly used for prediction, they have trouble with longer time-series because of a, regardless of the cost function, disappearing gradient and, additionally also due to their limited features w.r.t. parallelization [21].

The RNN approach nevertheless has some deficiencies, most prominently its lacking ability to capture long-term dependencies in sequential data, as lined out in [22]. In the Long Short-Term Memory (LSTM) RNNs recurrent hidden layers, so-called 'memory blocks' are contained; they are made of memory cells that store the network temporal state using self-connections and control the exchange of information through 'gates', which are multiplicative units. Namely, these are the input, output, and forget gates, which, respectively control the inflow or output of activation's to or from the cell or scale its internal states before using them recurrently, which can be interpreted as forgetting. This makes LSTM RNNs an interesting approach when working with longer time series. This is also due to the LSTMs' ability to filter non-relevant inputs through using their gates giving it an advantage when modelling dependencies that vary over time [23],[24].

Another approach to model long-term dependencies better are Gated Recurrent Unit (GRU) RNNs; they address the same vanishing gradient issues as the LSTM approach when back-propagating the gradient of the cost function through time using a simpler structure. Only two gate types are employed by the GRU; an update gate that controls the inflow of information as well as a reset gate that decides over forgetting past information [25]. In contrast to the LSTM architecture, the GRU architecture allows for discarding of past information entirely. Still exploding gradients remain an issue, which is however tackled by gradient clipping. This makes the GRU RNN have fewer parameters in comparison to an LSTM RNN and is, therefore, more lightweight and has been observed to outperform the latter in several tasks. This is also the case for univariate time series classification, which is applicable for classification problems in power systems when, for example, only voltage data is available [26]. Because of these properties, GRU RNNs could also be interesting for a distributed application in a detection mechanism and also for frequent retraining if needed.

Feed-forward Architectures with Attention Mechanisms

An alternative is posed by so-called Transformer architecture [27]. Here, attention mechanisms are used that enable capturing of global dependencies between the input and output, regardless of the positions of the sample points in the time series or sequence. Here, no recurrent computation is used, allowing for better parallelization. Instead 'self-attention' is employed to reach a representation of a sequence through setting the positions of the sequence in relation . An encoder-decoder setup is used, where it performs a mapping of the input to an internal representation, which the decoder then processes to generate the output auto-regressively. The encoder and decoder both consist of feed-forward networks as well as multi-head self-attention mechanisms. This attention mechanism projects a query and key-value tuples on an output which is calculated using the weighted values. These weights are computed in turn on the query and the respective key. This yields an attention value for every query-key-value item and therefore a representation of the sequence. Multi-head attention now enables processing information from a higher dimensional query-key-value set at various positions in contrast to a single attention head, which is helpful. Additionally, positional encodings are simply added to the initial inputs to insert some hint about the positions of the points of the sequence for the feed-forward networks. This non-recurrent approach could be also a computationally interesting option.

Recurrent Architectures with Attention Mechanisms

The R-Transformer concept follows a similar idea as the aforementioned transformer approaches [28]. The main improvement proposed over the regular transformer consists in additional capturing of the sequential information in the data. This is done by positional encoding in the regular transformer, which only yields a scant impact [29], and is often limited to a certain sequence length to be able to set into a context [30]. If positional information is to be retained for a flexible sequence length in an effective manner, high efforts are required to tailor such solutions [31]. These disadvantages take their toll on the robustness of a solution built on traditional attention mechanisms. Furthermore, local structures are neglected because of the sheer number of other positions which allows only for a small signal at a local position, even if these structures might be of quite an importance. To combat these flaws, the R-Transformer uses local RNNs sliding over the sequence, applying windows of a defined length to encode the sequential information in the data and capture local structures in the time series. Thus, latent representations are generated equally for each of the windows treated by the local RNN and are not dependent on any of the other windows. Therefore, information about its local surroundings is ingrained in each data point's representation. Additionally, by sliding the RNN over the time series, the global sequentiality of the data is taken into account as well. The effect of the local RNNs can be compared to a one-dimensional convolution operation, which has the advantage of being parallelizable, but also taking into account sequential information. The gained and encoded local information of one position is then, like in the aforementioned transformer, directly connected to all other positions in the sequence through the multi-head attention mechanism. In a similar application to the one at hand (MNIST dataset with 784x1 sequences), the R-Transformer outperforms both the regular, convolutional Transformer as well as simple recurrent approaches such as LSTM and GRU, whereas an RNN performed significantly worse than all other approaches. This makes the R-Transformer an interesting approach.

2.2.3 Summary and Open Issues

Summarisingly, the work on anomaly detection (see Table 2.1) in the electrical grid domain shows that there are approaches that are not flexibly applicable to new devices or are only applicable at a transformer level or with more information or data of properties which is not available. However, the domain of DL-based approaches offers methods that are, at least in theory, well suited for developing a solution to bridge this gap. Nevertheless, no applications to this specific problem can be found in the literature, and therefore, explorations and assessments of these have to be conducted. This is done by the introduction of a novel framework allowing for generating and/or handling data that is specific to the detection problem at hand. The framework also allows for the development and assessment of detection applications, in order to set up, pick and refine data-driven methods.

Table 2.1: Non-functional requirements (NFR) fulfilled (X) or unfulfilled (–) by approaches in related publications cited.

	Reference							
NFR	[14]	[13]	[15], [16]	[17], [21]	[22] - [24]	[18], [25], [26]	[27], [29] -[31]	[28]
Scalability	–	–	–	–	–	–	X	X
Adaptability	–	–	–	X	–	X	X	X
Integrability	X	X	X	X	X	X	X	X
Usability	X	X	–	X	X	X	X	X
Data Retention	X	–	X	X	X	X	X	X
Robustness	–	X	–	X	X	X	–	X
Quality	X	X	–	–	X	X	X	X

2.3 Scenarios for Monitoring and Detection

2.3.1 Employed Framework

To overcome the shortcomings of present approaches for detecting misconfigurations by the development of a new method, an environment is introduced which is able to handle different detection scenarios, grid setups, and data properties. In general, the approach to detecting devices in misconfigured states is novel in itself. This kind of framework (see Figure 2.3) is used to either synthesize or clean, process, and analyze data as well as apply ML and DL methods to it. Either real-world operational grid data or data of simulations using grids that are specifically designed for simulation purposes – like the ones that form the SIMBENCH[32] project – are being used.

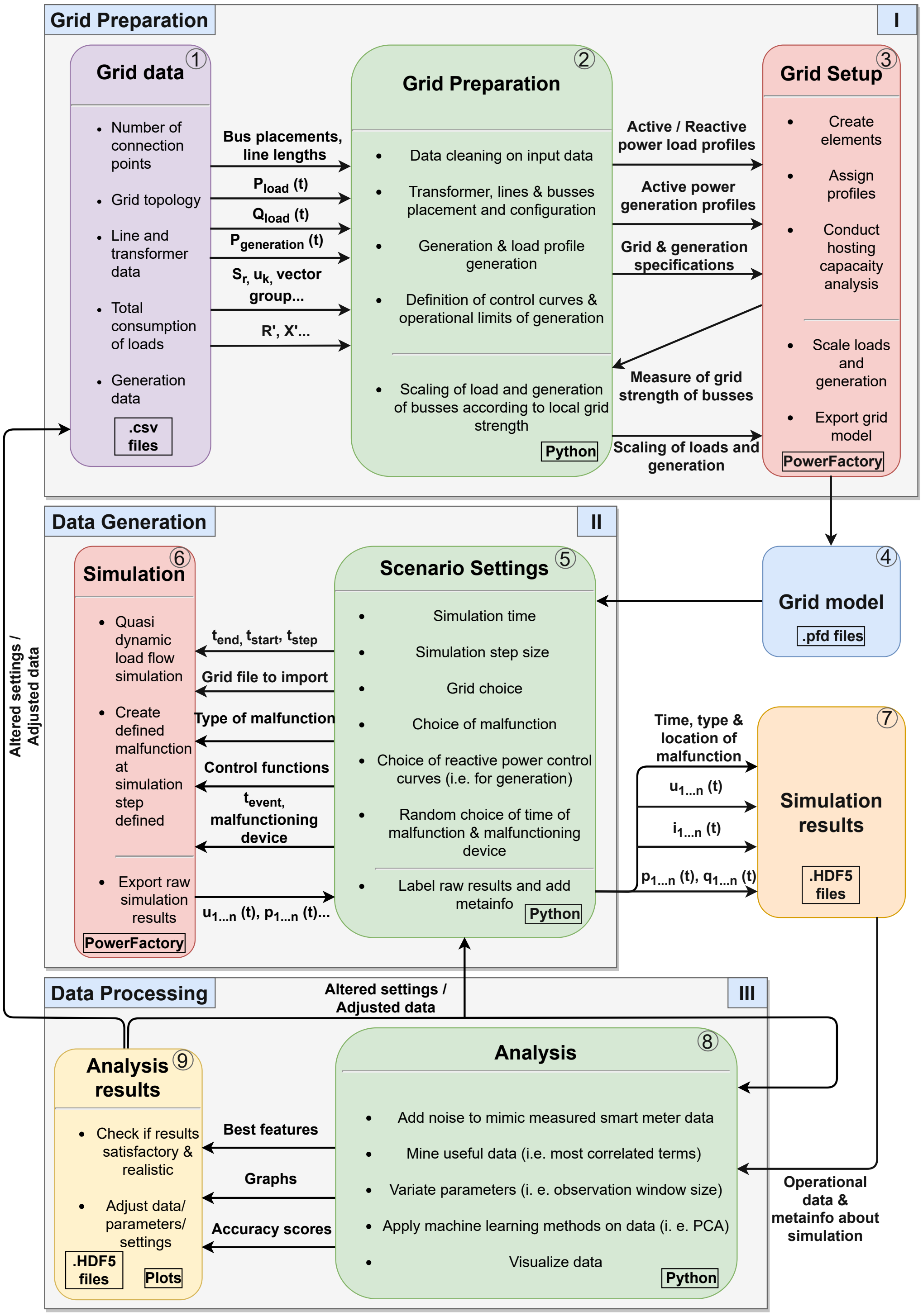


Figure 2.3: Framework used for generation and handling of data of misconfigured devices in power grids as well as for assembling datasets using this data and applying and assessing methods and algorithms for misconfiguration detection.

If operational data is to be synthesized, the grid data used is extracted from the respective files and prepared for further use in simulations, as indicated in Figure 2.3 under ‘Grid Data’. Those are data such as the number of connection points and the specifications of their connections and the substations as well as the consumption and dispatch of loads and generation in the grid. In this manner, the grid topology is checked and generation and load profiles, as well as control curves, are defined and handed over to the grid simulation software. This is done in the ‘Grid Preparation’ box. The next step is ‘Grid Setup’: a grid model is set up in a grid simulation software by placing elements, and adding specifications and profiles to these elements. Using simulations another plausibility and – if necessary – scaling of, for example, loads is conducted and a final grid model is yielded. This model is then used for running simulations in which parameters like the time resolution of the data synthesized, the misconfiguration of interest and its position, as well as the control curve to be monitored, can be varied, as indicated under ‘Scenario Settings’. The simulation then delivers operational data of the grid including data of a malfunction, which is then labeled and saved. This is represented by the ‘Simulation’ box which specifies the simulation method as quasi dynamic load flow simulations, which can also be altered to be a simulation of individual load flows. These individual load flows are necessary to implement voltage-dependent controls, as the one applied to EVSEs, as described later. These voltage-dependent controls run through inner control loops in order to find an adequate setpoint for the operational state. These inner loops slow down the simulation and, therefore, the entire data generation, making it very time-consuming to collect large amounts of data in this manner. To solve this, the framework also allows for the use of so-called ‘quasi dynamic simulation language models’ (‘QDSL models’) in combination with individual load flows. These QDSL models perform the inner loops of device controls, speeding data generation up by a factor of 7. Moreover, the misconfiguration is set up and the raw load flow data of the grid simulation is exported as well as information about when and where a defined misconfiguration occurred. These results are finally used to pick relevant data such as data of connection points having a PV unit or an electric vehicle charging station, add noise to it and, therefore, create datasets. These datasets are used to assess the applicability of machine learning detection methods, especially DL approaches in this case. This is done in the last step two steps, ‘Analysis’ and ‘Analysis Results’, of the framework; training of machine learning classifiers can be conducted as well as architecture exploration or hyperparameter optimization using grid search. The results can be used to make statements about the best-suited methods as well as to gain insights into the quality and property of the data on which the classification has been conducted.

2.3.2 Tackled Scenarios

What this looks like in practice, is illustrated by the schematic of a distribution grid with household loads and PV generation in Figure 2.4; one possible misconfiguration is shown. Here all PV inverters follow a certain control curve regarding to the power factor. As mentioned above, this is meant to help regulate the voltage in case of high active power infeed through the variation of reactive power dispatch. One of the PV units inverts its control curve, it is therefore misconfigured and the voltage is not controlled as intended anymore, which is to be detected. For PV inverters, other possible misconfigurations involve a flat control curve, which equals no control, and different maximal or minimal power factors. This allows an assessment of how grave a misconfiguration has to be found as to be detected by certain approaches.

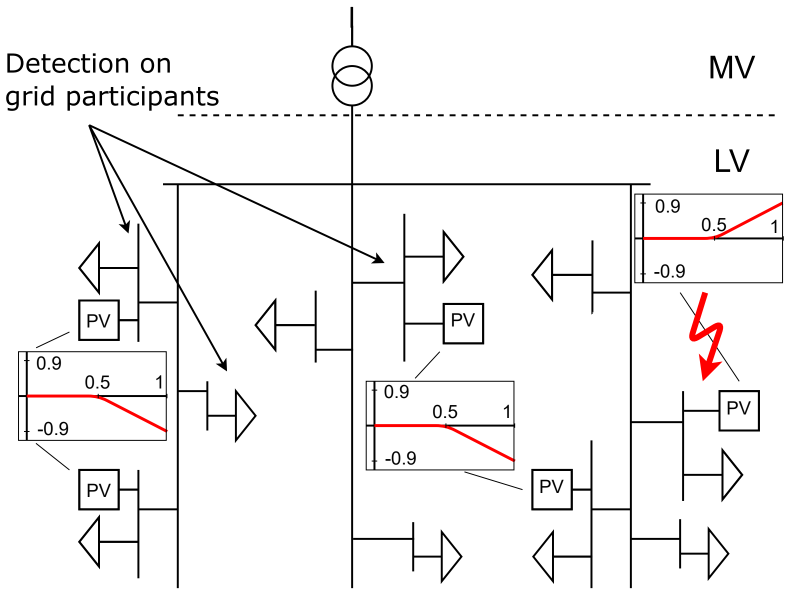


Figure 2.4: Schematic grid used to generate data [7].

Another misconfiguration scenario concerns the EVSEs; the control curve employed in the electric vehicle charging station simulated is a p(U) control, which is an active power control depending on the voltage. The curve used in the simulations is depicted in Figure 2.5; the EVSE is charging at its rated power above a voltage of 1.05 per unit, whereas the charging power is gradually reduced if the voltage of the connection point is lower than this. At a voltage value of 0.95 per unit, this reduction is halted at the minimal charging power of 18.75% of the rated charging power. Therefore, this control should help keep voltages within limits. The misconfiguration is assumed as a flat control curve, meaning no reduction in charging power depending on the voltage level.

These misconfigurations, but also misconfigurations in other devices such as battery energy storages, are supposed to be equally detectable using this approach; being grid supporting, a change in behavior should leave a similar impact on the operational data, such as the voltage. The similarity of features should therefore make a detection possible.

The voltage at the coupling points of the loads and the grid-connected devices, such as EVSEs and PV units is recorded, for example, with a sample rate of 15 minutes to mimic smart meter data. This data is then turned into a dataset by creating samples of a certain sequence length, labeling the same in classes 0 (regular behavior) and 1 (misconfiguration present) as well as choosing the ratio of classes, either balanced or

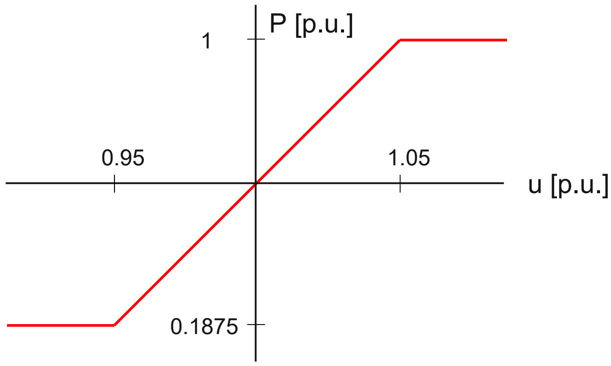


Figure 2.5: p(U) control curve applied to EVSEs.

unbalanced to an arbitrary degree, to fit the capabilities of the methods applied later. Finally, these labeled samples are fed into a data-driven detection method to train on them and assess its performance in detecting a malfunction by recognizing the correct classes. The datasets compiled and used consist of either weekly or daily time-series sampled in 15 min intervals (i.e., common for power system applications), which leaves us with either 96 or 672 data points per sample sequence. This allows for an assessment of the impact of sequence length on the performance of the applied DL methods, which is supposed to stem from their respective handling of long-term dependencies in a sequence.

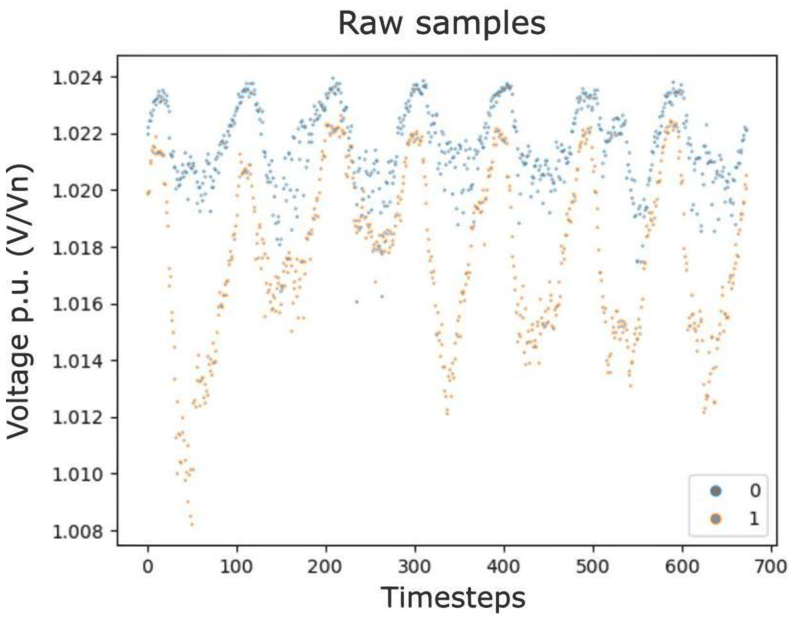


Figure 2.6: Samples of both classes (0 (blue): regular; 1 (orange); misconfiguration present in grid connected PV device) used for Deep Learning.

The novel data used for this work are created using up to 5 of the aforementioned SIMBENCH grids, which are either classified as rural or semi-urban since in such networks voltage issues are prevalent over current issues, making the misconfigurations described relevant in these grids. For the first scenario of a PV misconfiguration, Figure 2.6 shows in two weekly time-series samples the impact left by the misconfiguration on the operational data gained, namely the voltage. The variation in voltage for class 0 ('regular behavior') is much smaller than for the malfunctioning class 1 ('malfunction/misconfiguration present'). This behavior is what is expected here since the control is implemented to keep the voltage within certain admissible limits. Therefore, this different impact of the misconfigured power factor control curve is to be detected. For this case, various datasets with up to 200,000 samples of these kinds with balanced classes, to enable proper learning of features and classification using DL [33], were split into a train and test set and used for the adaption and assessment of the DL detection approaches described in the following. Furthermore, a dataset containing 20,000 samples sourced from a single grid containing PVs and EVSEs was created to assess the applicability of the DL methods in detecting the above-described malfunction of EVSEs.

2.4 Applied Learning Methods and Achieved Results

2.4.1 Data Used

The data shown above has been slightly preprocessed; before its usage in the different DL-based methods by subtracting its mean from every sample to eliminate the influence of a grid feeder-based voltage offset, as well as scaled to a range between -1 and 1. The scaler for this was fit on the training set, scaling all zero-meaned training samples between -1 and 1, and then later applied to the test set. Such samples were assembled to datasets of different sample sizes (1 day and 7 days respective 96 and 672 positions time-series length) and sample numbers (1,000, 5,000, 10,000 for preliminary analysis, and 20,000 respective 200,000 for the method comparisons). Bigger sample sizes imply more data in this case, but longer time-series might not be able to propagate back the gradients through time through the entire time-series.

2.4.2 Method Implementation

For a baseline and benchmark for the DL methods, traditional machine learning algorithms were applied. Namely, these are the Support Vector Machine (SVM, NuSVM), k Nearest Neighbor (kNN) as well as Decision Trees (DT) algorithm. for all of them the implementations found in the Scitkit-learn python library were used[1,2,3,4]. All of these algorithms are supervised learning algorithms, which are applicable to the labeled data at hand. For the SVM and NuSVM, the kernels used to form the decision boundary were varied. For the DT the purity measure was varied, meaning the measure by which data is segmented into classes. For the kNN algorithm, the distance measured to the next neighboring samples was varied to either count all neighbors equally or weighted based on their distance.

As a loss criterion for the DL models, PyTorchs CrossEntropyLoss[5] is applied, which combines the LogSoftmax and negative log likelihood loss (NLLLoss). The input is expected to be the raw, untreated score of each of the two classes, as well as a class label. The CrossEntropyLoss function can be denoted as

$$loss(x, class) = -log(\frac{exp(x[class]))}{\Sigma_j exp(x[j])} = -x[class] + log(\Sigma_j exp(x[j])) \tag{2.1}$$

where *x[class]* denotes the output for the true target class and j spans across all classes, meaning that *x[j]* is the output for the j^{th} class.

Figure 2.7 depicts the most basic structure of the Elman network trained. There, a simple RNN with 2 layers with 6 features in the hidden states each as well as a fully connected layer with 6 neurons and 2 output neurons is presented. The output neurons obviously predict the classes 0 and 1. Each time step is fed into the network, and the output of the final time step, as it is the 'most informed' output, is used for calculating the loss and updating the weights as well as for making a classification. This approach was used during the first assessments of recurrent approaches implemented, as described in the following.

The first goal was to train at least a weak learner, meaning that the output of the classifier should be more accurate than guessing. The initial assessments described in the following were performed with regard to scenarios of misconfigured PVs; in the case of the malfunction detection task presented before, this was achieved at a sample number of 5,000 for the 1 day time-series dataset as well as for the 7 days time-series dataset. This was achieved only using data created using one grid to be able to tell if there was even enough information in the data to make a meaningful classification (i.e., for this task the F-score using the most data reached by the network was slightly over 0.5). Furthermore,

[1]https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html#sklearn.svm.SVC

[2]https://scikit-learn.org/stable/modules/generated/sklearn.svm.NuSVC.html#sklearn.svm.NuSVC

[3]https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html

[4]https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html

[5]https://pytorch.org/docs/stable/generated/torch.nn.CrossEntropyLoss.html

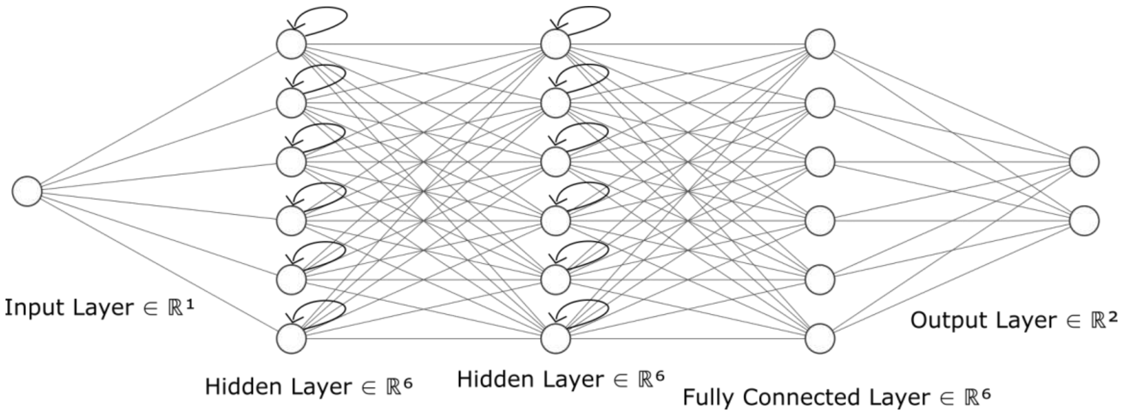


Figure 2.7: Schematic depiction of the RNN trained and used.

a very small learning rate of 10^{-6} had to be chosen to reach sufficiently good results with a standard stochastic gradient (SGD) optimizer. The learning rate was controlled in a manner so as to increase the learning rate by a factor of 1.1 in case the loss between epochs diminishes, and decrease it in turn by a factor of 0.9 at an increasing loss. Training was conducted for up to 100 epochs. A comparison with a linear model showed that the linear classifier did no better than guessing and therefore only reached an F-score of 0.33 on the balanced datasets. The RNN architecture put to trial here consisted of 5 RNN layers each consisting of 20 hidden units and a feed-forward layer with 20 neurons as well. Training here and in the following experiments is always conducted for 20 epochs with a learning rate of 10^{-3} if not stated otherwise. The RNN approach was trained using SGD and Adam optimizer on the 1-day and 7-day samples datasets, with 200,000 samples from 5 grids and 20,000 samples from 1 grid.

As a first alternative to the simple RNN structure, an LSTM RNN was tried out. The architecture used also consisted of 5 LSTM layers and a feed-forward layer with 20 hidden units, respectively neurons per layer, arranged in the same manner as for the simple RNN. An SGD optimizer was used for training.

To be able to compare the 'improved' simple recurrent approaches, for the GRU RNN, the same architecture was chosen as for the LSTM RNN. As optimizers, SGD and Adam were used when training on the same data as above. The transformer as the only non-recurrent detection approach using an attention mechanism was used with an architecture of 5 feed-forward layers with 20 neurons each. The attention mechanism constituted of a multi-head attention with one head at first. Here, an SGD optimizer was used.

Finally, the most sophisticated architecture used is the so-called R-Trans-former, following [28] which incorporates both attention mechanisms as well as recurrent and feed-forward neural networks, as lined out in Figure 2.8. The multi-head attention approach allows to relate a part of a sequence to any other part of the sequence as it treats them all equally but encodes them positionally at the same time. This helps to learn global dependencies while neglecting local structures, which might also be of great interest during the course

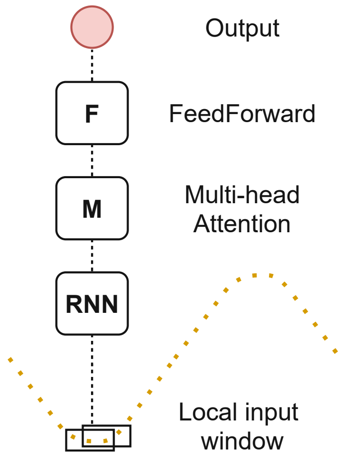


Figure 2.8: Structure of the R-Transformer used.

of a day. Therefore, each part of the sequence is processed beforehand by an RNN; a window of a certain number of points is slid over the sequence capturing local sequential information. In this architecture, this window had a size of 7 data points. Furthermore, the local RNNs were GRU RNNs of which 4 layers with 3 hidden units each were used. This was decided following a singular experiment conducted on the 7-day 200k dataset in which GRU reached an F-score of 0.51 after training for 47 epochs at a learning rate of 10^{-5}, outperforming RNN and LSTM. The multi-head-attention used had one head to be able to assess the impact the recurrence has in comparison to the regular Transformer. In a first approach, only one block of stacking a local RNN, a multi-head-attention network, and a feed forward layer were used. An SGD optimizer was used.

After conducting experiments based on the initial strategy of using only the last, 'most informed' output for backpropagation as well as classification of samples, a 'majority vote' as described in [34] was implemented. This majority vote uses the outputs of a portion of the entire sequence, or of the whole sequence, and calculates a loss depending on them. The absolute loss is then divided by the number of outputs used to have a comparable loss in all cases. This also allows for an evaluation of how many outputs should be used ideally to perform the majority vote. This can be done as a hyperparameter optimization, performed, for example, as a grid search.

2.4.3 Achieved Results and Discussion

The code used to produce the datasets and results can be found in the corresponding GitHub repository[6]. For the comparison of the methods as a main result metric the expressive F-score was used, which combines and balances Precision (i.e., how many of the found misconfigurations are actually ones), and Recall (i.e., how many of the misconfigurations present have been found). This allows a quick understanding of how

[6] https://github.com/DavidFellner/Malfunctions-in-LV-grid-dataset

helpful a result is to a grid operator since a DSO wants to balance between false alarms and finding all occurrences.

To provide a baseline, traditional machine learning algorithms were applied. Table 2.2 gives an overview of the methods applied as well as their parameters and the results yielded. The depicted results apply to the dataset containing data of PV misconfigurations. This assessment was conducted to provide a baseline and serve as a benchmark and additional justification of DL approaches in this case. All experiments were run applying 3-fold cross-validation.

Table 2.2: Overview of the results found when detecting a PV misconfiguration using traditional machine learning methods

Model	Decision Trees		kNN		SVM & NuSVM
Parameter varied	Impurity measure		Distance measure		Kernel
1 day-dataset (sequence length: 96)	Entropy	Gini	Uniform	Distance	Linear, Sigmoid, RBF Polynomial (degree 2-6)
Better than guessing (better than linear model)	No	No	No	No	No

As this assessment makes clear, various common traditional machine learning algorithms fail in delivering meaningful results, even if parameters are varied to optimize their performance.

The aforementioned majority vote classification was assessed using a grid search hyper-parameter optimization. As can be seen in Figure 2.9, the so-called 'calibration rate' was varied for this purpose: this rate determines what portion of the sequence, meaning how many of the first data points of the sample processed, are used for calibration. The outputs of these first data points are not used for the majority vote classification. This means that a calibration rate of 0.8 corresponds to the last 20 percent of the sequence's outputs being used for the classification. A calibration rate of 1 corresponds to using only the last 'most informed' output for classification. On the left side of the figure, we can see the performance of the R-Transformer architecture, whereas the right side depicts the score of the LSTM architecture as described before. The dataset used consists of 20,000 one-day samples, which are samples with 96 data points, sourced from a single grid containing only loads and PVs. Therefore, the misconfiguration under scrutiny here concerns a PV unit's control curve.

The results of the assessment , only using the last, 'most informed' output for classification, for the small dataset sourced from 1 grid as well as the big dataset collected from 5 grids when detecting a PV misconfiguration are summarised in Table 2.3. This is done for a setting with a PV proliferation of 25 percent, meaning every fourth load has a photovoltaic installation. In this context, a Weak Learner is performing better than the linear model which only guesses and therefore reaches an F-score of 0.33. The results achieved here are not good enough for actual usage, however, they provide a good orientation for further refinement of methods.

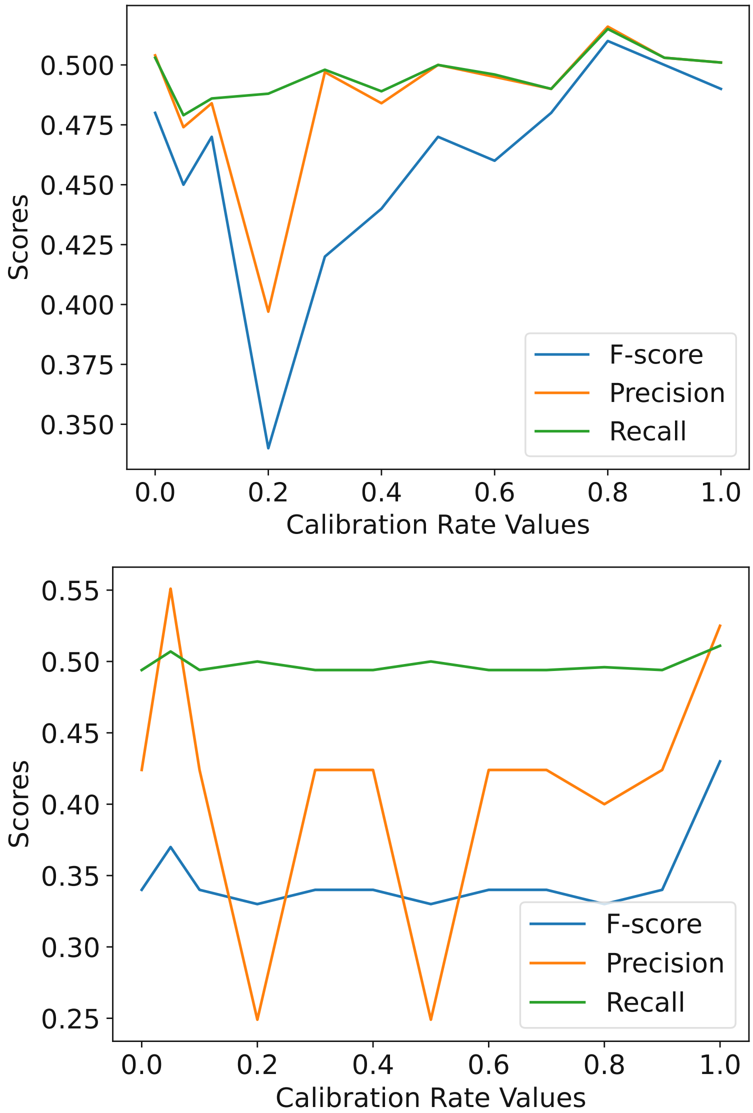


Figure 2.9: Grid search to assess the performance of the majority vote classification; top: RTransformer, bottom: LSTM

For the EVSE misconfiguration a less encompassing assessment was conducted; using data sourced from one grid with a PV and EVSE proliferation of 25 percent each, meaning every fourth load has solar generation and/or an electric vehicle charging station a dataset of 20,000 7-day samples was assembled. This dataset comprises, in contrast to the datasets used thus far, of samples of data of EV charging stations that are either misconfigured or in a regular state. Once again, only the last output is used for classification. The performances of the various methods applied to detect this EVSE misconfiguration are displayed in Table 2.4: once more the results are not satisfactory for a final solution but

Table 2.3: Overview of the results found when detecting a PV misconfiguration using different sequence length, dataset sizes and classifiers: the F-score balances Precision and Recall.

Model	RNN		LSTM RNN		GRU RNN		Transformer		R-Transformer	
Setup #grids & #samples	1 grid 20k	5 grids 200k	1 grid 20k	5 grids 200k	1 grid 20k	5 grids 200k	1 grid 20k	5 grids 200k	1 grid 20k	5 grids 200k
F-score 1 day-dataset (sequence length: 96)	0.33	0.33	0.34	0.33	0.47	0.33	0.33	0.33	0.49	0.47
F-score 7 day-dataset (sequence length: 672)	0.33	0.33	0.37	0.33	0.39	0.33	0.33	0.33	0.52	0.51
Weak Learner (better than linear model)	No	No	Yes	No	Yes	No	No	No	Yes	Yes

provide a guideline for further research on refined methods.

Table 2.4: Overview of the results found when detecting a EVSE misconfiguration using different classifiers on a single dataset: the F-score balances Precision and Recall.

Model	RNN	LSTM RNN	GRU RNN	Transformer	R-Transformer
Setup #grids & #samples	1 grid 20k	1 grid 20k	1 grid 20k	1 grid 20k	1 grid 20k
F-score 7 day-dataset (sequence length: 672)	0.20	0.27	0.47	0.47	0.46
Weak Learner (better than linear model)	No	No	Yes	Yes	Yes

After these assessments, a first phase of hyperparameter optimization was conducted on the dataset containing PV misconfigurations of 1 grid with a sample length of one day. As the R-Transformer architecture was found to be the best fit for this application, it was also the one chosen to be tuned for better performance. Amongst others, a grid search on the number of Attention Heads was conducted. The number of Attention Heads for the Transformer as well as the number of underlying RNN Attention Blocks were varied, either separately or on par with one another. A model dimension of 30 was chosen to accommodate a higher number of Attention Heads or Blocks. As the joined adjustment of the number of blocks showed the best results, Figure 2.10 shows the results of this assessment; the best number of heads was found to be 2 for both the Attention heads as well as the RNN Attention Blocks, yielding an F-Score of 0.53, which is a 4% improvement over the base configuration, which was setting the parameter to 1.

Based on the results of this first round of tuning, another round was conducted on the R-Transformer. This time the parameter Key Size of the underlying RNN blocks of the transformer was found to improve performance at a certain setting. The Key Size determines the length of the sequence that is processed by the underlying RNN. Figure 2.11 depicts the outcome of this exploration; a Key Size of 8, instead of 7, allows for an F-score of 0.60 which is another 7% increase in performance compared to the first

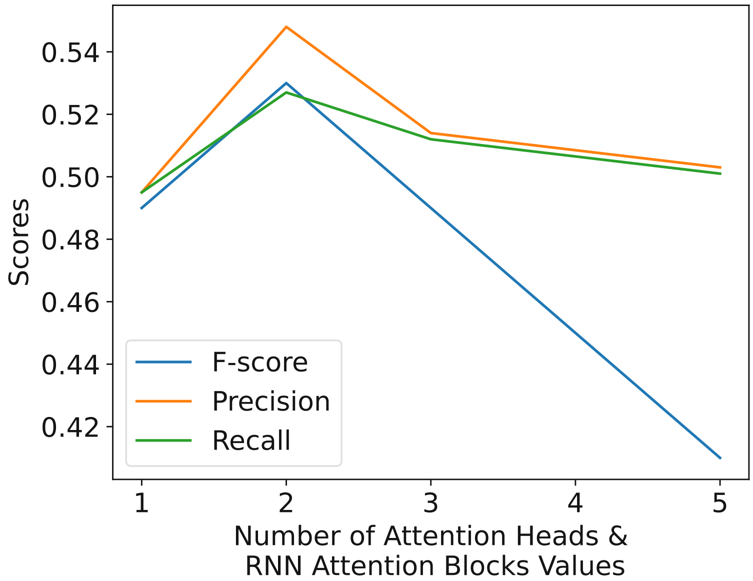


Figure 2.10: Hyperparameter optimization done on the number of Attention Heads of the R-Transformer

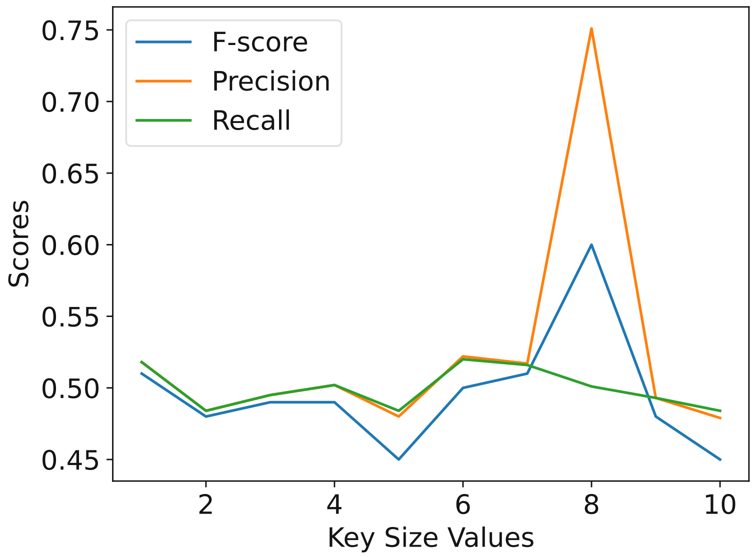


Figure 2.11: Hyperparameter optimization done on the key size of the RNN blocks of the R-Transformer

phase of tuning and a total of 11% enhancement over the base case.

These efforts on tuning show that the performance of the solutions can be augmented by extensive architecture exploration. This is to be done for every architecture for a specific use case, however, there are no additional hurdles except for increased computational demand.

2.5 Conclusions

2.5.1 Achievements

As the necessary integration of decentralized renewable energy generation and other newly introduced grid-connected devices proceeds, grid operators need novel ways to monitor the functionalities of these generation units and devices provide. They are crucial to the safe and reliable operation of power distribution grids. Thus, the framework described in this work allows for the development of such monitoring capabilities by extracting and handling data as well as using them for the development and assessment of machine learning methods for this purpose. By its implementation and usage to generate data the first two goals set initially were fulfilled. Several traditional ML, as well as DL-based approaches, have been described and compared in varying settings. In combination with the sensitivity analysis used to find a best-fitting solution, the two remaining objectives were tackled.

2.5.2 Discussion & Conclusion

The initial assessment of traditional machine learning methods for anomaly detection did not yield results pointing to the applicability of the same. Even after parameters of various methods have been varied conducting a sensitivity analysis in order to provide a meaningful benchmark, not satisfactory results could be achieved. This can be attributed to the low dimensionality of the data. Even the great amount of data does not allow the traditional machine learning algorithm to succeed here. This leaves DL methods to be explored as they pose the most promising option.

The assessment of the majority vote classification versus using only the last 'most informed' output for classification offers various findings. Generally, using only the last output, or only the last portion of the sequence, stably appears to give better results than using a major part of the sequence for the classification of PV misconfigurations. However, this overall advantage is mainly rooted in a higher F-score achieved, meaning that the methods can be fine-tuned by choosing a certain calibration rate to fulfill specific requirements: depending on the priorities of the user, a certain share of the sequence can be used for classification allowing for higher precision, in cases where no false positives are wanted, or also a higher recall, in cases when all occurrences of misconfigurations are to be found.

The results of the assessment of the different methods using only the 'most informed' last output lead to the following conclusions: The RNN approach presented already demonstrates the applicability of DL for this task. This quite simple approach already yielded a weak learner for the 1-grid case, that can be extended to an ensemble method or be replaced by more sophisticated algorithms and network structures. Nevertheless, training had to be conducted with a very low learning rate and for a long time. When only trained for fewer epochs and with a higher learning rate, the RNN can not tackle the problem and does no better than the linear model on the PV misconfiguration tasks. The RNN shows even worse performance for the EVSE case, seeming to misclassify samples. This could be due to the RNN learning wrong features, leading to indicating the improper class and showing that the RNN is not up to this task.

The LSTM and GRU RNN approaches both provide an improvement in the PV case, both yielding a weak learner for the 1-grid case. This shows that training can be done much faster with these approaches than with the simple RNN, probably because of the better back propagation of gradients through time. The GRU RNN performed significantly better than the LSTM RNN especially in the 1-day case, making it the more efficient structure. Therefore, GRU was chosen as the local RNN for the R-Transformer. Both approaches failed to provide a meaningful result on the dataset sourced from multiple grids. When put to the task of classifying the EVSE misconfiguration, the LSTM shows similar behavior as the RNN; it appears to fail to properly extract features and does not yield a weak learner within the given training frame. The GRU performs significantly better here, even yielding some of the best results in this setup. This could be attributed to the GRU's capability to discard past information, which is not of value anymore, more easily in comparison to the LSTM. Moreover, the GRU has fewer parameters than the LSTM. This might leave the GRU less confused after a shorter period of training.

The Transformer as the sole fully non-recurrent method showed that in the setting chosen feed-forward-only architectures do not yield satisfactory results as neither in the 1-grid nor in the 5-grid setup the linear model could be outperformed. At least this holds true for the PV case. In the EVSE case, the Transformer architecture yields, along with the GRU model, the most promising results. This might be due to the less sequential character of the features to be learned in the EVSE case in contrast to the PV case. Therefore, for this case, also non-recurrent approaches seem applicable.

The R-Transformer posed the most complex approach under scrutiny, which also yielded the best results for the 1 grid 20k samples dataset, remarkably showing better performance on the 7-day data for classification of a PV misconfiguration. This marks the impact the attention mechanism has as it improves the handling of longer sequences in comparison to the other recurrent approaches. Comparing the results of the feed-forward Transformer the advantage of using the local GRU RNN becomes obvious as the R-Transformer manages to provide meaningful classification. Especially on the 200k samples dataset from 5 grids the combination of these two features shows its strength as the R-Transformer is the only architecture that manages to gain traction in this setup and yield a weak learner. The performance is slightly higher for the smaller dataset though, probably due to the simple network architecture used and a resulting lack in capacity. A similar phenomenon might become obvious when applying the R-Transformer on the EVSE case; the results are slightly worse than for the regular Transformer as well as for the simpler GRU architecture. This could be attributed to the complexity of the R-Transformer. Since it has many parameters, it might not be able to learn within the training time given. Even if this complexity might not be needed here the R-Transformer yields at least a viable solution to the problem. Moreover, the results of the hyperparameter tuning for one use case showed that the performance of the R-Transformer can be increased significantly in this way. As the basic way of conducting such hyperparameter tuning is the same for all the use cases, it could be extended to all the other use cases. This would have to be part of a study focused on this specific problem, as computational resources are limited in the one at hand. For the practical application of the solution, this should not be a problem, as the architectures that are found to be optimal for a certain use case only have to be trained once. Table 2.5 summarises the approaches investigated as well as their assessment.

Table 2.5: Overview over approaches investigated.

Approach	Task	Comment
Most-informed output	Classification strategy	Best option in general for classification tasks as is yields the best scores overall
Majority vote	Classification strategy	Can offer an alternative classification method for specific goals i.e. avoiding false alarms
RNN	PV misconfiguration	Not able to extract features therefore, not better than guessing
	EVSE misconfiguration	Mislearning features, leading to even more misclassifications than through guessing
LSTM RNN	PV misconfiguration	Only partly able to extract features; slightly better than guessing in simple scenario
	EVSE misconfiguration	Mislearning features, leading to slightly more misclassifications than through guessing
GRU RNN	PV misconfiguration	Able to extract features making it better than guessing in simple scenario
	EVSE misconfiguration	Well able to extract features making it one of the best solutions
Transformer	PV misconfiguration	Not able to extract features therefore, not better than guessing
	EVSE misconfiguration	Well able to extract features making it one of the best solutions
R-Transformer	PV misconfiguration	Well able to extract features making it the best solution in all scenarios
	EVSE misconfiguration	Well able to extract features making it one of the best solutions

The study conducted shows how the framework can be utilized to explore methods, which lead in this case to the finding that the R-Transformer generally outperformed its competitors, which however still provided mostly functional solutions. Moreover, the applicability of solutions might differ between use cases. Additionally, the framework offers easy-to-use functionalities of tuning the architectures to obtain better performances, given the required computational power which was a limiting factor here.

2.5.3 Outlook

The presented work is a foundation for a future decision support tool for power grid operators which helps them to implement central monitoring of low voltage grids using DL detection approaches. Further work includes extensive architecture exploration in order to find the best fitting approach and an optimal model thereof for the tasks at hand. This architecture exploration was only conducted partly here since the computational resources available were limited. For a practical application, this would be no hurdle since the optimal architecture for a certain application only needs to be determined once, and only models with the best-suited parameters need to be trained then. When such models are found, a field trial in real-world grids for validation and further refinement of the method can be conducted. The sole availability of simulated data for this study can be understood as another limitation at this point. Furthermore, the range of use cases is to be expanded by training models on data of malfunctioning devices such as battery energy storage or heat pumps, which could also not be implemented yet due to the limitations of computational power already mentioned before. This would then lead to an implementation in said decision support tool and therefore integration into a grid operators toolbox for further monitoring capabilities.

Acknowledgement

This work received funding from the Austrian Research Promotion Agency (FFG) under the "Research Partnerships – Industrial Ph.D. Program" in DeMaDs (FFG No. 879017). Furthermore, this paper is an extended version of a contribution [10] to the 2021 International Conference on Smart Energy Systems and Technologies (SEST).

2.6 References

[1] E. Brown, J. Cloke, and J. Harrison, "Governance, decentralisation and energy: a critical review of the key issues." Loughborough University, 2015.

[2] C. Dharmakeerthi, N. Mithulananthan, and T. Saha, "Impact of electric vehicle fast charging on power system voltage stability," *International Journal of Electrical Power & Energy Systems*, vol. 57, pp. 241–249, 2014.

[3] J. Von Appen, M. Braun, T. Stetz *et al.*, "Time in the sun: The challenge of high pv penetration in the german electric grid," *IEEE Power and Energy Magazine*, vol. 11, no. 2, pp. 55–64, 2013.

[4] N. Mahmud and A. Zahedi, "Review of control strategies for voltage regulation of the smart distribution network with high penetration of renewable distributed generation," *Renewable and Sustainable Energy Reviews*, vol. 64, pp. 582–595, 2016.

[5] L. Wang, Z. Qin, T. Slangen, P. Bauer, and T. Wijk, "Grid impact of electric vehicle fast charging stations: Trends, standards, issues and mitigation measures - an overview," *IEEE Open Journal of Power Electronics*, vol. PP, pp. 1–1, 01 2021.

[6] P. P. Vergara, T. T. Mai, A. Burstein, and P. H. Nguyen, "Feasibility and performance assessment of commercial pv inverters operating with droop control for providing voltage support services," in *2019 IEEE PES Innov. Smart Grid Techn. Europe (ISGT-Europe)*, 2019, pp. 1–5.

[7] D. Fellner, "Data driven detection of malfunctions in power systems," in *Proceedings 9th DACH+ Conference on Energy Informatics*, 2020.

[8] S. Conti, R. Nicolosi, S. Rizzo, and H. Zeineldin, "Optimal dispatching of distributed generators and storage systems for mv islanded microgrids," *IEEE Trans. Power Delivery*, vol. 27, pp. 1243–1251, 2012.

[9] W. Luan, J. Peng, M. Maras, J. Lo, and B. Harapnuk, "Smart meter data analytics for distribution network connectivity verification," *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 1964–1971, 2015.

[10] D. Fellner, T. I. Strasser, and W. Kastner, "Detection of misconfigurations in power distribution grids using deep learning," in *2021 International Conference on Smart Energy Systems and Technologies (SEST)*, 2021, pp. 1–6.

[11] D. Fellner, H. Brunner, T. Strasser, and W. Kastner, "Towards data-driven malfunctioning detection in public and industrial power grids," in *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2020, pp. 1–4.

[12] N. Mehdiyev, J. Lahann, A. Emrich, D. Enke, P. Fettke, and P. Loos, "Time series classification using deep learning for process planning: A case from the process industry," *Procedia Computer Science*, vol. 114, pp. 242–249, 2017.

[13] W. Cui and H. Wang, "A new anomaly detection system for school electricity consumption data," *Information*, vol. 8, p. 151, Nov. 2017.

[14] D. D. Sharma, S. Singh, J. Lin *et al.*, "Identification and characterization of irregular consumptions of load data," *Journal of Modern Power Systems and Clean Energy*, vol. 5, no. 3, pp. 465–477, 2017.

[15] C. Cepeda, C. Orozco-Henao, W. Percybrooks, J. D. Pulgarín-Rivera, O. D. Montoya, W. Gil-González, and J. C. Vélez, "Intelligent fault detection system for microgrids," *Energies*, vol. 13, no. 5, 2020. [Online]. Available: https://www.mdpi.com/1996-1073/13/5/1223

[16] G. Cavraro, R. Arghandeh, A. von Meier, and K. Poolla, "Data-driven approach for distribution network topology detection," *CoRR*, vol. abs/1504.00724, 2015. [Online]. Available: http://arxiv.org/abs/1504.00724

[17] M. Hüsken and P. Stagge, "Recurrent neural networks for time series classification," *Neurocomputing*, vol. 50, pp. 223–235, 2003.

[18] S. Kanai, Y. Fujiwara, and S. Iwamura, "Preventing gradient explosions in gated recurrent units," in *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017.

[19] J. Chen, L. Cheng, X. Yang, J. Liang, B. Quan, and S. Li, "Joint learning with both classification and regression models for age prediction," *Journal of Physics: Conference Series*, vol. 1168, p. 032016, feb 2019. [Online]. Available: https://doi.org/10.1088/1742-6596/1168/3/032016

[20] R. Shoham and H. Permuter, "Amended cross entropy cost: Framework for explicit diversity encouragement," 2020.

[21] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, "Deep learning for time series classification: a review," *Data Mining and Knowledge Discovery*, vol. 33, no. 4, pp. 917–963, Jul 2019. [Online]. Available: https://doi.org/10.1007/s10618-019-00619-1

[22] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," *INTERSPEECH*, pp. 338–342, Jan. 2014.

[23] B. Lindemann, T. Müller, H. Vietz, N. Jazdi, and M. Weyrich, "A survey on long short-term memory networks for time series prediction," *Procedia CIRP*, vol. 99, pp. 650–655, 2021, 14th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 15-17 July 2020.

[24] Y. Wang, S. Zhu, and C. Li, "Research on multistep time series prediction based on lstm," in *2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE)*, 2019, pp. 1155–1159.

[25] B. C. Mateus, M. Mendes, J. T. Farinha, R. Assis, and A. M. Cardoso, "Comparing lstm and gru models to predict the condition of a pulp paper press," *Energies*, vol. 14, no. 21, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/21/6958

[26] N. Elsayed, A. S. Maida, and M. Bayoumi, "Gated recurrent neural networks empirical utilization for time series classification," in *2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, 2019, pp. 1207–1210.

[27] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *CoRR*, vol. abs/1706.03762, 2017.

[28] Z. Wang, Y. Ma, Z. Liu, and J. Tang, "R-transformer: Recurrent neural network enhanced transformer," *arXiv preprint arXiv:1907.05572*, 2019.

[29] M. Dehghani, S. Gouws, O. Vinyals, J. Uszkoreit, and L. Kaiser, "Universal transformers," *CoRR*, vol. abs/1807.03819, 2018. [Online]. Available: http://arxiv.org/abs/1807.03819

[30] R. Al-Rfou, D. Choe, N. Constant, M. Guo, and L. Jones, "Character-level language modeling with deeper self-attention," *CoRR*, vol. abs/1808.04444, 2018. [Online]. Available: http://arxiv.org/abs/1808.04444

[31] Z. Dai, Z. Yang, Y. Yang, J. G. Carbonell, Q. V. Le, and R. Salakhutdinov, "Transformer-xl: Attentive language models beyond a fixed-length context," *CoRR*, vol. abs/1901.02860, 2019. [Online]. Available: http://arxiv.org/abs/1901.02860

[32] S. Meinecke and et al., "Simbench—a benchmark dataset of electric power systems to compare innovative solutions based on power flow analysis." *Energies*, vol. 13.12:3290, 2020.

[33] B. Krawczyk, "Learning from imbalanced data: open challenges and future directions," *Progress in Artificial Intelligence*, vol. 5, no. 4, 2016.

[34] S. Satapathy, A. Jagadev, and S. Dehuri, "Weighted majority voting based ensemble of classifiers using different machine learning techniques for classification of eeg signal to detect epileptic seizure," *Informatica*, vol. 41, pp. 99–110, 03 2017.

CHAPTER 3

Data Driven Transformer Level Misconfiguration Detection in Power Distribution Grids

Publication: D. Fellner, T. I. Strasser, W. Kastner, B. Feizifar and I. F. Abdulhadi, "Data Driven Transformer Level Misconfiguration Detection in Power Distribution Grids," 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Prague, Czech Republic, 2022, pp. 1840-1847.

Abstract: As more novel devices are integrated into the electricity grid due to the changes taking place in the energy system, ways of detecting deviations from the intended settings are needed. If misconfigurations of, for example, reactive power control curves of inverters go unnoticed, the safe and reliable operation of the power grid can no longer be ensured due to possible voltage violations or overloadings. Therefore, methods of detection of misconfigurations of said inverters using operational data at transformers are presented and compared. These methods include preprocessing by dimensionality reduction as well as detection by supervised learning approaches. The data used is of high reliability as it was collected in a lab setting reenacting typical and relevant grid operation situations. Furthermore, this data was recreated by simulation to validate the simulation data, which could also potentially be used for detection applications on a bigger scale. The results for both data sources were compared and conclusions drawn about applicability and usability for grid operators.

Keywords: Power distribution, detection, device malfunctions, operational data.

3.1 Introduction

As the energy system is undergoing massive and quick changes, especially the electric power grid is experiencing a transformation. This leaves power distribution system operators (DSO) facing novel challenges. A major cause of these challenges in the transmission and storage of power is rooted in the decentralization of power generation [1]. One of the biggest effects is caused by the widespread use of photovoltaics (PV) in a grid. Violations of voltage limits as well as bidirectional power flows or overloading of components can be caused by generation exceeding demand in a grid segment [2]. High power infeed from PV can lead to voltage band violations due to elevated voltages. These are to be avoided through controls to allow for an extensive integration of renewable power generation in a decentralized manner. Therefore, the generation units implement some form of voltage regulation [3] that offers grid supporting functions. As the obvious measure of reducing active power dispatch is to be generally avoided to maximize renewable energy output, the voltage is mostly controlled through the variation of the reactive power generation. This is done through variation of the power factor following a droop control curve locally [4].

The behavior of PV inverters and other grid-connected devices has to be monitored to make sure these grid supporting functionalities are performed correctly. Otherwise, a stable and reliable grid operation can not be guaranteed by the DSO. However, limitations in data availability either set by a lack of sensors [5] or data protection regulations [6] have to be taken into account when developing a solution. Therefore, a data driven approach on transformer level to this is advantageous for DSOs as information about components in the grid is frequently lacking [7]. Misconfigurations of grid connected devices are a mismatch between the configuration implemented and the one laid out in the specifications, which is itself defined by grid codes. This mismatch can have two causes; either a different configuration was implemented on purpose or the configuration can change as a result of, for example, malfunctions. The case under scrutiny here is the latter one, meaning the configuration – the control curve – is expected to be initially the correct one. A more extensive discussion on this was already conducted in [8]. This makes obvious that misconfigurations can be detected by a solution using only the operational data collected in the grid, since a misconfiguration leaves a different impact on this data in comparison to a correct configuration of a grid connected device. For this reason, only operational data is used here.

The main contribution of this work is the detailed description of grid operational data, as well as methods applied on it to detect misconfigurations. Data was collected both in a laboratory environment as well as through simulation, allowing for a validation of the simulation data. Furthermore, data processing methods as well as detection methods were applied on the data so as to assess their performance in the task at hand. Dimensionality reduction methods for processing data as well as supervised learning approaches are employed as their applicability is suggested by previous work [9] as well as literature [10].

This work has the following content: In Section 3.1, a discussion of issues in power distribution grids and monitoring needs is conducted. Section 3.2 describes the state-of-the-art related to malfunctions in power systems as well as the relevant usage of artificial intelligence for detecting them. In Section 3.3, the data collected and the means thereof are laid out and in Section 3.4 a description and results of the approaches applied are presented. Finally, Section 3.5 provides the conclusions and an outlook about potential further work.

3.2 Related Work

In [11], energy consumption characterization of buildings of a university campus is presented with the aim of finding anomalies on building level. Features are extracted as well as data reduction methods applied during the characterization. Following, normal patterns for certain times of the day are identified by estimating the most probable one using globally optimal Evolutionary Trees. These, in addition to being more accurate than standard Decision Trees (DT), offer full interpretability of the results. After anomalies are detected on building level, underlying causes are discovered by an unsupervised approach based on Association Rule Mining (ARM). The data used stem from a medium-low voltage transformer, however, in 15 minutes resolution, reducing the applicability of this approach for the higher resolution data at hand.

The work in [12] presents a Deep Learning based anomaly detection method for finding outliers using a Light Gradient Boosting Machine. This machine has the advantage to be less computationally expensive due to the lower number of parameters. Even though the data used in the work has almost the same resolution as the data available here, the approach does not provide extraordinary good results. Moreover, a very big data set is required and therefore also used. This dataset is compiled using only active and reactive power data. This points towards the usage of feature-based approaches that can also handle higher dimensional multivariate data for the problem present.

An approach focusing on phase measurement unit (PMU) data is sketched in [13]; very good results are achieved using Gaussian Mixture Model (GMM) to estimate the probability density function of regular PMU data streams to define the minimum and maximum thresholds for anomalous data streams. Initially Principal Component Analysis (PCA) is used for feature selection as well as k-Means Clustering for clustering of the PMU streaming data. The anomalies under scrutiny here are only faults such as line-to-line or line-to-ground faults. Moreover, the anomalous data is merely simulated, and it is available in a very high resolution, which is not the case for the problem to be treated here. Nevertheless, the approaches to data treatment are relevant. Also [14] indicates that PCA is of use when treating data; it is used here to find the principle components in voltage sags allowing for clustering of them and then assessing the quality of this clustering. The results show that PCA is capable of extrapolating features from voltage data in addition to revealing that the ward linkage method is the best fit for clustering substation power quality data. Both results can be of help in the task presented here.

Methods for anomaly detection on transformer data that are also multivariate are elaborated in [15]. Support vector machines (SVM) as well as k Nearest Neighbors (kNN) and Decision Trees (DT) appear to deliver promising results here. However, the data set is not as high dimensional as the present one and the application described is cybersecurity. Additionally, an ensemble learner of three models is used which makes the approach complex. This makes the utilized supervised machine learning approaches such as SVM, kNN, and DT of interest, leaving nevertheless to be investigated how they perform in the particular case at hand. Another application of SVM and kNN to PMU data can be found in [16]. Here, both show good results when put to the task of detecting voltage magnitude anomalies in feature extracted data. The data used stem both from synthesizing as well as from real world sources and is therefore noised as it has realistic properties in general. However, the detection is only applied to voltage anomalies such as sags, ramps, and steps.

These anomalies do not necessarily have the same properties as the subtle changes in behavior that are to be detected in this work. One more example from the cybersecurity domain can sooth concerns raised by this; in [17], features are also first selected and then the SVM, kNN, and DT algorithms are applied to find anomalies in substation data. Here, these are constituted by, for example, false data injected. These attacks are recognized with a very high probability, showing that these machine learning algorithms are very well capable of detecting all sorts of anomalies in the present work setting.

Summarizing, the work in the electrical grid domain on anomaly detection (see Table 3.1) shows that there are approaches that are either very well suited to certain time resolutions of data, fit for particular dimensionalities of time series data, or require very big amounts of data and computational resources. What becomes clear is that a pre-processing that allows for feature selection appears to be very helpful. Along with classic machine learning algorithms for anomaly detection this approach yields very good results in various applications. However, no solution to the posed problem can be found in the literature, which is why explorations and assessments of approaches to such a solution have to be conducted.

Table 3.1: Non-functional requirements (NFR) fulfilled (X) or unfulfilled (–) by approaches in related publications cited.

NFR	**Reference**						
	[11]	[12]	[13]	[14]	[15]	[16]	[17]
Scalability	X	X	X	X	X	X	X
Adaptability	–	X	X	X	–	X	X
Integrability	–	–	X	–	–	X	X
Usability	–	–	–	–	–	X	–
Data Retention	X	–	X	X	X	X	X
Robustness	X	X	–	X	X	–	X
Quality	X	–	X	X	X	X	X

3.3 Data Collection & Properties

This section is intended to describe the motivation for collecting data in a laboratory setting as well as through simulation and elaborates the respective aims and functionalities it should help develop. The detailed ways of obtaining the lab and simulation data is elaborated along with their properties. Finally, the results produced by it are depicted and analyzed.

3.3.1 Laboratory Data

Data collection in a laboratory setting complements data collection conducted through simulations in an important way. Laboratory data is as close to real-world data as one can hope for, since real-world field data is practically impossible to obtain during the regular operation of a distribution power grid. This is because the occurrence of misconfigurations is not noted in time by the system operators, and therefore, the data collected can not be labeled. When using this data, one would not know whether it stems from regular or erroneous behavior of a grid connected device.

The data collected here concern the PV inverter reactive control curve addressed in Section 3.1 in the case of intended configuration as well as in two relevant misconfiguration cases. These data are very useful for the development of detection approaches at the transformer level. Only operational data was collected and is used in the following as explained and justified above.

For this purpose, low voltage distribution grids, or representative parts of these, were imitated in a laboratory, where grid participants were parameterized and malfunctions were enacted at a given time, allowing for the creation of a labeled validation dataset. Such a facility was found through the H2020 ERIGrid 2.0 project at the Power Network Demonstration Center (PNDC) at the University of Strathclyde in Glasgow, Scotland. To conduct the experiments and recordings infrastructure like controllable loads, substations, and inverters, lines as well as measurement devices, such as smart meters, were necessary. These were then set up in a typical way for grids to be exhibiting sought-after malfunctions, for example, in a radial topology for rural grids. Loads and generation were parameterized to follow certain consumption or generation profiles, as well as certain control schemes regarding their energy consumption or dispatch behavior. The profiles were created following the profiles used by the SIMBENCH[18] project, which provides grids that are specifically designed for simulation purposes. Profiles of consecutive days were chosen to mimic the data collected during grid operation in the course of about 2 weeks.

The operational data such as voltages and currents were then recorded by the grid participants to mimic smart meter data and their power flows to be able to validate the scenario settings. Additionally, readings were recorded at the substation connecting the grid to the medium voltage level. In this manner, one data point would be gained by a quick measurement at a certain setting of generation and load profiles. Given that at a 15 min resolution there are 96 data points per day, 96 tests would be necessary to collect data for one day. As already pointed out, the generation and load profiles, as well as dispatch and charging control patterns, were to be controlled, whereas operational data was measured. This measurement was made using Fluke measurement devices, which delivered 398 different variables per time step, which is 0.25 seconds. In a first selection step, this was manually reduced to 84 relevant variables for further use. As the setup was as close as possible to a real-world power distribution grid, the experiments yielded as realistic results as possible, which should guarantee the highest robustness for the detection methods and monitoring mechanisms developed using this data.

In the experiments conducted, 15 sets of time series that each match a day from 9 am to 3 pm were collected. This time span was chosen to save on valuable laboratory access time and still have as much data with meaningful PV contribution, since the night hours are not expected to contain much valuable information. 15 scenarios, each one consisting of a set of load and generation patterns, were applied to two grid setups depicted in Figure 3.1; both setups consist of a substation in Dyn configuration with an apparent power of 315 kVA, two individually configurable load banks and a PV inverter, as well as cables of up to 100 meters length each connecting them. Measurements are taken at 3 points; at the substation (corresponding to measurement point F2), as well as at both connection points of the loads (measurement points F1 and B1) and the inverter (situated at, and therefore corresponding either to measurement point F1 or B1, depending on the setup). The positions and connections of the measurement points are indicated in the figure.

For the first setup, Setup A, the reactive power control curve was either parameterized correctly or just set to a flat curve, which is called 'wrong' in the following. A flat control curve setting does not provide reactive power at all. Running the 15 scenarios for both control configurations yielded 30 sets of time series for this setup. For the second setup, Setup B, the control curve was, in addition to the correct and wrong options, inversed, yielding 45 sets of time series. An inversed curve setting provides the same amount of reactive power as the correct one, however, with a wrong sign. In total, 75 sets of time series were obtained.

In Setup A, the inverter is closer to the substation, whereas it is further away from it in Setup B. This is done to be able to later assess the impact of grid strength on the detectability of the misconfiguration in the data. In both setups, the misconfiguration is applied to the inverter, as the different exemplary control curves in Figure 3.1 indicate; one is correct, the other is inversed. Because of laboratory access time limitations, only two control configurations were implemented for Setup A, as Setup B is deemed the more interesting case.

3.3.2 Simulation Data

Data collection through simulations complements data collection conducted in a laboratory setting in an important way. It allows to create more data that can be of guaranteed quality when validated through comparison with laboratory data. As some parameters

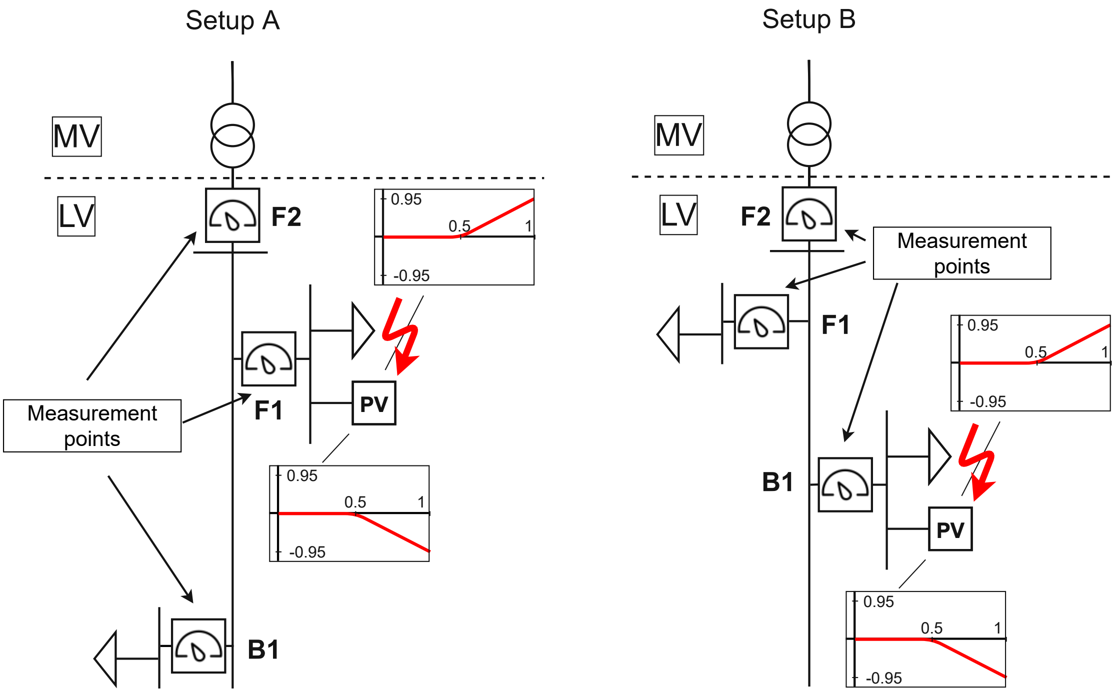


Figure 3.1: Setup A (left) and Setup B (right) with the corresponding names (F2, F1, B1) of the measurement points used in the following.

about the lines of the laboratory were not fully known, assumptions about the line parameters that reflect the most likely properties of the lines were made. Moreover, any modeling of imbalances in the grid was neglected since none were known. These inaccuracies might still have an influence on the simulation quality. The simulations were conducted using the same profiles and setups as in the laboratory setting, recreating the same 75 sets of time series.

In addition to these, simulations with an inversed control curve were also carried out for Setup A, yielding another 15 sets of time series. For the simulation, 30 relevant variables, chosen among the ones available in the software, were selected to be contained in the results. The grid data generation capabilities developed in the course of preceding work [8] were used here.

3.3.3 Outcomes

In summarizing, the experiments were conducted using 15 sets of load and generation profiles in both setups under up to 3 different inverter settings; a regular working control curve, a flat control curve ('wrong control'), and an inversed control curve. An example of the voltage data collected per measurement point in one of these scenarios can be seen in Figure 3.2; as one can see, the voltage is mostly higher in cases where the control curve is wrong or inversed, as is to be expected. For the Setup A, the difference is not as grave since the inverter has, as was expected, as well, a lower impact at a closer position to the substation where the grid is stronger. The difference in the simulation data between the two curves is also smaller than in the lab data. This can only be attributed to the inaccurate modeling of imbalances and possible reactive power consumption that results thereof.

Figure 3.3 shows data, again from the lab and simulation, for the individual cases of control configuration. Again, the impact of the control appears smaller in the simulation data however, it is still noticeable, especially again in Setup B where the PV is farther from the substation and therefore in a weaker point of the grid.

To visualize all scenarios as well as the relationships between each other, clustering was employed, namely, hierarchical ward clustering as described in [19]; first a similarity matrix is computed using the Pearson correlation coefficient. Then a dendrogram is built linking similar time series using the ward linkage method, which is a variance minimization algorithm. The results comparing the data in case of a correct or wrong control curve are shown in Figure 3.4. It becomes obvious that for both lab and simulation data rather the data of the same scenario, in terms of loads and generation, than of the same control setting, such as correct or wrong, are similar. Furthermore, the individual lab data samples seem less similar to each other than the simulation data samples, which are still quite dissimilar. This makes the detection task at hand a nontrivial one.

The simulation model used in a grid simulation software as well as all data and analysis produced using it can be found in the corresponding GitHub repository[1].

3.4 Methods & Results

3.4.1 Preprocessing

To assemble the dataset, all m multivariate time series data samples, each one representing a scenario with a certain control curve configured in one of the grid setups, which have t rows for t timesteps and n columns for n variables as represented by 1) in Figure 3.5, are flattened into single rows of a dataframe having $t * n$ columns. Each column, therefore, represents the value of one variable at a certain timestep of a measurement. The resulting dataset is a m x $t * n$ matrix, each of the m rows representing the data of one of the m measurement samples. Only measurements at the substation level (measurement point F2) are used, as a transformer level detection solution is to be developed.

This data is then scaled to have a mean of zero and a standard deviation of 1 along all $t * n$ features, which again are variables at a certain timestep. This is step 2) in Figure 3.5. The scaled data is then fed into a PCA as described in [14]; PCA is an orthogonal linear transformation, aiming to create a new coordinate system in which the first coordinate, which is the first principal component (1^{st} PC), represents the greatest variance in the data. There can be as many PCs as there are feature vectors in the data however, usually

[1] https://github.com/DavidFellner/Malfunctions-in-LV-grid-dataset

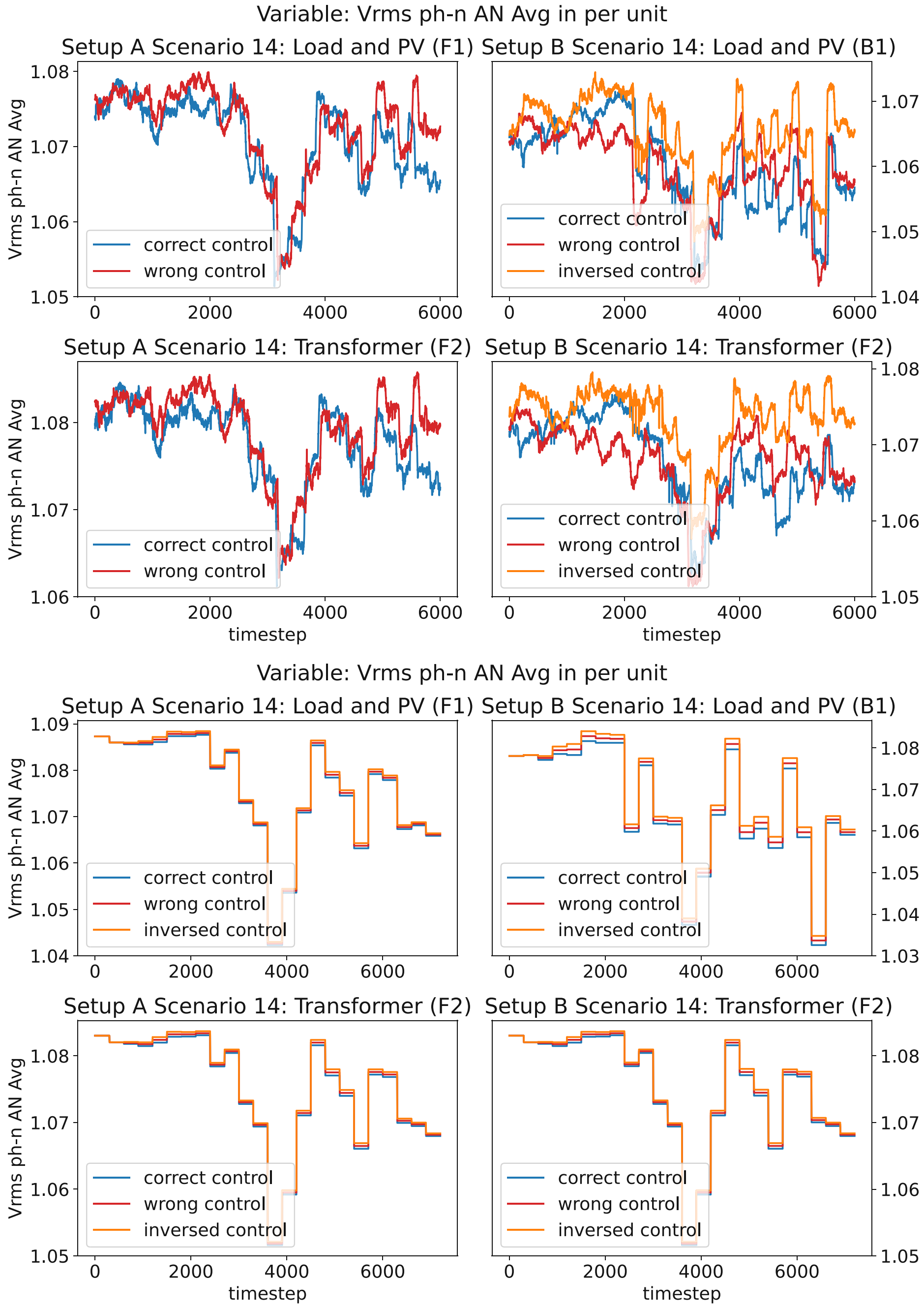


Figure 3.2: Laboratory (top) and simulation data (bottom) by measurement point (note that the measurements in Setup A with an inversed control curve are not available due to lab access time limitations).

fewer PCs than features are retained to achieve a dimensionality reduction and thereby

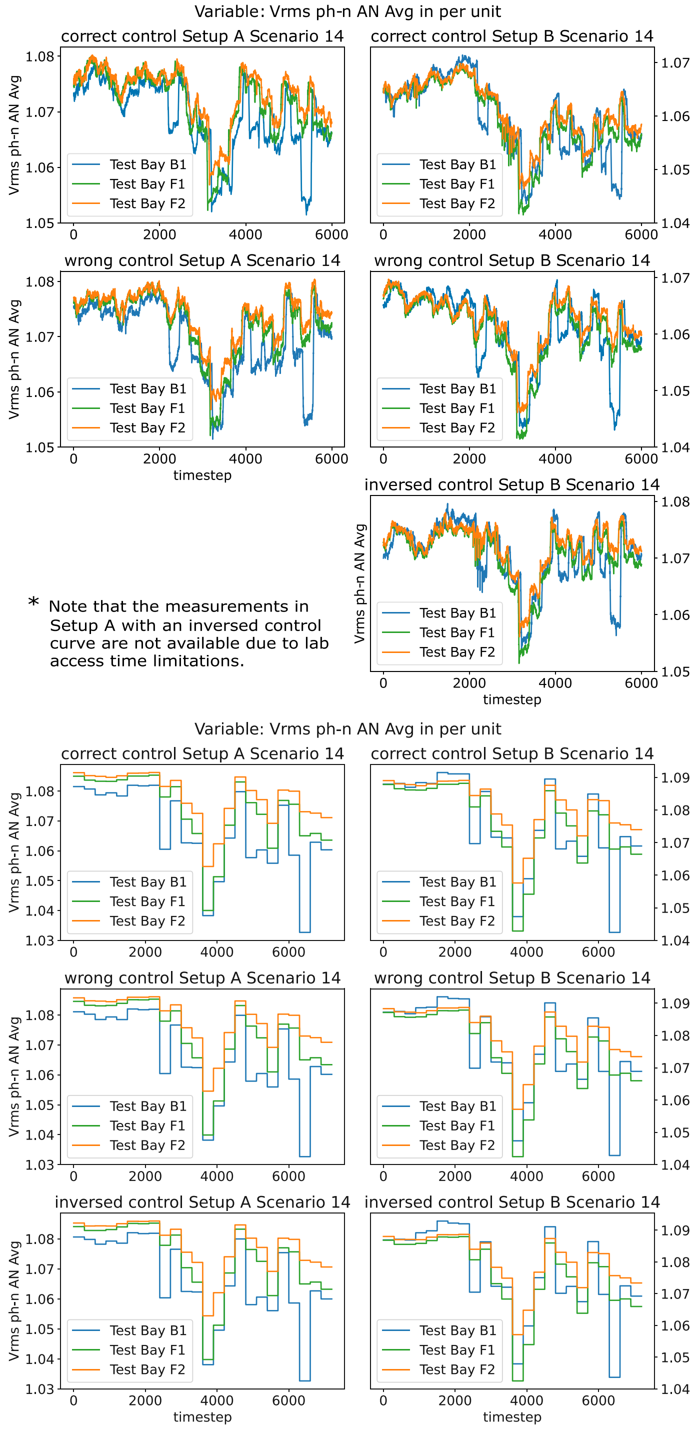


Figure 3.3: Laboratory (top) and simulation data (bottom) by control curve.

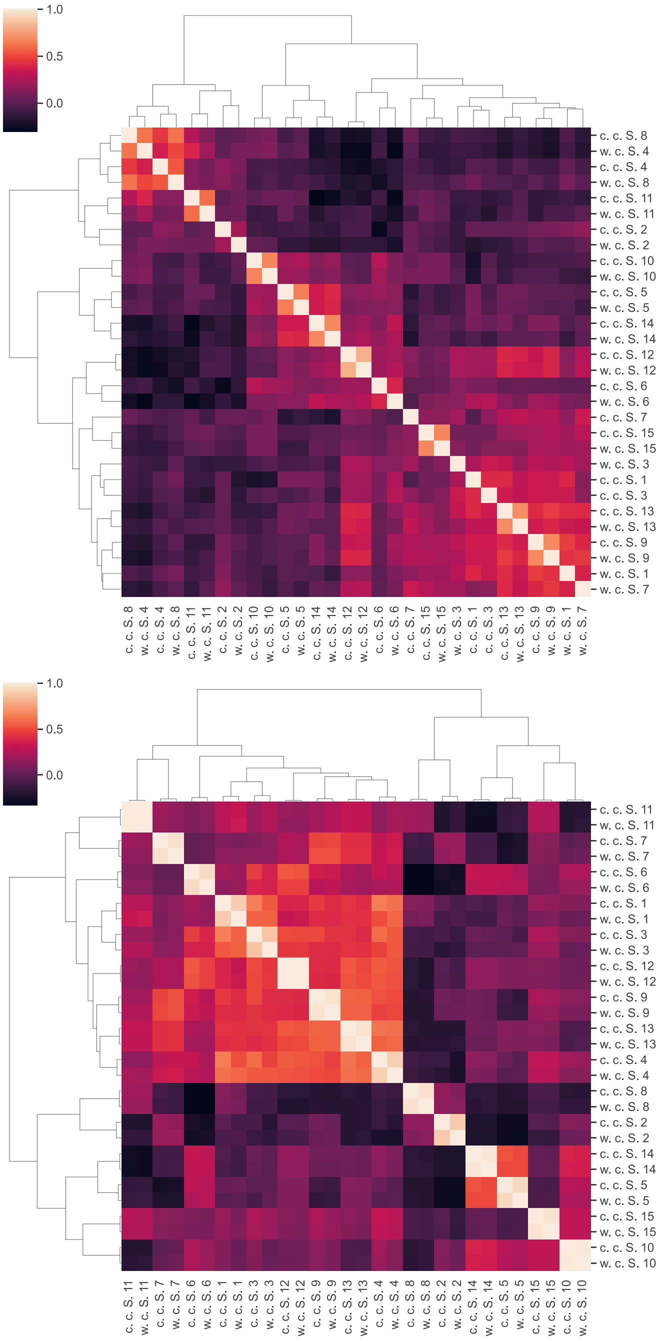


Figure 3.4: Laboratory (top) and simulation data (bottom) of setup B at measurement point F2 clustered; 'c. c. S. 1' and 'w. c. S. 1' stand for 'correct control Scenario 1' or 'wrong control Scenario 1' respectively.

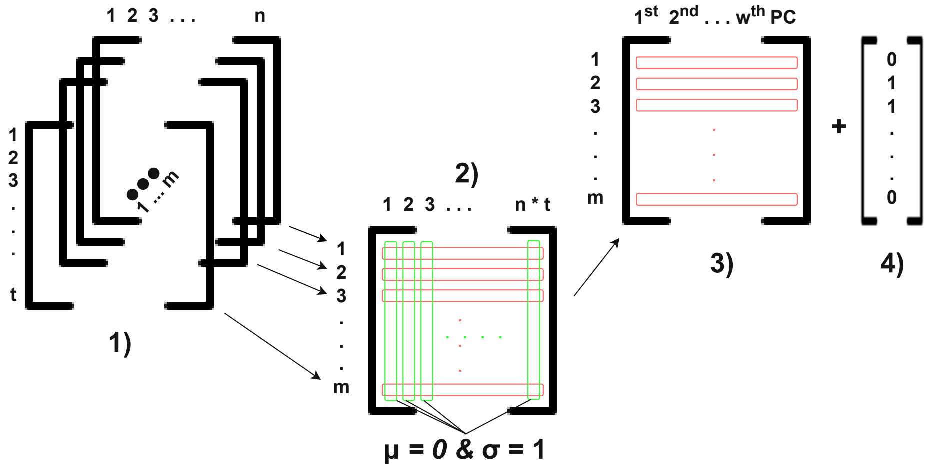


Figure 3.5: Preprocessing and dataset creation.

select important parts of the data, such as the 2^{nd}, 3^{rd}, and so on as higher order PCs retain decreasingly much variance and therefore less information.

As many principal components are kept so as to retain 99% of the variance in the data, which ends up being 17 or 27 components for the simulation, the respective laboratory data of Setup A. Step 3) of Figure 3.5 depicts this unlabeled dataset.

Lastly, the samples are labeled depending on the state of the control curve applied during the measurement yielding the final dataset, as can be seen in step 4) of Figure 3.5.

3.4.2 Detection

Based on the assessment done in Section 3.2, supervised machine learning algorithms are employed for the misconfiguration detection task at the transformer level. Additionally to the mentioned hyperparameter combinations below, additional sensitivity analyses on hyperparameters were conducted. In cases where little or no impact of variating these could be observed, the respective hyperparameters were set to common values as the default ones defined by the specific library implementation used.

As prompted by [16], SVM and kNN are used. The SVM is capable of binary as well as multiclass classification by finding a hyperplane in an arbitrary dimensional space that guarantees as big as possible separation margins between the classes. This makes the SVM especially suitable for high dimensional applications as the one at hand and therefore attractive. Scikit-learn's SVM classifier is used here[2], varying the kernels used (linear, polynomial, radial, sigmoid) and their degrees (1 to 6), as preliminary examinations have shown this parameter to have a significant impact on performance. Kernels define how the separation margin is formed, and therefore, how the decision boundary is adjusted to

[2] https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html#sklearn.svm.SVC

the data. Moreover, another variant, the NuSVM[3], was used. It has the same properties, only that it controls the number of support vectors that are used to find the decision hyperplane to avoid overfitting.

The kNN algorithm[4] uses the Euclidian distance of a data sample to its k-nearest neighbors and decides based on the majority of the neighbor's labels, which class the given sample should be attributed to. This makes kNN a lazy learner, therefore being a very time efficient method. This makes kNN beneficial especially for adding new samples. For this method, the number of neighbors to be taken into account (1 to 4) was varied as well as the weighting of their distances to the data sample. Either the distance of a neighbor would be taken into account, which is called distance weighting, or all neighbors would count equally, called uniform weighting.

Additionally, DTs were applied on the data, as suggested in [11]; a tree is built from the root by recursively partitioning the feature space, until areas, the leaves of the tree, of a certain purity in terms of class labels of the samples in this area are defined. Depending on the splits rule, which in this case the gini impurity, as well as information gain, were used for, the best splits of the feature space are performed. A new sample is then classified following the branches of the tree, which represent the decision rules until it is labeled according to the leaf it ends up with. The DT has a high degree of explainability, which incentivizes its usage in cases where decisions should be justified. Also here the Scikit-learn implementation[5] varying the splits rule was used.

All experiments were done implementing 7 fold cross-validation with balanced classes in all training and test batches. Using the data in this way is intended to reflect the behavior of a detection system using the operational data of the previous days to decide on whether a misconfiguration is present or not looking at the current data.

3.4.3 Results

The code used to produce the datasets and results can also be found in the GitHub repository[1]. The aforementioned datasets were fed to the detection methods, hyperparameters were varied, as well as results cross-validated as mentioned above. The datasets consist of 30 samples for the laboratory data of Setup A labeled as correct or wrong and 45 samples for the lab data of Setup B as well as the simulation data of both setups, which are labeled as correct, wrong, inversed or simply abnormal, meaning of class wrong or inversed.

Table 3.2 summarizes the best results found for a certain dataset using the F-score as a result metric. It represents a balanced combination of Recall, how many of the

[3] https://scikit-learn.org/stable/modules/generated/sklearn.svm.NuSVC.html#sklearn.svm.NuSVC

[4] https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html

[5] https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html

misconfigurations present have been found, and Precision, how many of the found misconfigurations are actually cases of erroneous configuration. As a grid operator using this application would want to balance between finding all occurrences of misconfigurations and false alarms, the F-score is an expressive metric of how useful the approach is to a DSO.

Table 3.2: Comparison of best detection results on laboratory and simulation data.

F-Score Result	**Grid Setup and Data Source**			
	Grid Setup A		***Grid Setup B***	
Case	***Lab Data***	***Sim data***	***Lab Data***	***Sim data***
correct vs. wrong	0.93	0.91	1	0.90
correct vs. inversed	*	0.97	1	0.97
correct vs. wrong vs. inversed	*	0.88	0.96	0.90
correct vs. abnormal	*	0.95	1	0.95

* Not available due to lab access time limitations

The methods and their hyperparameter settings leading to the best results for each case are listed in Table 3.3.

3.5 Conclusions

3.5.1 Achievements

Electricity grid operators need to be able to guarantee safe and reliable grid operation, also in the future of widespread decentralized generation of renewable energy. Therefore, better monitoring of the distribution grid becomes necessary. The data collected and described allows for the development of a validated solution for monitoring the behavior of PV systems in such a grid. Furthermore, the methods applied to this data show the applicability of such a solution.

In general, better performance for Setup B can be observed, which is in line with expectations because the PV is installed at a weaker point of the grid here. Therefore, the impact of the control curve is bigger and a misconfiguration of the same easier to detect. The scores reached on the laboratory data are also higher in all cases than on the scenario

Table 3.3: Comparison of best approaches on laboratory and simulation data.

Best Approach	**Grid Setup and Data Source**			
	Grid Setup A		***Grid Setup B***	
Case	***Lab Data***	***Sim data***	***Lab Data***	***Sim data***
correct vs. wrong	NuSVM: linear kernel	SVM: linear kernel	SVM: linear kernel	SVM: linear kernel
correct vs. inversed	*	NuSVM linear kernel	SVM: linear kernel	SVM: linear kernel
correct vs. wrong vs. inversed	*	SVM: linear kernel	SVM: linear kernel	SVM: linear kernel
correct vs. abnormal	*	SVM: linear kernel	SVM: linear kernel	SVM: linear kernel

* Not available due to lab access time limitations

data. As already discussed above, the simulation data showed only smaller impacts of the different control curves, which also makes detection harder for the simulated data. This also explains only a small difference in performance in the simulation data between Setups A and B. However, this also implies that the results for the simulation data can serve as a lower estimate for the performance on real-world data for Setup A. Nevertheless, the performance is very good, or even perfect, for both setups and all cases. This is likely connected to the rather simple grid topologies and the performance might deteriorate in more complex settings.

In all cases, the method delivering the best results was found to be a form of SVM with a linear kernel, which can be explained by the high suitability of this algorithm for high dimensional data and for datasets with a high feature to sample ratio. This property also allows for the usage of only very recent data, meaning data of the previous days, for detection properties.

The work presented shows that a classical supervised machine learning approach, the SVM, applied to transformer level data can yield very good misconfiguration detection results. As this is the case for both laboratory and simulation data, wide applicability of the method is implied. The even better results on the laboratory data underline the robustness of such a solution.

3.5.2 Outlook

The collection and assessment of the data presented as well as the detection methods explored serve as a building block for the envisioned decision support tool for electric power grid operators, facilitating the monitoring of low voltage distribution grids centrally. In such a solution, as data is collected at the transformer level, it is checked for signs of misconfigurations. After passing this check by the detection methods, simulations of misconfigured cases would be conducted to form the kind of dataset used in this assessment. An incoming abnormal data sample would most likely be recognized by a detection method trained on such a dataset, as the real world samples showed a greater impact on the control curve compared to the simulated samples. The simulations would require the load and generation profiles of grid participants, which could be obtained through disaggregation of the transformer load profile into its components. An approach to this disaggregation is the most important task concerning further work. It could be developed in combination with the load and PV measurements that were at this point only used for validation of the transformer measurements. Other additional tasks are the assessment of additional use cases, such as monitoring of demand side management.

The combination of these methods would then allow for the creation of the already mentioned decision support tool, which would only require a few days of calibration along with regular grid operation before being operational. Such a solution would increase DSOs monitoring capacities in a substantial and feasible manner.

Acknowledgment

This work received funding from the Austrian Research Promotion Agency (FFG) under the "Research Partnerships – Industrial PhD Program" in DeMaDs (FFG No. 879017) and from the European Community's Horizon 2020 Program (H2020/2014-2020) in project "ERIGrid 2.0" (Grant Agreement No. 870620) under the Lab Access user project #115 at the PNDC of the University of Strathclyde.

3.6 References

[1] O. Wagner, M. Venjakob, and J. Schröder, "The growing impact of decentralised actors in power generation: a comparative analysis of the energy transition in germany and japan," *Journal of Sustainable Development of Energy, Water and Environment Systems*, vol. 9, no. 4, pp. 1–22, 2021.

[2] G. C. Mahato and et al., "A review on high pv penetration on smart grid: Challenges and its mitigation using fppt," in *2021 1st International Conference on Power Electronics and Energy (ICPEE)*. IEEE, 2021, pp. 1–6.

[3] A. S. Awad, D. Turcotte, and T. H. El-Fouly, "Impact assessment and mitigation techniques for high penetration levels of renewable energy sources in distribution

networks: voltage-control perspective," *Journal of Modern Power Systems and Clean Energy*, 2021.

[4] T. T. Mai, A. N. M. Haque, P. P. Vergara, P. H. Nguyen, and G. Pemen, "Adaptive coordination of sequential droop control for pv inverters to mitigate voltage rise in pv-rich lv distribution networks," *Electric Power Systems Research*, vol. 192, p. 106931, 2021.

[5] M.-Q. Tran, P. H. Nguyen, O. Mansour, and D. Bijwaard, "Utilizing measurement data from low-voltage grid sensor in state estimation to improve grid monitoring," in *2020 55th International Universities Power Engineering Conference (UPEC)*, 2020, pp. 1–5.

[6] BGBl. II Nr. 313/2012, "Datenformat- und verbrauchsinformationsdarstellungsvo," 2022, https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20007999.

[7] Z. Ma and et al., "The role of data analysis in the development of intelligent energy networks," *IEEE Network*, vol. 31, no. 5, pp. 88–95, 2017.

[8] D. Fellner, T. I. Strasser, and W. Kastner, "Detection of misconfigurations in power distribution grids using deep learning," in *2021 International Conference on Smart Energy Systems and Technologies (SEST)*, 2021, pp. 1–6.

[9] D. Fellner, H. Brunner, T. Strasser, and W. Kastner, "Towards data-driven malfunctioning detection in public and industrial power grids," in *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2020, pp. 1–4.

[10] S. Barja-Martinez and et al., "Artificial intelligence techniques for enabling big data services in distribution networks: A review," *Renewable and Sustainable Energy Reviews*, vol. 150, p. 111459, 2021.

[11] R. Chiosa, M. S. Piscitelli, and A. Capozzoli, "A data analytics-based energy information system (eis) tool to perform meter-level anomaly detection and diagnosis in buildings," *Energies*, vol. 14, no. 1, 2021.

[12] M. Lu, L. Que, X. Jin, J. Liu, and L. Pan, "Time series power anomaly detection based on light gradient boosting machine," in *2021 International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA)*, 2021, pp. 5–8.

[13] A. L. Amutha, R. Annie Uthra, J. Preetha Roselyn, and R. Golda Brunet, "Anomaly detection in multivariate streaming pmu data using density estimation technique in wide area monitoring system," *Expert Systems with Applications*, vol. 175, p. 114865, 2021.

[14] D. Almeida and et al., “A pca-based consistency and sensitivity approach for assessing linkage methods in voltage sag studies,” *IEEE Access*, vol. 9, pp. 84 871–84 885, 2021.

[15] V. K. Singh and M. Govindarasu, “A cyber-physical anomaly detection for wide-area protection using machine learning,” *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 3514–3526, 2021.

[16] J. Fuentes-Velazquez, E. Beltran, E. Barocio, and C. Angeles-Camacho, “A fast automatic detection and classification of voltage magnitude anomalies in distribution network systems using pmu data,” *Measurement*, vol. 192, p. 110816, 2022.

[17] X. Wang and et al., “Feature selection for precise anomaly detection in substation automation systems,” in *2021 13th IEEE PES Asia Pacific Power Energy Engineering Conference (APPEEC)*, 2021, pp. 1–6.

[18] S. Meinecke and et al., “Simbench—a benchmark dataset of electric power systems to compare innovative solutions based on power flow analysis.” *Energies*, vol. 13.12:3290, 2020.

[19] P. Zehetbauer, M. Stifter, and B. V. Rao, “Phase preserving profile generation from measurement data by clustering and performance analysis: a tool for network planning and operation,” *Computer Science - Research and Development*, vol. 33, no. 1-2, pp. 145–155, 2018.

CHAPTER 4

Data-Driven Misconfiguration Detection in Power Systems with Transformer Profile Disaggregation

Publication: D. Fellner, T. I. Strasser and W. Kastner, "Data-Driven Misconfiguration Detection in Power Systems with Transformer Profile Disaggregation," IEEE Access, vol. 11, pp. 80123-80136, 2023.

Abstract: Rapid and necessary changes in the energy sector are leading to the rise of new, decentralized devices for generation and consumption in the electrical distribution grid. Such devices are inverter-connected photovoltaic (PV) generators, heat pumps (HP), or electric vehicle supply equipment (EVSE). These new components make the power grid operation more difficult as they display volatile behavior and therefore also need to provide grid-supporting functionalities. Distribution System Operators (DSOs) need to make sure these grid-supporting functionalities are performed correctly, in order to guarantee a safe and reliable operation of the grid. However, especially the low voltage distribution grid is still ill-equipped with sensors and therefore difficult to monitor. This contribution, therefore, presents a data-driven application for detection of misconfigurations using the data available at metering points of substations and selected voltage measurement points in combination with a transformer load profile disaggregation approach. The assembled application outlined is both functional, scalable, and easy to integrate into current monitoring schemes. Such a monitoring application has not been designed yet and is therefore novel. The data used were collected in a life-like laboratory setup and recreated using simulations in order to be able to test and validate both the detection as well as the disaggregation method. Two monitoring use cases of control functions are considered;

the first one is a reactive power control of PV inverters, and the other one is a Demand Side Management (DSM) control of loads. The results presented offer insights into both the quality and performance of the application assembled. Furthermore, the influences of the individual methods of the approach are explored as well. The conclusions drawn show that a functional monitoring solution of reasonable reliability can be implemented using the methods presented and tested here. The application can serve as a decision support tool for DSOs requiring only minimal adjustments to the sensing infrastructure.

Keywords: Data-driven monitoring, detection, Machine Learning, device malfunctions, transformer profile disaggregation, load estimation, low voltage grids, misconfigurations, operational data, power distribution.

4.1 Introduction

Both ecological and economic pressures force major paradigm shifts onto the electric power system. One of these is the introduction of decentralized renewable energy generation on a grand scale [1]. Another one is the increased electrification of loads, spanning from heating systems to electric vehicles [2]. These are located decentrally as well, making their impacts on the electric power grid just as troublesome: for historic reasons the power grid is designed to transmit electric energy as well as distribute this energy to customers whose consumption is relatively static and easy to anticipate. However, on one hand, the availability of this energy is getting more volatile as it is linked to, for example, solar and wind yields. On the other hand, it is generated decentrally which may lead to production overtaking local demand [3]. This in turn can cause reverse power flows from the low voltage level of the grid to higher voltage levels, which was unconsidered before. It can also lead to local voltage and current problems, as the grid is not laid out to cater to the decentral infeed of this energy. Furthermore, the aforementioned electric loads are being installed in the low-voltage distribution grid. This means for example more electric vehicles are charged in grid locations that may also not be designed for such high additional loads [4]. This can also lead to voltage or current problems.

4.1.1 Problem Statement

To cope with these problems, for example, of over-voltage in the case of distributed generation, as well as under-voltage in the case of additional loads, the devices installed need to provide grid-supporting functionalities. These functionalities include power factor control curves depending on the active power infeed ($cos\phi(P)$) or reactive power control curves depending on the voltage (p(U)) for distributed generation units [5]. An example of the latter is shown in Figure 4.1; the left side of the figure depicts a reactive power infeed control depending on the local voltage. If the voltage is either too high or too low, reactive power of the according sign is fed in. On the right side of the figure, the impact of this reactive power is depicted. Capacitive reactive power, which is dispatched in what is called underexcited operation, helps to dampen overvoltages by lowering the voltage.

The opposite is true for inductive reactive power, dispatched in overexcited operation, which bolsters the voltage and helps control undervoltages by lifting the voltage.

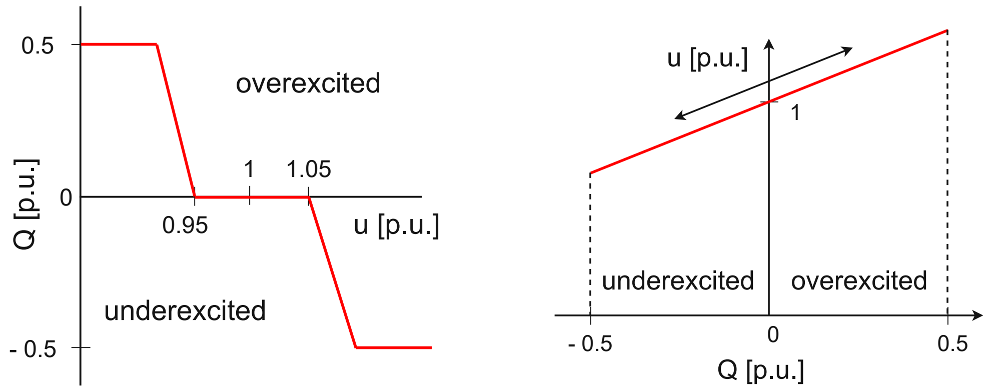


Figure 4.1: p(U) control (left) used for voltage control (right).

Both the $cos\phi(P)$ and the Q(U) control alter the reactive power dispatched by the inverter of a PV generation unit. The reactive power infeed can be used to control, or rather lower in this case, the local voltage and thus support the grid during operation to operate within the acceptable limits [6]. Similarly, loads can be equipped with control functions aiding the grid; EVSEs can follow a charging power control curve limiting the active power drawn in case the voltage drop is too low [7]. Also, household loads can follow patterns in order to shift their consumption to more favorable times of the day as far as the grid or possible self-consumption is concerned. The latter is relevant in case a rooftop PV system is installed that can be used to cover at least parts of the load's consumption. This is commonly referred to as DSM [8].

However, all of these generators and loads are usually installed decentrally in the low voltage distribution grid, which has, as already mentioned, not been designed for such use. This also manifests in the lack of sensor capabilities which are usually limited to substation measurements on the transformer or Smart Meter (SM) measurements [9]. The use of the latter is often restricted by data protection regulations, which leave the state of the distribution grid as a blind spot to the Distribution System Operator (DSO) [10]. Nevertheless, the DSO has to ensure the grid is working safely and is within acceptable limits of loading and voltages. To ensure this, the DSO needs to be able to monitor the correct execution of the discussed grid-supporting functionalities, as configurations might change in an undesired manner due to faults, software updates or user interference [11]. This would then lead to misconfigurations. At the moment, this is only possible through manual check-ups conducted by maintenance crews which are costly and unfeasible. As the rollout of the novel decentral generation and consumption devices proceeds, the need for an automated solution arises. This solution should be both easy to install and robust during use. Furthermore, it should require as little adaption to changes in the grid as possible, which can be used by the sole use of operational data, as discussed in previous work [12]. The need for insights into how such a solution could be designed and how well it could perform motivated the study conducted.

4.1.2 Contributions and Objectives

As there is no solution to this monitoring need, a new approach is needed. All of the requirements mentioned above lead to the question formulated: "Given the scarce availability and usability of data in low voltage distribution grids, what approach is the best fit to detect grid-supporting devices' misconfiguration in such grids?"

A number of objectives has to be fulfilled to answer this question:

- Detect unusual transformer profiles given only aggregated medium-low-voltage transformer data.
- Distinguish abnormal transformer operational data depending on the underlying cause.
- Based on the same aggregated transformer data, gain information on the behavior of decentral devices.
- Determine the data and its quality needed to offer a useful accuracy of detection.

These requirements and the goals that stem thereof led to the main contributions of the present work: a detection method using traditional Machine Learning (ML) methods on the transformer level is introduced and elaborated using a novel DSM use case. Furthermore, a disaggregation approach using load estimation is laid out that helps gaining information on the low voltage level given the transformer operational data. To conclude, the detection method and the disaggregation method are combined to form a detection application that is suited to be installed in DSOs' control rooms as a decision support tool. The quality and influences of the individual parts are assessed in the course of the work, helping to reach the objectives set.

4.1.3 Organisation

The remaining work's structure can be listed in the following manner: In Section 4.2, the state-of-the-art related to configuration monitoring in power systems and the usage of data-driven methods for the same are treated. Section 4.3 details on implementation and functionality of the detection and disaggregation method and lines out the assembled detection application. In Section 4.4, the use cases and deployed grid setups for assessing them are described. Section 4.5 presents the results achieved for each application and stage. Finally, in Section 4.6 the discussion, conclusions, and an outlook about potential further work are given.

4.2 Related Work

In the literature, no solution for the stated problem exists. However, approaches to solving parts of it can be found, even if they might not be straightforwardly applicable. In the following, these related contributions are assessed in this regard.

4.2.1 Background and Data Preprocessing

As the review conducted in [13] outlines there are various disciplines when it comes to monitoring power systems. The first applicable one is fault detection, which treats the detection of the occurrence of deviations from regular operating conditions. Another one is fault classification, which encompasses the fault type to be identified. Lastly, fault location is mentioned, which obviously means localizing the fault which is constituted by singling out the part of the grid covered by the substation. Therefore, the problem at hand falls into the category of fault diagnosis, as this discipline combines all the aforementioned challenges. However, the publication mentions explicitly the lack of automatic fault location methods implemented by DSOs. The same applies to fault classification, whereas there is even only a small number of publications in this field. What also remains to be said is that the review mentions only line fault location and classification scenarios. Also the review conducted in [14] does not address misconfigurations but only power quality disturbances and lists causes for numerous ones. It only provides methods on how to detect these. Another current related review article [15] only addresses power system frequency and control as an application of Deep Learning. It does not mention misconfiguration detection in the way it is regarded here. The last related review article to be found is [16], which treats condition monitoring of wind power systems. This is also somewhat related to the problem defined, however, this is a very general monitoring task. The specifics of it do not necessarily apply to the problem at hand. For the misconfiguration detection case treated in the work presented here, no reviews or related works apart from previous work by the authors could be found.

Treating data in the principal component subspace in order to reduce dimensions and filter for relevant features using Principal component Analysis (PCA) is a promising strategy. PCA assesses which components, meaning which features of a sample, have the highest impact on the feature vector. This is done by evaluating which projection of the data onto a vector retains the most variance in the data. This vector is the first primary component. If one wishes to keep a certain percentage of the variance of the data, an according number of primary components can be used to represent the entire data. This means the data can be projected onto these components and then use said components instead of the original data. In [17] this approach is employed, also using measurements at the substation bus where a data-driven operation model is assembled. However, once more this solution is used to detect line faults and not misconfigurations.

4.2.2 Monitoring Approaches

In [18], fault diagnosis of single-phase to ground and three-phase faults are conducted using gradient boosting trees. Even though the application is once again not congruent with the one at hand, Decision Trees (DT) are of interest for the detection and classification of misconfigurations. This method works by using a training set of data in order to divide the feature space along linear decision boundaries. This is done iteratively until, ideally, only samples of the same class remain in the so-called leaves of the formed branches of the decision tree. To avoid overfitting, the depth of the tree can be limited, leading to impure leaves but a better ability of the trained model to generalize. These decision boundaries can then be used to classify new, unseen samples. Even though this appears to be applicable, the method presented makes extended use of feeder data measurements, which is to be avoided here. Also, [19] uses Random Forest Decision Trees to localize faults. Here, they are used as regressors to estimate the distance on a feeder to a fault, but also to identify the faulted branch. Therefore, DTs are considered a detection method to be assessed.

Deep Learning (DL) is another approach for fault detection and location proposed in [20]. The properties of feeders are learned by a deep neural network such that it is able to generalize on fault location and occurrence. The main advantage of DL is that it is able to condense its own features from the data, making complex preprocessing of the input data unnecessary. The method is, as the authors elaborate, able to do this even if only measurements at the beginning and end of a feeder are available. For this reason, the method could be of interest. Nevertheless, the training of such a network is conducted with hundreds of thousands of time series covering every imaginable operation scenario. These data stem from simulation and basically constitute a look-up table that is engraved in the deep neural network. [21] applies a DL Attention Mechanism to voltage sag type and location detection. Attention Mechanisms offer weighting inputs according to their importance to the output, thus improving the learning of the relation between the two. The work presents good results, yet the problem of proper data sourcing remains. For the present work, an approach is to be found that is easy to integrate and scale without major adaptions for new grid setups or changes within the grid. This renders the DL approach impractical for the task at hand.

The work in [22] mentions k-Nearest Neighbour (kNN) as an instance-based learning method that can be used for fault detection and classification. The kNN algorithm classifies samples according to the labels of their neighbors: depending on the number of neighbors and, optionally, the distance to these neighbors, a vote is taken by all training samples on the class assigned to the unseen sample. This means no classifier has to be built per se, making kNN a non-parametric method. The only parameter is said number of nearest neighbors considered and whether their vote should be weighted according to their individual distance to the sample to be classified. A large number of data samples, however, lets the computational cost of kNN explode, as all of them have to be evaluated when making a prediction. Also, [23] uses kNN for fault detection, but also to classify events like PV outages. However, data at high resolutions from phaser measurement units

(PMU) positioned in the grid measure voltage magnitudes and also angles. Additionally, [24] finds kNN a fit solution for detecting voltage sags in distribution grids, which also shows that the method is not necessarily limited to line faults. Even though kNN seems to be an approach worth being explored, it remains unclear if it can also perform well under the present circumstances.

The authors of [25] use a Support Vector Machine (SVM) classifier for fault location. The SVM is suited for small datasets with data of high dimensionality and building the classifier has a low computational cost. Furthermore, the SVM can be used using various kernels, allowing for non-linear decision boundaries. In general, decision boundaries are found by the large margin principle: the decision boundaries are calculated in a way to maximize the margin of the samples to the decision boundaries. They also use dimension-decreased data for their solution. The results are promising, however, the solution works using micro-PMU data which is not available in that form for the problem at hand. [26] also uses the SVM for fault detection and location, but uses an online data bank of simulated fault locations to build the classifier. It is to be evaluated if this poses an interesting approach to solving the problem of data availability. Additionally also [27] uses the SVM approach, in this case, to classify power quality disturbances such as well, harmonics, flicker, or interruptions. This shows the wide range of applications the SVM can cope with, making the approach of particular interest to the problems stated initially.

4.2.3 Disaggregation Approaches

The problem of disaggregating a load profile into its contributing profiles without using and installing sensors that track them directly is generally known as Non-Intrusive Load Monitoring (NILM) [28]. In general, only disaggregation of household profiles into individual appliance profiles is found in literature [29]. NILM is only partly congruent with the approach to transformer load profile disaggregation needed here, as the origin and availability of the input data differ as well as the sought output profiles as these usually fall into application categories. Of particular interest here is energy estimation as elaborated in [30]: this estimation is further dissected into event-based and eventless-based NILM. It is to be said that, even though an arbitrary number of appliances can make up the profile to be disaggregated, the appliances are usually identified beforehand. Identification also means extracting a particular typical profile for each appliance, which can be time-consuming and requires a lot of adaptations. This might be cumbersome or not possible in a grid setup as loads appear very different depending on their position in the grid and the resulting influence of lines on these loads' consumption. This is especially a problem for the application at hand if not many adjustments are to be done for individual grids.

Approaches for disaggregation regarding entire transformer profiles in a distribution system using substation data only treat estimating PV or other distributed generation [31]. Here again, historic load data is used which corresponds to the appliance identification mentioned earlier. This can pose a problem again as not many manual adaptions to certain grids are to be conducted for the solution envisioned. Similar approaches that can be found only treat system-level disaggregation, meaning disaggregating even more aggregated power profiles such as national consumption. In [32], the authors disaggregate national consumption into substation-level contributions, which is still of too big granularity for use cases related to our work.

In [33] the NILM problem is reframed as a source separation problem, meaning that the source of an aggregated profile is to be determined. Mostly Neural Network (NN) architectures are proposed for the task. Here. only high-frequency signals are being treated, for example, 16 Hertz signals. If the method works on data with lower frequencies is still to be evaluated. The authors of [34] address the problem as feeder-level disaggregation, which means disaggregation of substation profiles into components. They use NNs for this task, however, to make quantile predictions and not point predictions, as would be more suited to the problems stated above. Lastly, [35] mentions a wide range of NN architectures to be fit for the task of disaggregation, even though the use case under scrutiny is not a feeder-level disaggregation but the disaggregation of a household profile into its devices' contributions.

Different regressors for disaggregation are compared in [36]. The results of this work show more or less equal performances for different regressors such as NN, SVR, or Random Forest with each of them having an edge over the others depending on the dataset they are applied on. Therefore, no solution seems to be outstandingly favorable over the others. Also, the work in [37] uses very simple regression methods for disaggregation: even Linear Regression (LR) seems to offer a good option. For that reason, also well-known and simple solutions such as LR should be considered for the problem at hand.

Table 4.1: Non-functional requirements (NFR) fulfilled (X) or unfulfilled (–) by approaches in related publications cited.

	Reference								
NFR	[17]	[18], [19]	[20], [21]	[22] - [24]	[25] - [27]	[28] - [30]	[31], [32]	[33] - [35]	[36], [37]
Scalability	–	–	–	X	X	–	–	X	X
Adaptability	X	X	–	X	X	X	X	X	X
Integrability	X	–	–	–	–	–	–	–	X
Usability	–	X	–	X	X	–	–	–	X
Data Retention	X	–	X	X	X	X	–	X	X
Robustness	–	X	–	–	X	X	X	X	X
Quality	X	X	X	X	X	X	X	X	X

4.2.4 Summary and Open Issues

Summarisingly, the work on monitoring with regard to misconfigurations as well as the works on disaggregation (see Table 4.1) in the electrical grid domain shows quite some gaps to be filled. Either the approaches don't treat the same problem as the one at hand as is the case for monitoring which covers only line faults and not misconfigurations. Here, some methods appear worth exploring, mostly traditional Machine Learning (ML) approaches such as kNN or SVM. Deep Learning seems not applicable as the data set size and availability of data, in general, is a problem. Regarding disaggregation mostly

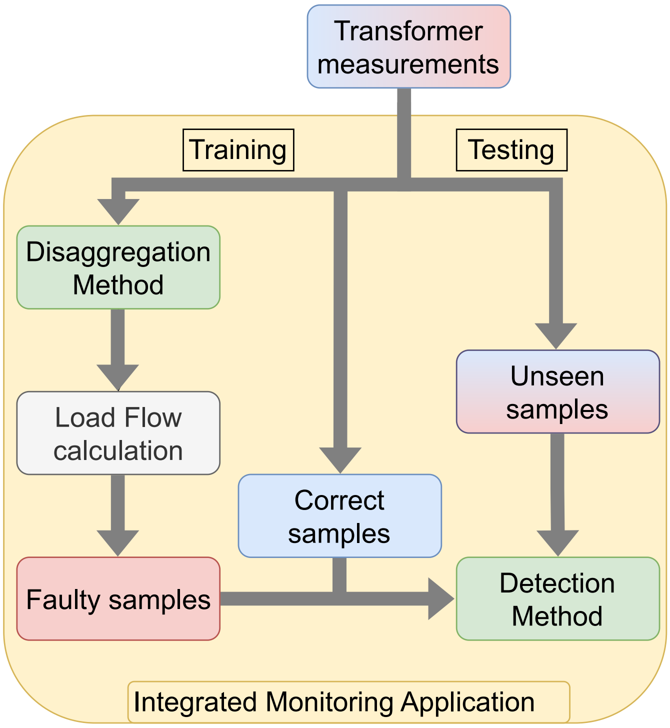


Figure 4.2: Flowchart linking the methods of the monitoring application.

the household level NILM is the focus of the works found in the literature. A wide range of methods, traditional regressors but also NNs seem to be valid solutions. It remains to be seen if they can also be applied in the same manner on the feeder-level disaggregation problem present here, as the constraints are quite different from the ones usually found with NILM problems. The main contributions of the work here are to apply detection approaches in the misconfiguration use case, but also combine them with methods applied to the feeder-level disaggregation task in order to yield an integrated and deployable misconfiguration monitoring solution.

4.3 Methods and Algorithms

The approach presented in this work integrates a transformer profile disaggregation method in the form of a load estimation with a detection method for specific misconfigurations. First, this detection method is explained in detail. Then, the disaggregation method is elaborated. Ultimately, the complete detection application and its functionality are illustrated. How these individual parts are linked is also depicted in the overview flowchart of Figure 4.2. The entirety of the code used to develop, test, and assess all of the methods presented can be found in the corresponding repository[1].

[1]https://github.com/DavidFellner/Malfunctions-in-LV-grid-dataset

4.3.1 Detection Method

The detection method used has been presented in detail in previous work [38]: here the applicability of the method was shown using laboratory data as well as their respective recreations using simulation of substation data. The misconfiguration to be detected was a PV reactive power control misconfiguration. Such misconfigurations can be, for example, a flat $cos\phi(P)$ curve which leads to a lack of voltage control. The method works on substation transformer measurement data. It constructs a classifier using samples of regular grid operations as well as of the corresponding operation circumstances when a misconfiguration is present. The classifier is then used to assess new, unseen samples making a statement on whether a misconfiguration is present or not. The assessment yielded good results, both on the recorded laboratory data as well as on the simulation data. The misconfiguration could be detected in all, or almost all cases, depending on the specific misconfiguration, its position in the grid, and the data source. This paved the way for further usage of this method and its transformation into other use cases as well as its integration into the final detection application.

The approach for detection can be briefly summarised as follows: the highly dimensional substation measurement data, which consists of voltages and currents but also active and reactive power flows are used. These data were recorded at a 4 Hertz rate, and also recreated at the same frequency. The data of one entire day are regarded as a single data sample. This is made possible by flattening the data into a single row, whose columns are marked as certain variables at a specific time step, such as the voltage at phase A at 10 am. This yields a number of columns equivalent to the product of the number of rows times the number of channels recorded. 15 days' worth of data were recorded in all configurations, totaling 15 data samples for the assumption of regular operation as well as for cases where a certain misconfiguration is present. Taking a simple example, 15 sets of load and generation profiles were applied and the substation recorded both for the PV reactive power control to be on or off. The resulting data set has 30 samples in total, 15 correct ones, and 15 samples of malfunctioning cases. PCA was then applied to this data set in order to filter for the most important features. The PCA was specified to retain 99% of the variance in the data, which still reduced the dimensionality of the individual samples significantly. After this step, the final data set is assembled. This data set is then fed to classification methods, namely the aforementioned kNN, DT, and SVM. The latter two build a classifier using the training set and then use the found decision boundary to classify unseen examples as either stemming from regular operation or a misconfigured operational state. kNN, as elaborated before, performs the classification by looking at each testing sample's neighbors and classifying it in accordance with the majority of them. For the PV use case under scrutiny in [38], the best-fitting solution was the SVM.

4.3.2 Disaggregation Method

Usually, only cases of correct operation are recorded or assumed as such as grid operators are unaware of the occurrence of a misconfiguration. A real application would need to know what the faulty samples look like to be able to build the classifier. As only substation data is to be used for the detection method, and some form of recreation of misconfiguration cases is necessary, also some form of data mining to gain information about loads' consumption in the underlying grid is needed.

The approach chosen is a load estimation approach. In order to be able to conduct this estimation, the properties of the grid the loads are situated in have to be captured. In order to achieve this, a training set of generic load flow results is generated using grid simulations. This is done by running 10,000 load flows where loads and generation units are assigned profiles with uniformly distributed values. The results are saved as a data set. The only properties necessary to know here are the minimum and maximum power values of the loads and generation, which should both be available to grid operators since they are either needed for billing or installment of devices. The load flow results, in addition to power flows, then also contain the voltage values for each combination of load and generation settings.

This training set is in turn used to train a NN or build the regressor used for LR as a benchmark. The training set obviously contains the same inputs as used later for the estimation, which is depicted in Figure 4.3: voltages at the substation and at neuralgic points in the grid, as well as active and reactive power, flows at the substation are measured. Furthermore, the production of generation units is assumed as known through external estimation. This estimation is, for example, straightforward for PVs, as radiation models in combination with the installed rated power yield very accurate estimations of production. In the case presented, the estimation is done in hindsight, meaning that the historic radiation data is easy to obtain. The outputs in the training set, the labels, are the active and reactive power consumption of the loads, marked as estimated.

The NN trained is a very simple one, made up of only 1 hidden layer with ReLU as an activation function and Adam as an optimizer. ReLU, in contrast to the Sigmoid activation function, avoids the vanishing gradient problem, which was encountered during developing the solution. Adam optimizer has the advantage of computing individual adaptive learning rates for different parameters which speeds up learning compared to using classic gradient descent for the optimizer. The learning rate was set to 10^{-3} and the batch size to 32. The voltage input data were scaled using a standard scaler which scales the data around the mean divided by the standard deviation. The standard scaler is used here since the voltage values are expected to be clustered around a nominal value, like 230V. The load inputs and outputs were scaled using a minimum maximum scaler, scaling the data between 0 and 1. For the loads, the minimum-maximum scaler was chosen as these values are easy to determine for grid operators from historic billing data, and therefore a minimum and maximum value can be defined for the uniformly distributed inputs. Scaling in these two forms allows for consistent inputs for the NN without any outliers that might inhibit the learning of the model.

Table 4.2 summarises the requirements and the outputs of the disaggregation method by device class. As mentioned before, the voltages and power flows at the substation,

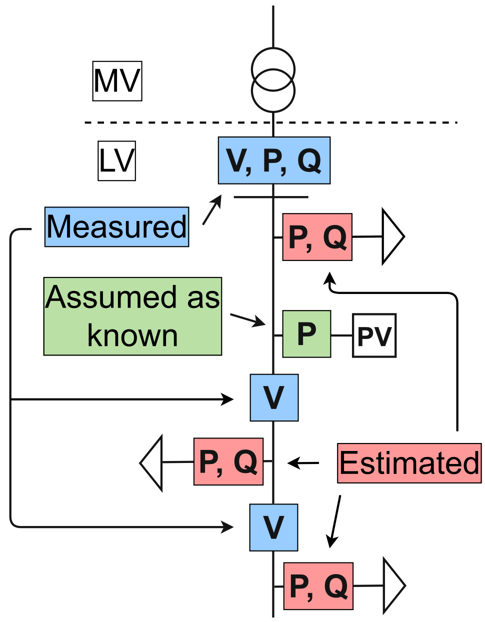


Figure 4.3: Requirements for the disaggregation method.

Table 4.2: Requirements and output of disaggregation method by device.

Device	known	estimated
Transformer	V, p, p	–
PV	p	–
Voltage sensor	V	–
Load	–	p, p

voltages at points between loads, and the infeed of generation devices are needed. The estimation then yields active and reactive power values of the loads in the grid.

4.3.3 Integrated Monitoring Application

In order to merge the aforementioned detection method and disaggregation method into a monitoring application, the two have to be integrated. The functionality of the application is sketched in Figure 4.4 as well as described in the following.

The substation data are used for detection as elaborated above. In order to build a classifier employed for monitoring, a certain calibration period is necessary. During this calibration period, new unseen samples of transformer level data, the data of one day constitute a sample, are assumed to have been collected during regular operation without any misconfiguration present, as in part 1 of Figure 4.4. In order to obtain the corresponding faulty sample consisting of data collected under misconfigured circumstances, grid simulations are employed. These simulations recreate the grid operation of the previous day by setting the load and generation values to the corresponding historic

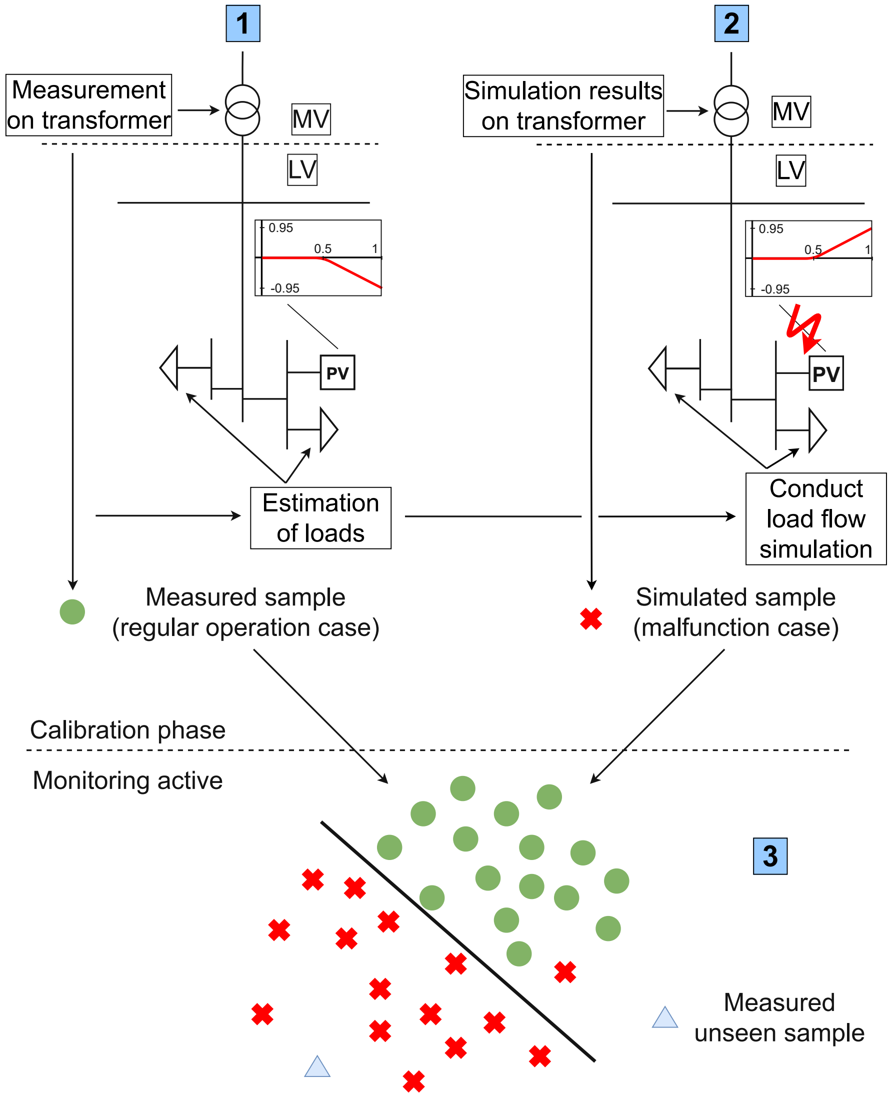


Figure 4.4: Scheme of the integrated monitoring application.

values. During the simulation, the control curve to be monitored is set to a misconfigured setting. Multiple simulations are conducted to cover an arbitrary number of misconfigurations in this way by producing grid operational data under such circumstances. To attain a complete data set, the measured 'correct' samples and the simulated 'faulty' ones are combined. To be able to simulate the 'faulty' samples though, the values of the individual loads in the underlying grid need to be determined. For this task, the already described disaggregation of the measured transformer power profile is used. The simulation is done using the load and generation profiles, the latter are assumed to be known from external sources, and simply changing the configuration under scrutiny to the misconfigured setting. The load flow simulation then yields the transformer data for this respective case, as shown in part 2 of Figure 4.4. This can be repeated for an arbitrary number of misconfigurations without high computational cost and only requires modeling the misconfiguration once.

After the calibration was conducted, which in the presented case was done for 14 days, the monitoring application is ready for use. New, unseen data are then treated, as lined out in the description of the detection method: a day's data is flattened into a single row and treated by PCA to form one sample. This sample is then classified as either stemming from regular operation or not, which is depicted in part 3 of Figure 4.4. Therefore, the monitoring application delivers a diagnosis of device's configuration status once a day. The classifier can then also be updated each day, in case the sample collected is deemed to be of regular operation. This leaves the application with a rolling window of historical data making up the classifier which also accounts for possible drifts in the grid operational data.

4.4 Monitoring Examples

Here the two applications the monitoring approach was tested on, as well as the grid setups the data for these were collected and the corresponding data properties are presented. Both applications are highly relevant to the integration of renewable energy sources into the power system. The PV use case aims at the monitoring of the direct mitigation of the impact of decentral integration through a reactive power control. Such controls are widely configured at PV inverters. The DSM monitoring use case tends to the detection of incorrect load shifting looking to maximize PV self-consumption, which is a more indirect grid-supporting functionality since it also mitigates stress on the grid.

4.4.1 DSM Use Case

The first application is the monitoring of a DSM functionality of loads. The DSM in question aims to shift the load in a way, as to maximize PV generation self-consumption as depicted in Figure 4.5: the load profile assigned to a household load (the red profile) which has a PV generator attached is shifted (the green profile) so that the biggest consumption peak coincides with PV production (the blue profile). Therefore, it is shifted

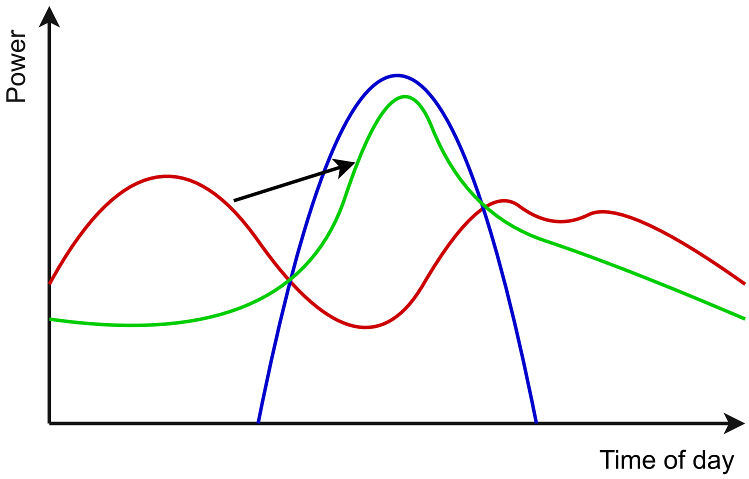


Figure 4.5: DSM working principle; red: original load; green: load after DSM; blue: PV generation.

to sometime during the day, lowering energy demand from the grid. The overall energy consumed throughout the day remains unchanged, though.

This is considered a correctly configured load that implements DSM control. In case the load is not shifted in the way described, the load is considered to have no DSM control. Data were collected in a laboratory environment using two grid setups shown in Figure 4.6: both setups contain a transformer, 3 loads, and one PV unit. In setup A, the PV is located at a load close to the substation and therefore close to the start of the feeder. In setup B, the PV can be found at the end of the feeder. Data were measured at the connection points of the loads as well as at the substation. The data was collected at a 4 Hertz rate, measuring a multitude of signals such as voltages, currents as well as active and reactive power flows.

Using these grid setups, data corresponding to 15 days of grid operation were collected by assigning load and generation profiles and measuring the grid data. Each of the profile combinations is referred to as a scenario in the following. This was done twice as the data were collected once with the DSM control in place, and once with no DSM control. This yielded 30 samples, 15 of which were 'correct' and 15 were 'faulty'. Figure 4.7 allows a glimpse at the data collected: the top part of the figure depicts the voltage measured in the lab environment at the load with attached PV as well as at the transformer. The DSM control helps curtail over-voltages by raising the self-consumption of otherwise excessive PV generation. This effect is more pronounced in the setup where the PV is closer to the substation. The lower part of the figure shows the recreation of these measurements by simulation: the basic behavior is the same, however, the effects of the DSM are less pronounced than in the real-life data.

In order to get a full picture of all scenarios with both DSM control and no control, a

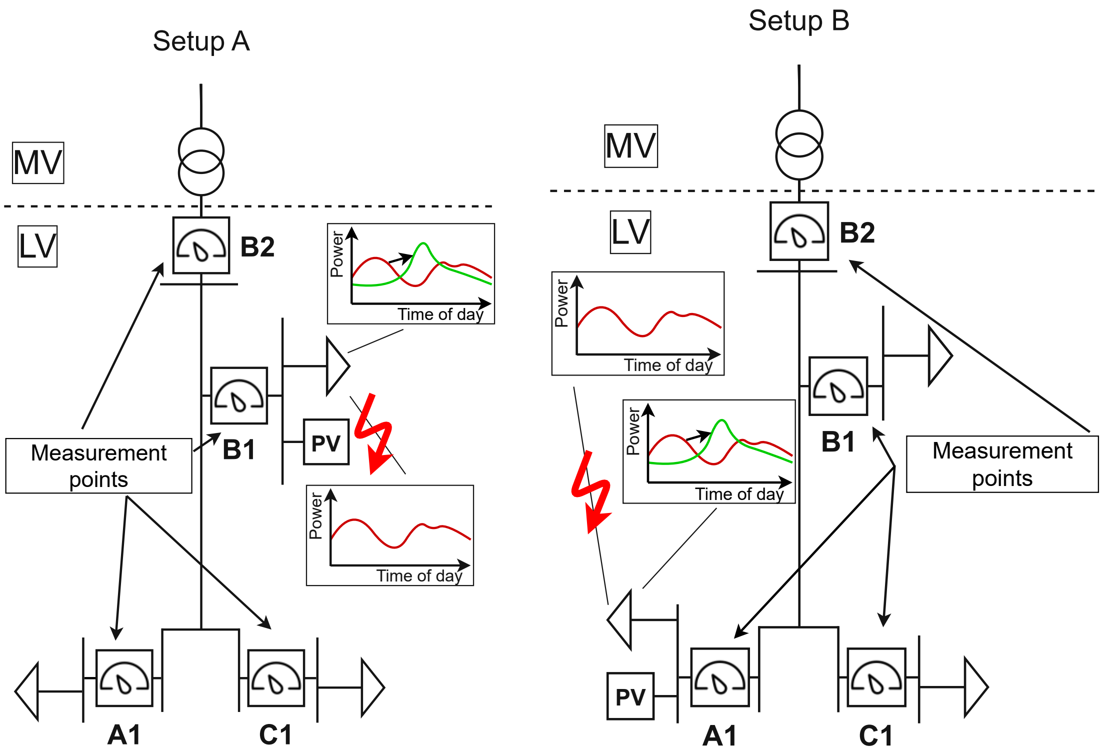


Figure 4.6: Test setups used for data collection.

cluster map can be created. The clustering was conducted using ward clustering [39] which creates a similarity matrix using Pearson correlation and then builds a hierarchical dendrogram linking together the most similar time series. This is done by the ward linkage method, an algorithm minimizing the variance. The clustering was done for the laboratory data, which can be seen in the top part of Figure 4.8, as well as for the simulation data, which is depicted in the bottom part of the figure. Two aspects can be derived from these cluster maps: first that the data from the same scenario are more similar to each other than the data collected with the same control setting. This means that data samples from the DSM and Pv use case are not trivial to separate. Secondly, the laboratory data are in general less similar to each other than the simulated data, which is in accordance with the observation made earlier that the impact of the control is less pronounced in the simulated case. This could have implications for the performance of the monitoring application, as it combines real-world measurement data with simulated data. However, a sample from the real world shows more pronounced effects of the control, meaning it should be easier to detect in case of a misconfiguration than its simulated peers used to build the classifier, as elaborated above.

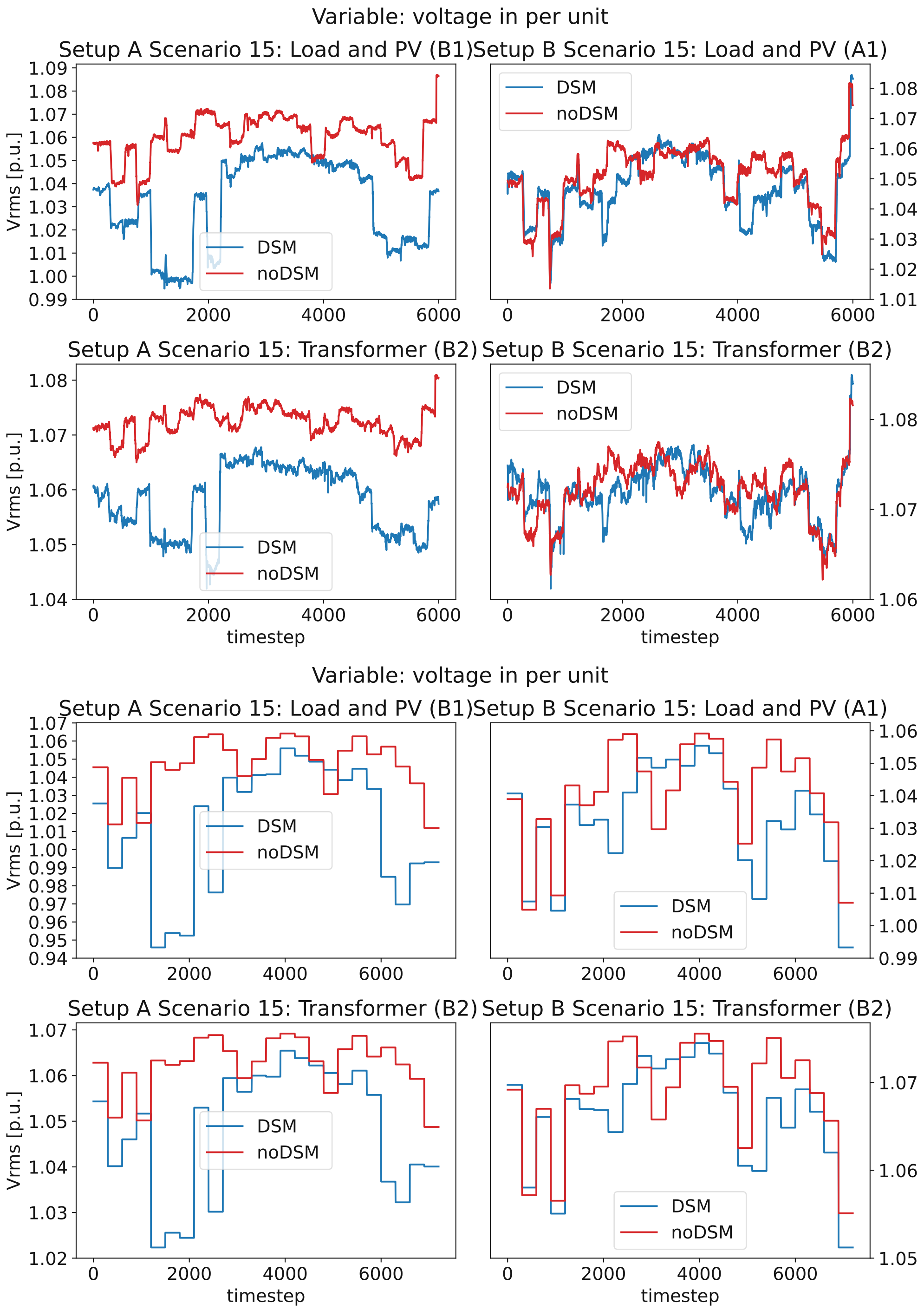


Figure 4.7: Laboratory (top) and simulation data (bottom) by measurement point.

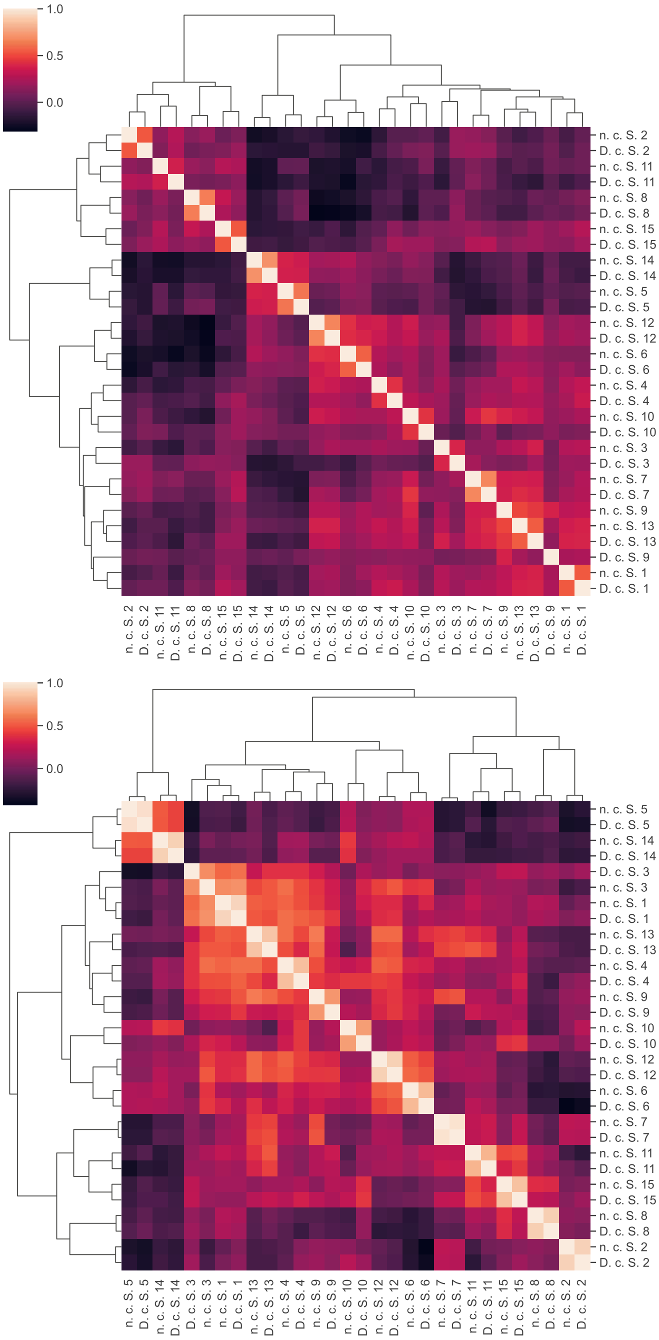


Figure 4.8: Laboratory (top) and simulation data (bottom) of setup B at measurement point B2 clustered; 'D. c. S. 1' and 'n. c. S. 1' stands for 'DSM control Scenario 1' or 'no control Scenario 1' respectively.

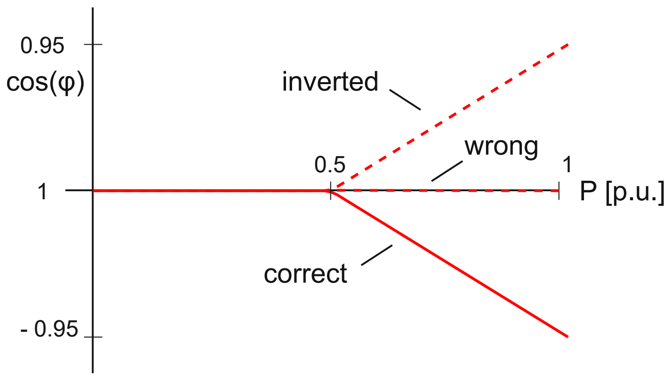


Figure 4.9: cosϕ(P) control curve and its abnormal configurations.

4.4.2 PV Use Case

The second application is the monitoring of a PV inverter and its reactive power control curve. The curve under scrutiny is a cosϕ(P) control curve. The misconfigurations, sketched in Figure 4.9, are either a flat control curve, called 'wrong' in the following, which means no reactive power infeed, or an inversed curve leading to an infeed of the opposite sign. As mentioned above, this power factor control is used to dispatch reactive power in order to avoid or mitigate overvoltages at high PV active power infeed. The same data as described above was collected for this use case, also using two grid setups. Both consisted of a transformer, two loads, and a PV generation unit. In one setup, this PV is located closer to the substation, in the other one, at the end of the feeder. A detailed description of these setups, the control curve as well as its misconfigurations, the data collected, and the results of the detection method achieved on this data can be found in previous work [38]. Also here, the number of samples collected for each case of configuration is 15, meaning 15 days' worth of data were collected.

4.5 Results and Discussion

The results achieved by the individual parts of the monitoring application, but also by the entire application are shown here.

4.5.1 DSM Detection Method

First, the performance of the detection method is evaluated. This is done for the DSM use case, as the results for the PV use case can be found in previous work [38]. Table 4.3 shows these results: the first row shows the F-score achieved in both grid setups and when using data collected in the laboratory as well as through simulation. The results were achieved by conducting a 7-fold cross-validation. The F-score is calculated using Recall, how many of the misconfigurations present were also found, and Precision, how many of the found misconfigurations are actually misconfigurations. The F-score, therefore, balances the two. The second row lists the classifier yielding the best result.

Table 4.3: Comparison of best detection results on laboratory and simulation data of the DSM use case.

Metric	Grid Setup A		Grid Setup B	
	Lab Data	Sim Data	Lab Data	Sim Data
F-Score	0.91	0.70	0.85	0.68
Best method	NuSVM: RBF kernel	SVM: sigmoid kernel	SVM: linear kernel	NuSVM: sigmoid kernel

The results clearly show a good performance on the data collected in the laboratory setting, whereas the misconfigurations appear harder to detect in the simulation data. This can be explained when considering the aforementioned higher similarity between samples in the simulated cases compared with the samples collected in the lab environment. In general, the results in setup A have an edge over the results in setup B. In setup A, the DSM-controlled load is closer to the substation and therefore has a higher impact on the transformer data, as discussed already. This makes the misconfiguration easier to detect. However, in both cases the detection is feasible. Furthermore, either the SVM or NuSVM, which constrains the number of support vectors making up the decision boundary depending on the so-called nu parameter ranging between 0 and 1, is found to be the best-performing algorithm for detection. This was to be expected, both considering the results of previous work as well as the properties of the SVM which shows good performance on small, highly dimensional datasets.

4.5.2 Disaggregation Method

The performance of the disaggregation method and its load estimation as the next building block of the monitoring application is assessed here. This was done for both use cases, so for four grid setups in total. Figure 4.10 shows examples for the estimation of a load's active and reactive power consumption for grids used for the PV use case. The active power value is depicted on the left, and the reactive power value is on the right. The estimation is done using both a NN as well as LR. The active power estimation follows the actual value quite accurately, whereas the reactive power seems to be underestimated generally with some peaks in the NN estimation that are off.

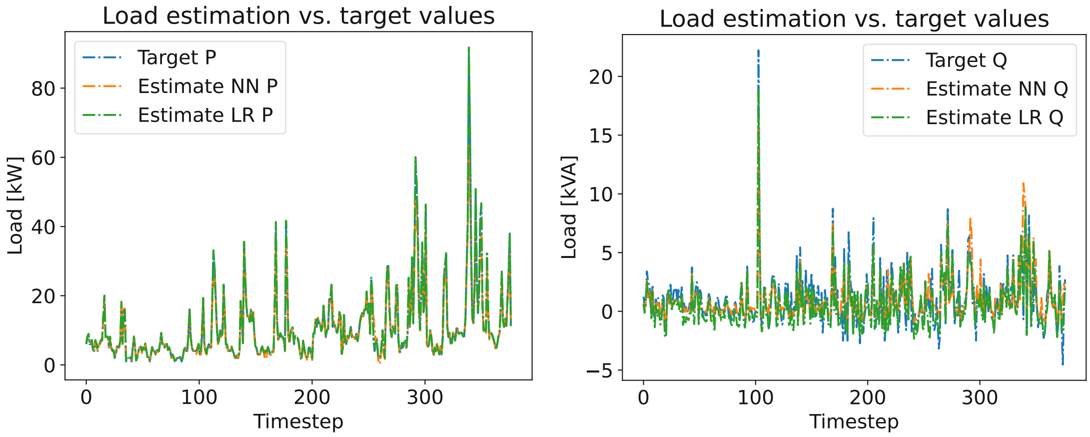


Figure 4.10: PV use case: estimation of load profile.

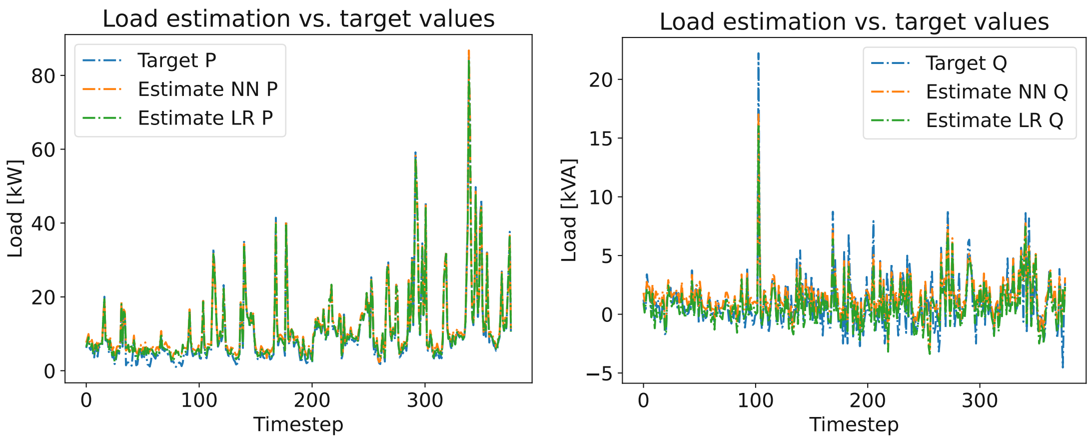


Figure 4.11: DSM use case: estimation of load profile.

The estimation of the consumption of a load, which has the same active and reactive power profile as the one shown before but in a grid used for the DSM use case can be found in Figure 4.11: here the active load estimation seems to be too high in a few instants, whereas the reactive power appears to be more accurately estimated with the LR estimation being farther away from the actual value when it comes to peaks. This allows for the conclusion that the estimation is generally better for the active power values than for the reactive power consumption, which might have to do with the properties of the grid and the varying reactive power consumption of the lines therein.

The complete results on all grid setups are listed in Table 4.4. The metric used is the mean squared error, which in this case is based on the scaled values ranging from 0 to 1. The results for the first two grid setups used for the PV use case are much worse than the results for the second two used for the DSM use case. The grids used for the DSM use case have more loads than the ones used for the PV use case, having 3 instead of

Table 4.4: Comparison of disaggregation error results for both the PV and DSM use cases.

Mean squared error (MSE)	PV Use Case		DSM Use Case	
	Setup A	Setup B	Setup A	Setup B
Neural Network	$23 * 10^{-3}$	$20 * 10^{-3}$	$0.8 * 10^{-3}$	$0.85 * 10^{-3}$
Linear Regression	$17 * 10^{-3}$	$16 * 10^{-3}$	$0.9 * 10^{-3}$	$0.87 * 10^{-3}$

2 voltage measurements and fewer lines without any measurements in general. This is likely to be the cause of the better performance on the 'denser' grids used for the DSM use case. The performance of the NN and LR are almost the same here, with the NN having a slight edge over the LR performance. For the less 'dense' grids employed in the PV use case, the LR shows better performance, pointing to the LR being the more robust option. Whether this has an impact on the overall performance of the monitoring application remains to be determined in the following.

4.5.3 Integrated Monitoring Application

Finally, the detection method and the disaggregation approach were combined and the resulting monitoring application was put to test. The performance results for the PV and the DSM use case are both evaluated using the aforementioned F-score as well as by pointing out the best scoring algorithm for detection. Furthermore, the impact of the disaggregation approach is evaluated. this is done by comparing the performance of the monitoring application using the actual load data as inputs for the simulation of the misconfigured samples to the performance using estimated load data as inputs. Both the NN as well as the LR estimation are considered inputs. As there are 15 samples available for each of the use cases and grid setups, 15 combinations of training and test sets were formed. In each of them, all but one sample of regular operation as well as one of grid operation with a misconfiguration present are used for training. The two test samples originate from the transformer measurements. In this way, the real operation of the monitoring application is emulated.

Table 4.5 and Table 4.6 show the results for the PV use case. The results consider the detection of the flat reactive power control curve, called 'wrong', as well as the detection of an inversed curve, or simply an abnormal curve which means either a flat or inverted curve. The top part of the table shows the results for grid setup A where the PV is closer to the substation, and the bottom half lists the results for grid setup B which contains a PV at the end of the feeder. The detection of the 'wrong' control curve works equally well in both setups, also with regard to the origin of the load data for the simulation. The best-performing detection method here is a form of SVM with a non-linear kernel. The inversed control curve can be better detected in grid setup A when the actual load data are used for simulations, even though the best detection approach is a kNN considering two neighbors indiscriminate of their distance in both cases. There is no difference in performance between the grid setups when the estimated load data are used, however, an SVM with a sigmoid kernel is the best-performing method then. The best performance is achieved when detecting both misconfigurations at once. The performance is the same for both grid setups, with the best performing algorithm being kNN with two neighbors weighted for their distance in the case of the actual load data being used and again SVM with a sigmoid kernel when either the NN or LR estimated load data are used for simulation. In this use case, the detection results are generally significantly worse in case the estimated load data are used. However, it does not seem to matter whether they stem from the NN or LR estimation even though the LR estimation was more accurate as already discussed. This general drop in performance can be attributed to the poor estimation quality for the grid setups used here. In general, the results are acceptable for individual misconfigurations or even good when trying to detect any misconfiguration of the PV reactive power control curve.

The results of the monitoring application on the DSM use case can be found in Table 4.7: The results for both grid setups are quite similar with the only difference being the best method found. In grid setup A, this is a NuSVM with a polynomial kernel of the 4^{th} degree, whereas in grid setup B, it is an SVM with a Radial Basis Function (RBF) kernel. What is of particular interest here is that the performance is the same when using the actual load data as input for the simulation as when using the estimated load data, regardless of whether NN or LR is employed. This can be traced to the much better estimation accuracy in the grids under scrutiny here. This allows defining the MSE as sufficiently small at about 10^{-3} for the estimation not to have an impact on the performance of the detection. The overall results are decent, with them matching the performance of the detection of a specific PV misconfiguration.

4.6 Conclusions

4.6.1 Achievements and Conclusion

The problems raised by the transformation of the electric energy grid need novel solutions such as controls on a device that support the grid to work within operational limits. Due to a lack of sensors in the distribution grid, DSOs need solutions for monitoring the correct execution of these controls, in order to be able to guarantee a reliable and safe operation of the grid. The integrated monitoring application presented delivers

Table 4.5: Comparison of best detection methods using the original or the estimated input data of the PV use case for Grid Setup A.

PV use case: Grid Setup A				
Best score / method		**Data Source**		
Case	Metric	Original	NN Estimated	LR Estimated
correct vs. wrong	F-score	0.71	0.62	0.62
	Best method	NuSVM: polynomial 4^{th} degree kernel	SVM: sigmoid kernel	SVM: sigmoid kernel
correct vs. inversed	F-score	0.80	0.67	0.67
	Best method	kNN: 2 neighbors uniform weights	SVM: sigmoid kernel	SVM: sigmoid kernel
correct vs. abnormal	F-score	0.83	0.80	0.80
	Best method	kNN: 2 neighbors euclidian weights	SVM: sigmoid kernel	SVM: sigmoid kernel

just that tackling the first two objectives set initially. The detection method, as well as the disaggregation method, were both evaluated and then combined to form an easy-to-integrate and deploy monitoring solution that can act as a decision support tool for DSOs pointing them to misconfigurations of controls at a regular interval. A PV inverter and a DSM use case were presented and used for the experiments. The application presented

Table 4.6: Comparison of best detection methods using the original or the estimated input data of the PV use case for Grid Setup B.

PV use case: Grid Setup B				
Best score / method		**Data Source**		
Case	Metric	Original	NN Estimated	LR Estimated
correct vs. wrong	F-score	0.71	0.62	0.62
	Best method	NuSVM: polynomial 4^{th} degree kernel	SVM: sigmoid kernel	SVM: sigmoid kernel
correct vs. inversed	F-score	0.76	0.67	0.67
	Best method	kNN: 2 neighbors uniform weights	SVM: sigmoid kernel	SVM: sigmoid kernel
correct vs. abnormal	F-score	0.83	0.80	0.80
	Best method	kNN: 2 neighbors euclidian weights	SVM: sigmoid kernel	SVM: sigmoid kernel

makes use of data already available to the grid operator, with the sole extension of a PV generation estimation and voltage measurements at certain points in the grid to mine information about loads' consumption. This satisfies the third goal defined. The former are considered rather easy to obtain as they are historic for the past day. The latter constitutes only a small extension to sensing capabilities in the low-voltage distribution

Table 4.7: Comparison of best detection methods using the original or the estimated input data of the DSM use case.

DSM use case				
Best score / method		**Data Source**		
Case	Metric	Original	NN Estimated	LR Estimated
Setup A:	F-score	0.67	0.67	0.67
DSM vs. no DSM	Best method	NuSVM: polynomial 4^{th} degree kernel	NuSVM: polynomial 4^{th} degree kernel	NuSVM: polynomial 4^{th} degree kernel
Setup B:	F-score	0.69	0.69	0.69
DSM vs. no DSM	Best method	SVM: RBF kernel	SVM: RBF kernel	SVM: RBF kernel

grid. Figure 4.12 sketches the problems stated and the contributions made in a condensed way. The contributions include the development of a detection and classification of abnormal transformer measurement data as well as an assessment of necessary data quality and an approach to Data Mining through Disaggregation. Life-like data were collected in a laboratory environment and recreated through simulation to give more validity to the results. These results give insights into the performance of the individual parts as well as of the complete monitoring application. The performance achieved in all scenarios is sufficiently satisfying to serve as a reliable and helpful tool for better monitoring of distribution grids, which fulfills the last aims set in the beginning. The limitations are mainly set by the assumptions on a correctly configured initial state before the calibration of the monitoring solution is conducted. This means previously present misconfigurations can not be detected, only newly occurring ones.

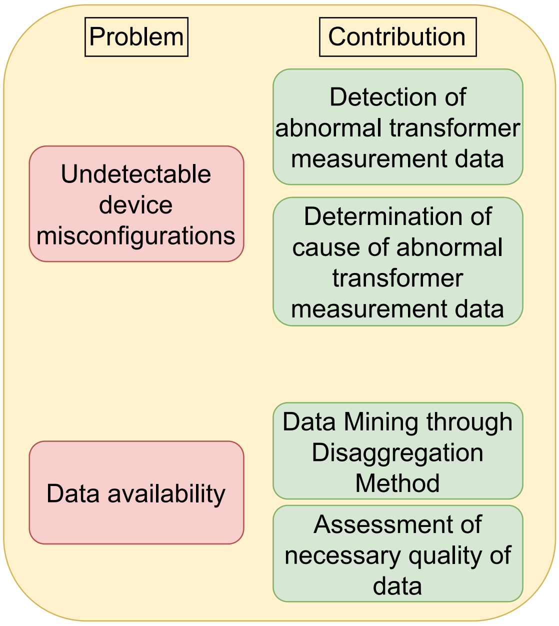


Figure 4.12: Problems stated and the corresponding contribution by the work presented.

4.6.2 Outlook

The application is meant to be working online as a decision support tool for DSOs. Therefore, a field trial assessing the transformer profile disaggregation approach as well as the complete monitoring solution would be beneficial to further improve the application as well as check its robustness. Furthermore, a trial in more diverse grid setups as well as larger grid setups is of interest, to be able to judge the application's scalability. To test the application's adaptability, the integration of new use cases regarding misconfigurations or devices is to be done in future work as well. The examples presented can also serve as templates for integrating other devices' misconfigurations. These are to include EVSEs and HPs.

Conflict of Interest Statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper or pose a conflict of interest of any sort.

4.7 References

[1] A. Q. Al-Shetwi, "Sustainable development of renewable energy integrated power sector: Trends, environmental impacts, and recent challenges," *Science of The Total Environment*, vol. 822, p. 153645, 2022.

[2] Y. Fu, Z. O'Neill, J. Wen, A. Pertzborn, and S. T. Bushby, "Utilizing commercial heating, ventilating, and air conditioning systems to provide grid services: A review," *Applied Energy*, vol. 307, p. 118133, 2022.

[3] J. Caballero-Peña, C. Cadena-Zarate, A. Parrado-Duque, and G. Osma-Pinto, "Distributed energy resources on distribution networks: A systematic review of modelling, simulation, metrics, and impacts," *International Journal of Electrical Power & Energy Systems*, vol. 138, p. 107900, 2022.

[4] S. Rahman, I. A. Khan, A. A. Khan, A. Mallik, and M. F. Nadeem, "Comprehensive review & impact analysis of integrating projected electric vehicle charging load to the existing low voltage distribution system," *Renewable and Sustainable Energy Reviews*, vol. 153, p. 111756, 2022.

[5] R. Stanev, P. Dzhumaliyski, N. Stoychev, K. Viglov, N. Nikolov, and N. Polihronov, "Voltage control strategies for distribution networks with distributed energy resources," in *2022 14th Electrical Engineering Faculty Conference (BulEF)*, 2022, pp. 1–8.

[6] D. Stanelytė and V. Radziukynas, "Analysis of voltage and reactive power algorithms in low voltage networks," *Energies*, vol. 15, no. 5, 2022.

[7] T. Unterluggauer, J. Rich, P. B. Andersen, and S. Hashemi, "Electric vehicle charging infrastructure planning for integrated transportation and power distribution networks: A review," *eTransportation*, vol. 12, p. 100163, 2022.

[8] A. T. Dahiru, D. Daud, C. W. Tan, Z. T. Jagun, S. Samsudin, and A. M. Dobi, "A comprehensive review of demand side management in distributed grids based on real estate perspectives," *Environmental Science and Pollution Research*, Jan 2023.

[9] A. Koirala, T. Van Acker, R. D'hulst, and D. Van Hertem, "Hosting capacity of photovoltaic systems in low voltage distribution systems: A benchmark of deterministic and stochastic approaches," *Renewable and Sustainable Energy Reviews*, vol. 155, p. 111899, 2022.

[10] J. Breitenbach, J. Gross, M. Wengert, J. Anurathan, R. Bitsch, Z. Kosar, E. Tuelue, and R. Buettner, "A systematic literature review of deep learning approaches in smart meter data analytics," in *2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC)*, 2022, pp. 1337–1342.

[11] P. Mastny, J. Moravek, J. Drapela, M. Vrana, and M. Vojtek, "Problems of verification of operating parameters of dc/ac inverters and their integration into the distribution system in the czech republic," in *CIRED Porto Workshop 2022: E-mobility and power distribution systems*, vol. 2022, 2022, pp. 304–308.

[12] D. Fellner, T. I. Strasser, and W. Kastner, "Applying deep learning-based concepts for the detection of device misconfigurations in power systems," *Sustainable Energy, Grids and Networks*, vol. 32, p. 100851, 2022.

[13] P. Stefanidou-Voziki, N. Sapountzoglou, B. Raison, and J. Dominguez-Garcia, "A review of fault location and classification methods in distribution grids," *Electric Power Systems Research*, vol. 209, p. 108031, 2022.

[14] G. Bonde, S. Paraskar, and S. Jadhao, "Review on detection and classification of underlying causes of power quality disturbances using signal processing and soft computing technique," *Materials Today: Proceedings*, vol. 58, pp. 509–515, 2022, international Conference on Artificial Intelligence & Energy Systems.

[15] Y. Zhang, X. Shi, H. Zhang, Y. Cao, and V. Terzija, "Review on deep learning applications in frequency analysis and control of modern power system," *International Journal of Electrical Power & Energy Systems*, vol. 136, p. 107744, 2022.

[16] M. Benbouzid, T. Berghout, N. Sarma, S. Djurović, Y. Wu, and X. Ma, "Intelligent condition monitoring of wind power systems: State of the art review," *Energies*, vol. 14, no. 18, 2021.

[17] L. Souto, J. Meléndez, and S. Herraiz, "Fault location in low voltage smart grids based on similarity criteria in the principal component subspace," in *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2020, pp. 1–5.

[18] N. Sapountzoglou, J. Lago, and B. Raison, "Fault diagnosis in low voltage smart distribution grids using gradient boosting trees," *Electric Power Systems Research*, vol. 182, p. 106254, 2020.

[19] P. Stefanidou-Voziki, D. Cardoner-Valbuena, R. Villafafila-Robles, and J. L. Dominguez-Garcia, "Feature selection and optimization of a ml fault location algorithm for low voltage grids," in *2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, 2021, pp. 1–6.

[20] N. Sapountzoglou, J. Lago, B. De Schutter, and B. Raison, "A generalizable and sensor-independent deep learning method for fault detection and location in low-voltage distribution grids," *Applied Energy*, vol. 276, p. 115299, 2020.

[21] Y. Deng, X. Liu, R. Jia, Q. Huang, G. Xiao, and P. Wang, "Sag source location and type recognition via attention-based independently recurrent neural network," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 5, pp. 1018–1031, 2021.

[22] O. T. Ibitoye, M. O. Onibonoje, and J. O. Dada, "Machine learning based techniques for fault detection in power distribution grid: A review," in *2022 3rd International Conference on Electrical Engineering and Informatics (ICon EEI)*, 2022, pp. 104–107.

[23] M. MansourLakouraj, M. Gautam, H. Livani, and M. Benidris, "A multi-rate sampling pmu-based event classification in active distribution grids with spectral graph neural network," *Electric Power Systems Research*, vol. 211, p. 108145, 2022.

[24] Y. Yalman, T. Uyanik, I. Atli, A. Tan, K. C. Bayindir, O. Karal, S. Golestan, and J. M. Guerrero, "Prediction of voltage sag relative location with data-driven algorithms in distribution grid," *Energies*, vol. 15, no. 18, 2022.

[25] H. Mirshekali, R. Dashti, A. Keshavarz, and H. R. Shaker, "Machine learning-based fault location for smart distribution networks equipped with micro-pmu," *Sensors*, vol. 22, no. 3, 2022.

[26] A. Keshavarz, R. Dashti, M. Deljoo, and H. R. Shaker, "Fault location in distribution networks based on svm and impedance-based method using online databank generation," *Neural Computing and Applications*, vol. 34, no. 3, pp. 2375–2391, Feb 2022. [Online]. Available: https://doi.org/10.1007/s00521-021-06541-2

[27] A. Yilmaz, A. Kuecueker, G. Bayrak, D. Ertekin, M. Shafie-Khah, and J. M. Guerrero, "An improved automated pqd classification method for distributed generators with hybrid svm-based approach using un-decimated wavelet transform," *International Journal of Electrical Power & Energy Systems*, vol. 136, p. 107763, 2022.

[28] M. Kaselimi, E. Protopapadakis, A. Voulodimos, N. Doulamis, and A. Doulamis, "Towards trustworthy energy disaggregation: A review of challenges, methods, and perspectives for non-intrusive load monitoring," *Sensors*, vol. 22, no. 15, 2022.

[29] P. A. Schirmer and I. Mporas, "Non-intrusive load monitoring: A review," *IEEE Transactions on Smart Grid*, vol. 14, no. 1, pp. 769–784, 2023.

[30] S. Dash and N. Sahoo, "Electric energy disaggregation via non-intrusive load monitoring: A state-of-the-art systematic review," *Electric Power Systems Research*, vol. 213, p. 108673, 2022.

[31] M. Toro-Cárdenas, I. Moreira, H. Morais, P. M. Carvalho, and L. A. Ferreira, "Net load disaggregation at secondary substation level," *Renewable Energy*, 2022.

[32] M. G. Pinheiro, S. C. Madeira, and A. P. Francisco, "Short-term electricity load forecasting—a systematic approach from system level to secondary substations," *Applied Energy*, vol. 332, p. 120493, 2023.

[33] P. A. Schirmer and I. Mporas, "Double fourier integral analysis based convolutional neural network regression for high-frequency energy disaggregation," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 3, pp. 439–449, 2022.

[34] X.-Y. Zhang, C. Watkins, and S. Kuenzel, "Multi-quantile recurrent neural network for feeder-level probabilistic energy disaggregation considering roof-top solar energy," *Engineering Applications of Artificial Intelligence*, vol. 110, p. 104707, 2022.

[35] C. Nalmpantis, N. Virtsionis Gkalinikis, and D. Vrakas, "Neural fourier energy disaggregation," *Sensors*, vol. 22, no. 2, 2022.

[36] P. A. Schirmer, I. Mporas, and M. Paraskevas, “Evaluation of regression algorithms and features on the energy disaggregation task,” in *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*, 2019, pp. 1–4.

[37] D.-W. Kim, K.-U. Ahn, H. Shin, and S.-E. Lee, “Simplified weather-related building energy disaggregation and change-point regression: Heating and cooling energy use perspective,” *Buildings*, vol. 12, no. 10, 2022.

[38] D. Fellner, T. I. Strasser, W. Kastner, B. Feizifar, and I. F. Abdulhadi, “Data driven transformer level misconfiguration detection in power distribution grids,” in *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2022, pp. 1840–1847.

[39] P. Zehetbauer, M. Stifter, and B. V. Rao, “Phase preserving profile generation from measurement data by clustering and performance analysis: a tool for network planning and operation,” *Computer Science - Research and Development*, vol. 33, no. 1-2, pp. 145–155, 2018.

CHAPTER 5

The DeMaDs Open Source Modeling Framework for Power System Malfunction Detection

Publication: D. Fellner, T. I. Strasser and W. Kastner, "The DeMaDs Open Source Modeling Framework for Power System Malfunction Detection," 2023 Open Source Modelling and Simulation of Energy Systems (OSMSES), Aachen, Germany, 2023, pp. 1-6.

Abstract: Modeling and simulation of electrical power systems are becoming increasingly important approaches for the development and operation of novel smart grid functionalities – especially with regard to data-driven applications as data of certain operational states or misconfigurations can be next to impossible to obtain. The DeMaDs framework allows for the simulation and modeling of electric power grids and malfunctions therein. Furthermore, it serves as a testbed to assess the applicability of various data-driven malfunction detection methods. These include data mining techniques, traditional machine learning approaches as well as deep learning methods. The framework's capabilities and functionality are laid out here, as well as explained by the means of an illustrative example.

Keywords: Data-driven approach, malfunction detection, modeling and simulation, electric power systems, smart grids.

5.1 Introduction

The development of new smart grid capabilities for electric power grids is essential these days. The transformation towards a sustainable, yet still resilient energy system entails various challenges. These demands can only be faced by novel functionalities [1], which allow the grid to react to the current situation. In order to implement them, but also to test and monitor them, realistic testbeds are needed. However, there are various obstacles to using the electrical power grid as a testbed. The reasons for this are mainly domain-specific: as the power grid is a vital building block of modern life, it is regarded as a critical infrastructure. Any meddling or introduction of non-fully elaborate functionality could compromise its reliability [2]. Moreover, the power grid can not be rebuilt in a scaled-down version that would fully reflect its properties. Furthermore, due to the historical development of the power grid as a hierarchical system, the lower tiers of the network are fairly ill-equip with sensors [3]. These circumstances make data collection and testing in the field, or on a replica of the actual power grid, either difficult or next to impossible.

This leads to modeling and simulation being the only feasible option for early-stage development and assessment of smart grid solutions. This is especially true if these approaches are not only to be tested in a very limited lab setting. Regarding grid models, there is free material available to facilitate these tasks. Very prominent representatives there are the IEEE radial test feeders [4] which are widely used in power system analysis under novel circumstances [5]. Even though the IEEE test feeders feature load profiles, they lack renewable generation profiles and an approach for future scenarios in general. The SIMBENCH project [6] is an open-source project providing specifically designed power grids that allow for the simulation of distribution grids. These models also include scenarios and consumption or generation profiles for electric mobility, battery storage, and novel forms of power generation. In combination with load flow solvers or power grid simulation software [7], these resources can be used to assess the impact and behavior of new techniques in grid operation. The state-of-the-art on these solvers and tools is quite advanced [8] and allows for high computational efficiency [9]. The data generated in the course of this could also be used to develop means of monitoring grid-connected devices.

However, the integration of these solvers with grid simulation and the modeling of specific applications as well as their malfunctions is missing from the literature. This is a prerequisite for the development of monitoring applications. The current approaches are often solely mathematical models not integrating data-driven approaches [10]. In case they do integrate approaches such as machine learning, they only target very common issues and applications; in [11] the authors present a model for predicting general power consumption. The work presented in [12] is more specific focusing on combined heat and power as well as electrical vehicle integration into the power grid. Demand response in a smart grid environment is under scrutiny in [13], however, with mere attention paid to its implementation and not to its monitoring functionalities with regard to correct execution. When it comes to monitoring, significant contributions can be found in the field of security with respect to malicious attacks on the power grid [14]. Nevertheless, this does for example not cover misconfigurations occurring during regular operation. These misconfigurations can lead to malfunctions of the grid-connected device.

The framework presented now aims to fill this gap by providing modeling and data generation, processing, and analysis capabilities. It is designed to serve as a testbed aimed to develop and assess functional monitoring solutions. The approach strives to detect malfunctions during the regular operation of grid-connected devices. The grid setups to be used can be arbitrary. Also, the malfunctions under scrutiny can be modeled freely, as well as a variety of detection methods employed. This is demonstrated in detail in the previous works of [15] and [16]. These features allow for the easy expansion of monitoring use cases. Furthermore, the final detection application can be parameterized freely to facilitate development.

The manuscript has the following content: In Chapter 5.1, the general motivation and background for the work the and field of application of the software framework are presented. Chapter 5.2 provides an overview of the framework, its architecture, and its functionalities. Chapter 5.3 provides insights into the application of the framework by illustrating an example use case in detail. Chapter 5.4 outlines the impact the framework has as a testbed for the development of monitoring solutions for power system operators. Finally, Chapter 5.5 provides the conclusions and an outlook about potential further work.

5.2 Framework Description

The framework is entirely written in Python and the implementation can be found on the corresponding GitHub repository[1]. The most important dependencies regarding external libraries and their use in the framework are illustrated in Chapter 5.1; almost all libraries used are free and open-source libraries, with the exception of a library to interface the here-employed power grid simulation software, DIgSILENT PowerFactory. As there is sample data provided in the repository, the use of such software is not mandatory. Furthermore, any grid modeling and simulation solution can be used in combination with the rest of the framework. In addition, a script which is under development is used for load estimation.

However, other implementations of this functionality can be used as well. This means there are no crucial parts of the framework that are not openly accessible. The common Python libraries are made for data handling and path allocations, whereas for the classic machine learning capabilities Scikit-learn [17] is used. For deep learning, especially for the recurrent neural networks employed, Pytorch [18] is being used. For regular neural network applications, Tensorflow [19] is applied. The choice of using different libraries for the implementation of artificial neural networks depending on their type was made in order to allow for increased flexibility when developing a solution. Pytorch enables the developer to adjust and craft the desired architecture in greater detail in comparison to TensorFlow. This is especially interesting when trying to craft a monitoring solution in a setting like the power grid, as the relevant properties of the data and features are widely unknown beforehand.

[1] https://github.com/DavidFellner/Malfunctions-in-LV-grid-dataset

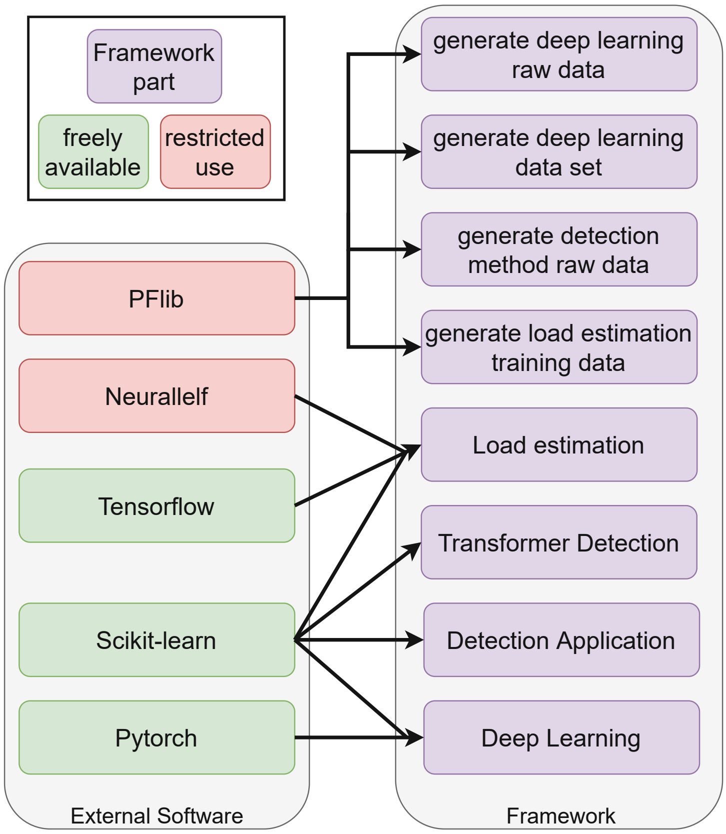


Figure 5.1: The dependencies of the framework.

5.2.1 Software Architecture

The architecture of the framework differs depending on the use case of the respective part of the software. Chapter 5.2 depicts the software architecture of the framework.

The basic settings for the experiment to be conducted by the framework are defined in the configuration file. These settings include data paths and directories as well as configurations for the machine learning, or deep learning approaches that are to be used. The settings also define the neural network models and classifiers to use and how many layers or what type of kernel they should be parameterized with. Also, settings for the loading or creation of data and the assembly of datasets can be specified. These include the specification of the grid models or malfunctions, in order to define the use case the detection is applied to. Further settings include the mapping of data to align real-world measurements with simulation results, for cases in which these two data sources are to be combined.

Then, the data set generation or import of the defined use case is done via functions. Functions are chosen here in order to allow for easier integration of different data sources or grid simulation tools. The functional interfaces are easier to adjust or exchange in

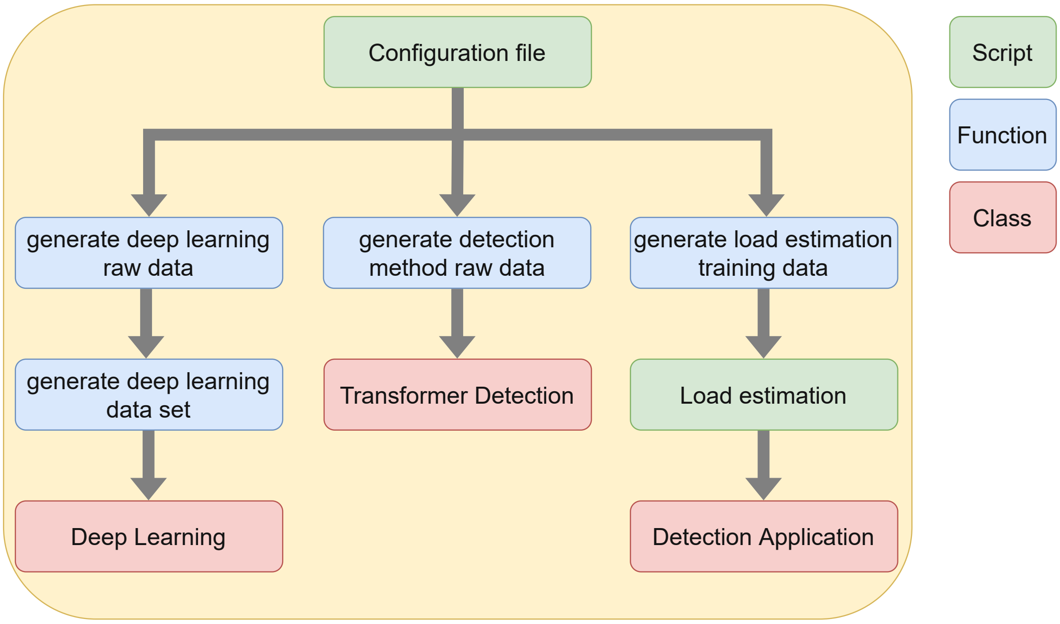


Figure 5.2: The software architecture of the framework.

comparison to an integration of these data handlers within classes. Depending on the use case, this data is then saved. In the case of deep learning, the created data sets are also saved as their compilation is more computationally expensive compared to the data sets used for other approaches.

For experiments testing not a single detection method but a pipeline of methods that form an approach to a practical detection application, load estimation is done via an external script. This script is still under development and therefore not fully integrated with the Detection Application class. This also allows for the use of alternative load estimation or generally data mining approaches more easily.

The main functionalities regarding malfunction detection are bundled into three classes: Deep Learning and Transformer Detection both serve as isolated test beds for methods. These are either grid-unspecific using device-level data in the case of the Deep Learning approaches or grid-specific using transformer data for Transformer Detection. The last class, Detection Application, then allows for the integration of the individually assessed methods into a practically applicable detection application.

5.2.2 Software Functionalities

The framework, as already mentioned above, allows for a great variety of scenarios in which the modeled malfunctions are to be detected. The malfunctions are modeled as incorrect control curves of devices whose behavior are reflected in grid operational data. This detection can be tested and validated in different grid topologies, using differently sized and composed data sets of different origins. The data can originate both from simulation or real-world settings such as lab environments. Furthermore, the approaches to preprocessing and data-driven detection can be varied. Also, various options for metrics and visualization of results are given.

When developing deep learning-based detection methods, data generation allows for the generation of large amounts of data. This is done by using an arbitrary number of grid models for simulation. These simulations can be parallelized, to swiftly yield operational data of a certain type of grid-connected device experiencing the malfunction modeled. Moreover, operational data of the correct behavior of these grid-connected devices is extracted as well. These data can not be obtained in the real world, especially not in a labelled manner as the occurrence of a misconfiguration goes unnoticed at the moment. The results are saved in a CSV format. This data is used to form data sets of the misconfiguration under scrutiny in the use case, which are stored in an hd5 format as they contain up to 200,000 samples. These data sets can now contain data stemming from a single grid or multiple grids. This allows for the assessment of whether the applied deep learning method is able to extract fundamental properties from the data. This is done in order to assess if a specific method can recognize a malfunction without any grid-specific context. The data and the individual samples therein can also be plotted. The framework allows for data preprocessing such as scaling as well as training in various deep-learning approaches. In addition, it enables a comparison to traditional statistical methods. Furthermore, hyperparameter tuning can be conducted. The performance is assessed using common measures such as the F-score, and scores can also be visualized.

Another monitoring approach is provided by transformer-level detection. Here only operational data gathered at the transformer is used. Data is loaded from, or generated and saved to CSV files. Also, both loading of, for example, real-world data, and generation of data is implemented to merge data of different origins. Then the data is preprocessed via Principal Component Analysis (PCA) and combined into datasets. Again, these datasets contain grid operational data of cases in which a malfunction is present or have their origin in regular grid operation. As there are data in a higher resolution as well as more data channels available in this setting, traditional machine learning approaches are to be tested here. This is due to the meters at substations measuring more variables, and these at a higher rate, than smart meters in the distribution grid. As this case is grid-specific, also more advanced tools of data analysis such as hierarchical clustering are available. This clustering helps to assess whether possible real-world data from a specific grid aligns with simulated data. Various classifiers can be applied which can then be assessed by the aforementioned range of result metrics and their plots.

The so-developed and assessed methods can be tested in a near-to-life setup which is represented by the detection application. Here, in order to fill gaps in data that were assumed to be known in the isolated method testbeds, also load estimation is conducted. A load estimation approach using a neural network is trained. Therefore, training data is generated in a similar manner to the cases described before and saved in a CSV file. Moreover, this load estimation is compared to a linear regression estimation to benchmark it. Using this estimation for data mining, data sets can be assembled in a manner similar to what they could also be collected like in the field. This aims at testing the performance of the detection methods under more realistic conditions. The data mining approach is also kept flexible in order to test the methods under different assumptions on which data is available. The result metrics can be inspected at every step of this pipeline to identify the potential for enhancements.

5.3 Illustrative Example

To complement the above-elaborated description of the software with a more tangible example, one use case is described in detail below (cf. Chapter 1).

The crucial parts of a sample configuration file for testing a deep learning application on an electric vehicle charging station use case are presented. At first data paths are defined, both for the grid data used as well as for results and the dataset. Then the specific dataset to be used is defined along with the use case, which is done by choosing the device type that is to be monitored for malfunctions. Following, parameters for the type of neural network used for detection are specified along with training parameters such as the number of epochs, or the optimizer. The great flexibility in the choice of these parameters is made possible by the before-mentioned use of Pytorch. Also, the result metrics can be chosen, as well as settings for a grid search in order to be able to tune hyperparameters.

In the next section of the configuration file, the dataset to be created can be specified. If a dataset is already set to be available no new dataset is created. If not so, the number of samples the dataset created should contain, or how long a sample is, is defined. Also, the number of grids the samples should be drawn from can be specified. Lastly, settings on the grid simulation which creates the dataset can be customized. Parameters such as step size or how many cores should be used for parallelization can be set, along with the exact type of malfunction. In this case, as shown in Chapter 5.3, a generic active power control curve of an electric vehicle charging station is inverted, which is considered the misconfiguration to be detected. The curve depends on the voltage, meaning in the malfunctioning case active power consumption is not reduced at low voltages which therefore constituted the detectable anomalous behavior. The red line marks the correct control curve, whereas the blue line is the inverted, malfunctioning control curve.

These settings and parameters are then used to either create or import a grid model. Such a grid model is depicted in Chapter 5.4. The grid is modeled with the specified amount of, for example, photovoltaic units or electric vehicle charging stations. Some of them are then in turn modeled with the malfunction specified. Then grid simulations are run and data is collected at the devices' connection points to the grids, which are symbolized by the triangles, boxes, or circles connected to the lines. The data is then used to assemble datasets. These are then used for training and testing the specified deep neural networks.

The so-trained neural networks are used for the detection of malfunctions in the test set.

```python
import os
import math

# Sytem settings
grid_data_folder = os.path.join(os.getcwd(), 'raw_data_generation', 'input')
raw_data_folder = os.path.join(os.getcwd(), 'raw_data')
...

# Deep learning settings
learning_config = {
    "mode": "train",  # train, eval
    "dataset": "7day_200k",
    "type": "EV",
    # PV, EV, (PV, EV) > malfunction
    "RNN model settings": [1, 2, 20, 5],
    # dim of in&output, dim of hidden state, # of layers
    "LSTM model settings": [1, 2, 3, 5],
    "R-Transformer model settings": [1, 3, 2, 1, 'GRU', 7, 4, 1, 0.1, 0.1],
    # input size, dimension of model,output size, heads, rnn_type, key size,
    ↪  # local RNN layers, # RNN-multihead-attention blocks, dropout,
    ↪  emb_dropout
    "number of epochs": 20,
    "learning rate": 1 * 10 ** -6,
    "decision criteria": 'majority vote',
    ...
    "activation function": 'relu',  # relu, tanh
    "mini batch size": 60,
    "optimizer": 'SGD',  # Adam, SGD
    "k folds": 5,  # choose 1 to not do crossval
    "early stopping": True,
    "LR adjustment": 'warm up',
    "% of epochs for warm up": 10,
    "train test split": 0.3,
    "metrics": ['accuracy', 'precision_macro', 'recall_macro', 'f1_macro'],
    ...
    "plot samples": True,
    "classifier": "RNN",
    "save_model": True,
    "do grid search": True,
    "grid search": ("calibration rate", [0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5,
    ↪  0.6, 0.7, 0.8, 0.9, 1])
}

# Dataset settings
raw_data_available = True  # leave True if grid simulation is not available
sample_length = 7 * 96  # 96 datapoints per day
number_of_samples = 200000
number_of_grids = len([i for i in os.listdir(grid_data_folder)

# Grid simulation settings
parallel_computing = True
cores = 12
sim_length = 365  # simulation length in days
step_size = 15  # simulation step size in minutes
percentage = {'PV': 0,
              'EV': 25, 'BESS': 0,
              'HP': 0}  # percentage of busses with active PVs etc...
broken_control_curve_choice = 2  # 1 = flat curve, 2 = inversed curve
t_start = None  # default(None): times inferred from profiles in data
t_end = None
```

Listing 1: Configurations for a deep learning use case.

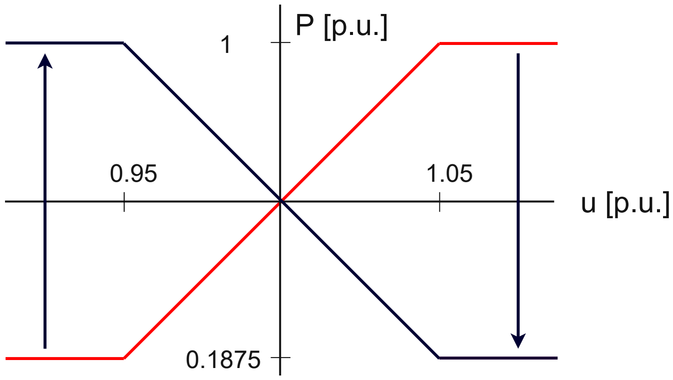


Figure 5.3: Malfunctioning p(U) control curve.

The performance results are then stored and also plotted, as Chapter 5.5 illustrates. Here, the F-score is listed as a metric. The Precision, how accurate label predictions are, as well as the Recall, signifying how many of the true positives were found, are used to calculate this score. The results allow drawing conclusions about the performance of a certain parameterization of a certain deep neural network architecture on a specific dataset. It also allows for easy hyperparameter optimization. The model scoring the best results is saved and can be exported for integration into applications to make demonstrations easy. This should also help facilitate possible field tests of the found solution.

5.4 Impact and Application

The framework's impact is mainly threefold: first of all, it allows for the development of detection methods on a device level, as shown in the previous practical example. This method is intended to work across grid setups; the deep learning approach is meant to extract fundamental properties from the data of devices in regular operation and of devices experiencing malfunctions. Pretraining a network for a certain malfunction then allows the incorporation of the detection solution of this use case into a distribution system operator's monitoring system. Such a solution also enables the operator to know which malfunction occurred. The second aspect aims at developing a detection solution at the transformer level. This is done by using data collected at the substation and applying traditional machine learning methods to it. This detection approach is grid specific. However, it requires no extensive prior training. Only a certain calibration phase would be necessary.

For both application cases, different data sources, data qualities, and data availability can be assessed. Furthermore, different neural network architectures, classifiers, and parameters of these can be compared as well benchmarked against classic statistical methods.

Lastly, the full detection application merges the approaches mentioned above with a full detection application. This means integrating the isolated approaches with data mining

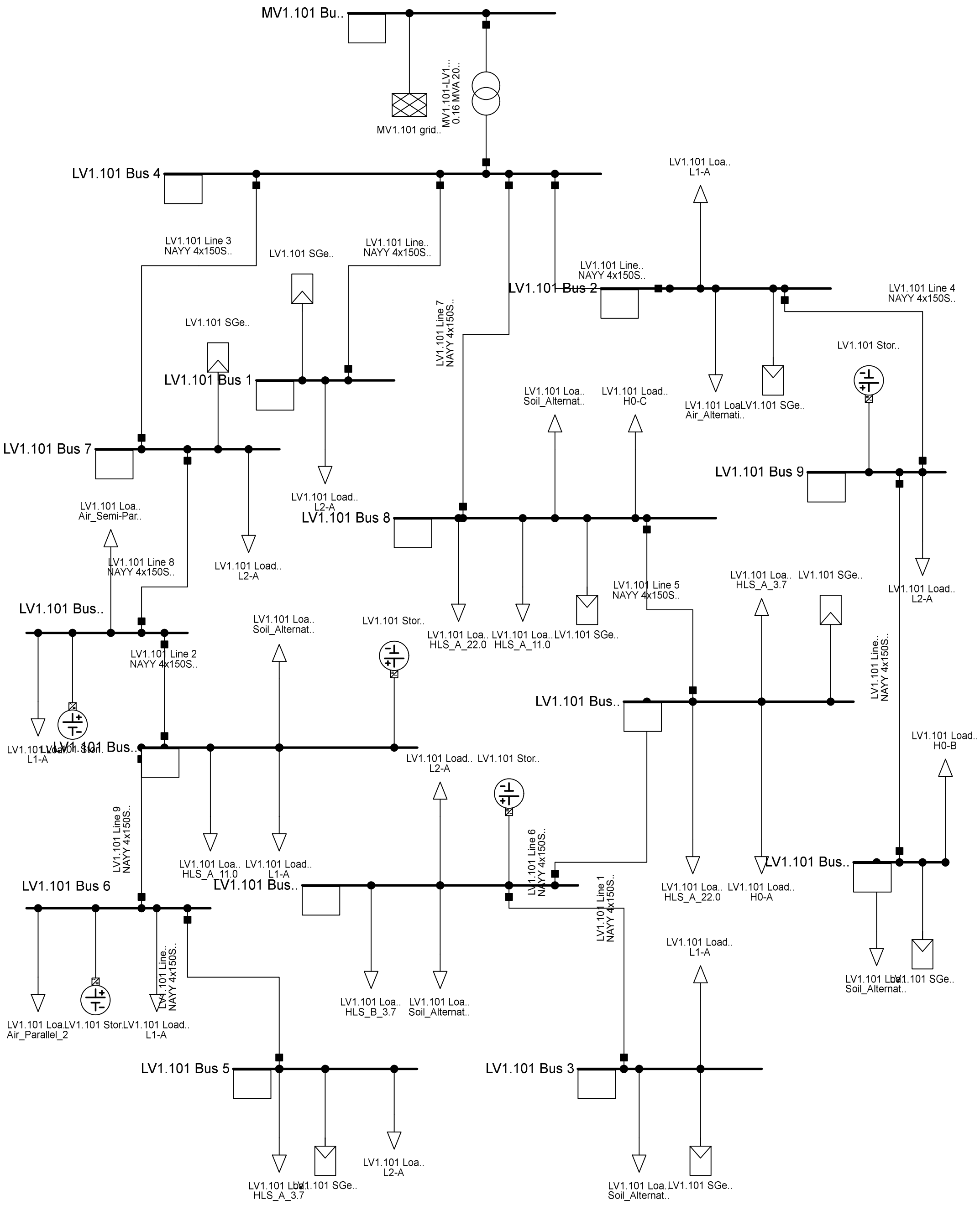


Figure 5.4: Sample power grid (taken from [6]).

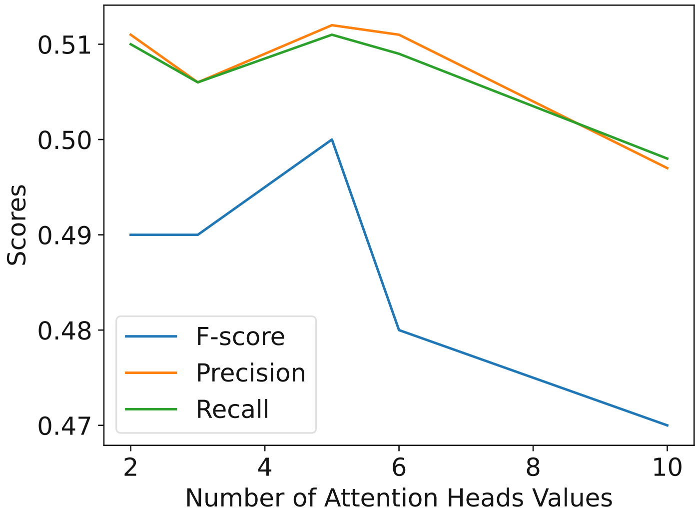


Figure 5.5: Results on hyperparameter tuning.

techniques such as load estimation. This data mining is in turn also either performed by a neural network or by traditional statistical approaches. It can also be tuned to allow for optimal solution development for real-world applications. A testbed of this form did not exist to this point, and as elaborated in the beginning, the real-world power grid can not be used as such. Currently, because of the assumed data availability, its applicability is limited to the adaptation of misconfiguration detection in an LV grid segment linked to the MV level by a substation. However, for this reason, this scope of use cases also has a big advantage in integrability, since few alterations to the grid infrastructure are needed. Therefore, the framework has an impact as an enabler of technology development.

5.5 Conclusions

The work presented describes the need for new monitoring capabilities for smart grids and points out the lack of possibilities to develop such with the means available. Therefore, a framework that can serve as a testbed for novel monitoring solutions for all sorts of new grid-connected devices is introduced here. Various approaches can be tested and integrated into a complete solution. This enables the development of a future detection tool for grid operators. The assessment of this solution can be conducted under as life-like circumstances as possible outside of the grid. The framework is designed in a flexible manner, as to allow users to exchange parts of it. Therefore, it is possible to use whichever means of grid simulation or data mining technique the user prefers.

In the future, more predefined use cases are to be added to reflect the characteristics of more malfunctions. Also, the choice and architectures of predefined machine learning algorithms ought to be updated regularly, in order to keep up with recent developments in these methods. Finally, a field test of the solution as a monitoring tool is envisioned.

5.6 References

[1] V. Saxena, N. Kumar, and U. Nangia, "Smart grid: A sustainable smart approach," *Journal of Physics: Conference Series*, no. 1, p. 012042, Aug. 2021.

[2] R. Cantelmi, G. Di Gravio, and R. Patriarca, "Reviewing qualitative research approaches in the context of critical infrastructure resilience," *Environment Systems and Decisions*, vol. 41, no. 3, pp. 341–376, Sep. 2021.

[3] S. K. Routray, D. Gopal, A. Pallekonda, A. Javali, and S. Kokkirigadda, "Measurement, control and monitoring in smart grids using nbiot," in *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, 2021, pp. 229–234.

[4] W. Kersting, "Radial distribution test feeders," *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 975–985, 1991.

[5] K. D. Shinde and P. B. Mane, "Analysis of radial distribution test feeders in presence of solar photovoltaic systems using powerfactory," in *2022 IEEE International Conference in Power Engineering Application (ICPEA)*, 2022, pp. 1–4.

[6] S. Meinecke and et al., "Simbench—a benchmark dataset of electric power systems to compare innovative solutions based on power flow analysis." *Energies*, vol. 13.12:3290, 2020.

[7] R. Villena-Ruiz, A. Honrubia-Escribano, and E. Gómez-Lázaro, "Learning load flow analysis in electric power systems: A case study in powerfactory," in *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2022, pp. 1357–1362.

[8] M. C. Herrera-Briñez, O. D. Montoya, L. Alvarado-Barrios, and H. R. Chamorro, "The equivalence between successive approximations and matricial load flow formulations," *Applied Sciences*, vol. 11, no. 7, 2021.

[9] Z. Wang, S. Wende-von Berg, and M. Braun, "Fast parallel newton–raphson power flow solver for large number of system calculations with cpu and gpu," *Sustainable Energy, Grids and Networks*, vol. 27, p. 100483, 2021.

[10] D. K. Panda and S. Das, "Smart grid architecture model for control, optimization and data analytics of future power networks with more renewable energy," *Journal of Cleaner Production*, vol. 301, p. 126877, 2021.

[11] S. Tiwari, A. Jain, N. M. O. S. Ahmed, Charu, L. M. Alkwai, A. K. Y. Dafhalla, and S. A. S. Hamad, "Machine learning-based model for prediction of power consumption in smart grid- smart way towards smart city," *Expert Systems*, vol. 39, no. 5, p. e12832, 2022.

[12] F. Calise, F. L. Cappiello, M. Dentice d'Accadia, and M. Vicidomini, "Smart grid energy district based on the integration of electric vehicles and combined heat and power generation," *Energy Conversion and Management*, vol. 234, p. 113932, 2021.

[13] A. M. Eltamaly, M. A. Alotaibi, A. I. Alolah, and M. A. Ahmed, "A novel demand response strategy for sizing of hybrid energy system with smart grid concepts," *IEEE Access*, vol. 9, pp. 20 277–20 294, 2021.

[14] H. Zhang, B. Liu, and H. Wu, "Smart grid cyber-physical attack and defense: A review," *IEEE Access*, vol. 9, pp. 29 641–29 659, 2021.

[15] D. Fellner, T. I. Strasser, and W. Kastner, "Applying deep learning-based concepts for the detection of device misconfigurations in power systems," *Sustainable Energy, Grids and Networks*, vol. 32, p. 100851, 2022.

[16] D. Fellner, T. I. Strasser, W. Kastner, B. Feizifar, and I. F. Abdulhadi, "Data driven transformer level misconfiguration detection in power distribution grids," in *2022 IEEE SMC)*, 2022, pp. 1840–1847.

[17] F. Pedregosa, G. Varoquaux, A. Gramfort, et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[18] Adam Paszke, Sam Gross, Francisco Massa, et al., "PyTorch: An Imperative Style, High-Performance Deep Learning Library," in *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019, pp. 8024–8035.

[19] Martín Abadi, Ashish Agarwal, Paul Barham, et al., "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: https://www.tensorflow.org/

List of Figures

List of Tables

Acronyms

ANN Artificial Neural Networks. 6, 14, 19, 23, 26

BESS Battery Energy Storage Systems. 2, 10

CSV Comma Separated Value. 116

DeMaDs Data Driven Detection of Malfunctioning Devices in Power Distribution Systems. xvi, 24, 26, 56, 76

DG Distributed Generation. 1, 2

DL Deep Learning. ix, xi, 5, 6, 13, 17, 18, 22, 23, 26, 27, 36, 37, 40, 42, 44–46, 49, 53, 56, 84

DSM Demand Side Management. 2, 19, 27, 81, 82, 92–94, 97–101, 104, 125, 126

DSO Distribution System Operators. ix, xi, 4, 6, 9, 26, 27, 34, 36, 49, 62, 74, 76, 81–83, 101, 102, 105

DT Decision Tree. 13, 14, 19, 37, 45, 46, 63, 64, 73, 84, 88

EV Electric vehicle. 1, 2, 10, 50

EVSE Electrical Vehicle Supply Equipment. 7, 17, 27, 34, 36, 42–45, 50, 51, 54, 55, 124, 126

FCR Frequency Containment Reserve. 2

FNN Feedforward Neural Network. 14

GRU Gated-Recurrent-Unit. 13, 38, 40, 47, 48, 51, 54, 55

HP Heat Pump. 10

HVAC Heating, Ventilation and Air Conditioning. 5

kNN k Nearest Neighbours. 13, 14, 19, 37, 45, 46, 49, 64, 72, 73, 84–86, 88, 101–103

kPCA Kernel Principle Component Analysis. 5

KPI Key Performance Indicators. 7, 10, 24

LOF Local Outlier Factor. 37

LR Linear Regression. 14, 19, 23, 86, 89, 98–104

LSTM Long-Short-Term-Memory. 13, 38, 40, 47–51, 54, 55, 124

LV Low Voltage. ix, xi, 8, 121

ML Machine Learning. ix, xi, 13, 17, 18, 22, 23, 26, 27, 36, 37, 40, 53, 82, 86

mRVM Multiclass Relevance Vector Machine. 6

MSE Mean Squared Error. 16, 24, 100, 101

MV Medium Voltage. 8, 121

NILM Non-intrusive Load Monitoring. 14, 85–87

NN Neural Networks. 13, 86, 89, 98, 100–104

NuSVM Nu Support Vector Machine. 45, 46, 49, 98, 101–104

OLTC On-Load Tap Changer. 2

PCA Primary Component Analysis. 5, 13, 19, 26, 63, 68, 83, 88, 92, 116

pSVM Partially Hidden Structured Support Vector Machine. 5, 6

PV Photovoltaic. 1–3, 5–8, 10, 12, 15–17, 19, 23, 27, 34, 36, 42–46, 49–51, 53–55, 62, 65, 66, 68, 74, 76, 81, 84, 85, 88, 90, 92, 93, 97–103, 124–126

RNN Recurrent Neural Network. 13, 18, 22, 37–40, 46–48, 51–55, 124

RTC Real-Time Charging Control. 2

SCADA Supervisory Control and Data Acquisition. 27

SM Smart Meter. 4, 6, 7, 15, 22, 26, 27

SRQ Sub Research Questions. 9–14, 124

SVM Support Vector Machine. 5, 6, 13, 14, 19, 21, 26, 37, 45, 46, 49, 64, 72, 75, 85, 86, 88, 98, 101–104, 124