

Entwicklung von Vorhersagealgorithmen für Ausfälle in komplexen leistungselektronischen Systemen in der Photovoltaik

Akronym: VORAUS PV
FKZ: 13FH094PB8

Schlussbericht des Teilsvorhabens JADE

Version vom 20. August 2024

Projektlaufzeit 01.07.2020 - 31.12.2023
Projektpartner Hochschulen
Jade Hochschule (Jade HS)
Hochschule für Angewandte Wissenschaften Hamburg (HAW)
Unternehmen
SMA Solar Technology AG (SMA)
Forschungs- und Wissenschaftseinrichtungen
OFFIS - Institut für Informatik (OFFIS)

Kontakt Prof. Dr.-Ing. Sebastian Rohjans
Jade Hochschule
Ofener Str. 16/19
26121 Oldenburg
E-Mail: sebastian.rohjans@jade-hs.de

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



**FORSCHUNG AN
FACHHOCHSCHULEN**

Inhaltsverzeichnis

Abbildungsverzeichnis	1
1 Kurzbericht	3
1.1 Aufgabenstellung	3
1.2 Voraussetzungen	4
1.3 Planung, Ablauf und Ergebnisse	4
1.4 Artefakte und Publikationen	5
1.5 Zusammenarbeit mit anderen Stellen	5
2 Eingehende Darstellung	7
2.1 Projektbeschreibung und Motivation	7
2.1.1 Stand von Wissenschaft und Technik	8
2.1.2 Wissenschaftliche und technische Arbeitsziele und angestrebte Innovationen	9
2.1.3 Zeit- und Arbeitsplan	9
2.2 Ausführlicher Arbeitsbericht	12
2.2.1 Standardkonforme Anwendungsfälle	12
2.2.2 Datensatz	14
2.2.3 Modellierung	16
2.2.4 Simulation	23
2.3 Veröffentlichungen und Verbreitung von Projektergebnissen	26
2.4 Fazit	27
Literaturverzeichnis	29
Anhang: Inverter Failure Detection Use Case (Auszug)	33

Abbildungsverzeichnis

2.1	Gegenüberstellung des geplanten und tatsächlichen zeitlichen Ablaufs.	10
2.2	UML Diagramm aus einem Anwendungsfall	14
2.3	Verteilungen der IGBT Temperatur verschiedener baugleicher Wechselrichter.	15
2.4	Gruppierung (clustering) der Wechselrichter in Abhängigkeit von Luftfeuchtigkeit und Halbleitertemperatur.	16
2.5	Schematische Darstellung einer Anomalieerkennung.	17
2.6	Schematische Darstellung des Aufbaus des Zentral-Wechselrichters.	17
2.7	Vergleich vier verschiedener datengetriebener Modelltypen.	18
2.8	Ergebnis der Analyse des Einflusses einzelner Kenngrößen in Abhängigkeit von der Tageszeit mittels Permutation Feature Importance.	19
2.9	Ergebnis der Hyperparameteroptimierung eines der one-to-many MLP-Modelle.	20
2.10	Vergleich des Fehlers verschiedener Modelle und Modell-Generationen.	21
2.11	Analyse eines one-to-one Modells hinsichtlich der Modellqualität in Abhängigkeit von der Häufigkeit der Datenpunkte.	22
2.12	Implementierte Benutzeroberfläche zum Pausieren und Visualisieren der simulierten Szenarios.	24
2.13	Schematische Darstellung der Mosaik Co-Simulation.	25
2.14	Vergleich zwischen realen Messwerten und simulierten Ergebnissen eines fehlerhaften Wechselrichters.	25
2.15	Ergebnisse der Anomalieerkennung der simulierten Szenarios mit dem LSTM-Modell.	26

Kapitel 1

Kurzbericht

1.1 Aufgabenstellung

Im Projekt VORAUS PV wurde untersucht, ob und wie moderne Datenanalyseverfahren mittels maschinellem Lernen das Erkennen und Vorhersagen von Ausfällen von Photovoltaik (PV)-Wechselrichtern ermöglichen. Ziel des Projektes war es anhand der vorhandenen Daten eines Wechselrichterherstellers Algorithmen zu entwickeln, die in der Lage sind einzelne in der Zukunft liegende Ausfallereignisse vorherzusagen. Dies ermöglicht eine rechtzeitige vorausschauende Wartung, Reparatur oder einen Austausch von Wechselrichtern bevor es zu einem Ausfall kommt. Somit kann die Anzahl der Stunden, die eine PV-Anlage kontinuierlich Strom liefert, sowie ihre über die Lebensdauer berechnete Gesamtleistung und damit ihre Wirtschaftlichkeit erhöht werden. Das Ziel des VORAUS PV Projekts war, durch den Einsatz moderner Analyseverfahren, hierzu einen grundlegenden Beitrag zu leisten.

Das Vorhaben wurde hierzu von den Projektpartnern in drei Teilprojekte unterteilt: Die Datenanalyse, die Co-Simulation und der Hardware-in-the-loop (HIL)-Teststand. Dieser Teilprojektbericht befasst sich lediglich mit dem von der Jade Hochschule durchgeführten Teilprojekt der Co-Simulation.

Das Ziel des Teilprojekts Co-Simulation lag darin, eine simulative Umgebung zu schaffen, um die im Teilprojekt Datenanalyse von der HAW entwickelten Algorithmen zur Detektion von sich anbahnenden Ausfällen szenarienbasiert evaluieren zu können. Hierzu sollten zunächst Anwendungsfälle entwickelt werden, welche die Basis für die Simulation bilden. Anschließend sollten Modelle entwickelt, parametrisiert und über eine Co-Simulation zusammengeschlossen werden. Im letzten Schritt sollten die Methoden zur Anomalieerkennung ebenfalls an die Co-Simulation angeschlossen und durch verschiedene Szenarien evaluiert werden.

1.2 Voraussetzungen

Das Projekt VORAUS PV ermöglichte erstmals eine systematische Auswertung einer sehr großen Menge von nicht standardisierten Daten von Hochleistungs-PV-Wechselrichtern aus mehreren Jahren Betrieb in weltweitem Einsatz. Dieses Projekt unterscheidet sich sowohl durch seine Zielsetzung (Vorhersagen von einzelnen Wechselrichterausfällen), durch die Anwendung von aktuellen Datenanalyse-Methoden (maschinelles Lernen), durch den starken Anwendungsbezug (Aufgabenstellung aus der Perspektive des Wechselrichterherstellers) als auch durch den Umfang des Datensatzes von vorherigen Arbeiten. Der Kooperationspartner SMA Solar Technology AG (SMA) stellte hierzu einen umfangreichen Datensatz aus dem Betrieb von weltweit installierten Wechselrichtern zur Verfügung. Anders als übliche, oftmals an Gebäuden oder Privathaushalten angebrachte, Wechselrichter im Kilowatt-Bereich, handelte es sich bei den im Projekt untersuchten Wechselrichtern um Hochleistungs- oder auch Zentral-Wechselrichter, welche im Megawatt-Bereich agieren. Die Art der Daten und die Komplexität des Wechselrichters war somit ein weiteres Alleinstellungsmerkmal.

Durch den Wechsel von Prof. Dr.-Ing. Rohjans von der HAW zur Jade HS vor Beginn des Projekts war das Projekt nicht mehr im vollen Rahmen an der HAW ansässig, sondern das Teilprojekt Co-Simulation wechselte mit Prof. Dr.-Ing. Rohjans zur Jade HS. Da die Spezialisierung der Jade HS nicht im Bereich Energieinformatik zu verorten ist, wurde eine fundierte wissenschaftliche und akademische Unterstützung durch das OFFIS - Institut für Informatik sichergestellt. Das Forschungsinstitut OFFIS wurde 1991 als An-Institut der Carl von Ossietzky Universität Oldenburg gegründet, beschäftigt aktuell knapp 300 MitarbeiterInnen und erhält institutionelle Förderung durch das Land Niedersachsen. Der FuE-Bereich Energie erforscht und entwickelt neue Formen computergestützter Informationsverarbeitung in zukünftigen Smart Grids und kann auf große Expertise im Bereich Co-Simulation sowie maschinellem Lernen zurück blicken.

1.3 Planung, Ablauf und Ergebnisse

Im Rahmen des Projekts an der Jade Hochschule wurde der Teilbereich "Co-Simulation" von einem wissenschaftlichen Mitarbeiter bearbeitet, unterstützt durch eine studentische Hilfskraft. Die Zusammenarbeit und der Ergebnisaustausch wurden durch die Corona-Pandemie beeinträchtigt, was zu Verzögerungen führte. Das Projekt war ursprünglich in sechs Sprints unterteilt, die aufeinander aufbauten. Die Einarbeitung und geplante Aufenthalte bei SMA wurden aufgrund von Reiseeinschränkungen während der Corona-Pandemie abgesagt und konnten auch im weiteren Verlauf des Projekts nicht nachgeholt werden. Dies führte zu Verzögerungen beim Know-how-Aufbau. Die Projektlaufzeit wurde daher ein halbes Jahr verlängert.

Die IEC/PAS 62559 konformen Anwendungsfälle konnten wie geplant in Sprint 1 erstellt werden. Die Modellanalyse im zweiten Sprint zeigte, dass die angestrebte physikbasierte Modellierung der Wechselrichter aufgrund ihrer Komplexität nicht ohne dem Hersteller eigenem, proprietären Wissen durchführbar war. Stattdessen wurde nach Rücksprache mit dem Konsortium auf eine datengetriebene Modellierung umgeschwenkt.

Die weiteren Sprints umfassten die Implementierung der identifizierten Modelle, die Anbindung an die Co-Simulation, die Simulation von Szenarien und die Bewertung verschiedener Modelle. Aufgrund von sich herausstellenden Leistungsproblemen bei der Co-Simulation wurde die Reihen-

folge einiger Sprint-Inhalte angepasst und Analysen vorgezogen. Die ursprünglich für den sechsten Sprint geplante Dokumentation und Verwertung der Ergebnisse wurde sukzessive im Verlauf des Projekts durchgeführt, um relevante Ergebnisse zeitnah zu dokumentieren und zu veröffentlichen. Da sämtliche Inhalte des Sprints auf andere Sprints aufgeteilt wurde, wurde der sechste Sprint aufgelöst.

Aufgrund der Größe und Diversität des Datensatzes kam es bei der datengetriebenen Modellierung immer wieder zu Schwierigkeiten, welche jedoch im Rahmen eines Forschungsprojekts zu erwarten sind und auch entsprechend behoben werden konnten. So wurde zum Beispiel eine Methode entwickelt mit welcher sich ähnelnde Datenpunkte im Datensatz zu identifizieren lassen, um den Rechenaufwand zu verringern. Schlussendlich konnten sowohl die Modelle erfolgreich implementiert, als auch die Anomaliedetektion an die Co-Simulation angebunden werden. Die anschließend durchgeführte Simulation der Szenarien ließ jedoch kein klares Ergebnis bezüglich der Vorhersage möglicher Ausfälle der Wechselrichter zu.

1.4 Artefakte und Publikationen

Die wesentlichen Ergebnisse des Projektes wurden bereits veröffentlicht beziehungsweise auf Fachveranstaltungen vorgestellt und stehen somit einem breiten Fachpublikum zu Verfügung:

- J. Wibbeke, P. Teimourzadeh Baboli und S. Rohjans. „Optimal data reduction of training data in machine learning-based modelling: a multidimensional bin packing approach“. In: *Energies* 15.9 (2022), S. 3092
- J. Wibbeke, D. Alves und S. Rohjans. „Estimating time-delayed variables using transformer-based soft sensors“. In: *Energy Informatics* 6.Suppl 1 (2023), S. 16

Darüber hinaus konnte durch die Arbeit mit dem Co-Simulationsframe Mosaik¹ zu dessen Verbesserung beigetragen werden.

Im Laufe des Projekts konnte zudem eine studentische Abschlussarbeit mit dem Titel „Performance Vergleich von LSTM und Transformer Modellen: ein PV Anwendungsfall“ erfolgreich durchgeführt werden.

Der im Projekt beschäftigte wissenschaftliche Mitarbeiter konnte sein angestrebtes Promotionsvorhaben mit dem aktuellen Arbeitstitel „Soft sensors from imbalanced data for safety-critical regression applications“ in Kooperation zwischen der Jade Hochschule, dem OFFIS - Institut für Informatik und der Carl von Ossietzky Universität Oldenburg beginnen. Das Vorhaben wird auch nach Projektende fortgeführt, mit dem Ziel, Ende 2025 erfolgreich abgeschlossen zu werden.

1.5 Zusammenarbeit mit anderen Stellen

Über die gesamte Projektdauer wurden Ergebnisse und Erfahrungen im Konsortium mit der HAW und dem OFFIS ausgetauscht. Ein besonders weitgehender Austausch fand zwischen der Jade Hochschule und dem OFFIS statt, welcher unter anderem das Promotionsvorhaben in Kooperation mit

¹<https://mosaik.offis.de>

der Carl von Ossietzky Universität Oldenburg hervorbrachte. Zudem fand ein Austausch mit der Universität Kassel statt, welche sich im Projekt Digital Twin Solar ebenfalls mit der Modellierung von Zentralwechselrichtern auseinandersetzten. Digital Twin Solar wurde ebenfalls von SMA unterstützt, wodurch die Universität Kassel am selben Datensatz arbeitete. Die von SMA initiierte Kooperation zielte zunächst auf einen Austausch von Wissen und Ergebnissen ab. Im Laufe des Projekts wurden jedoch auch die verfolgten Modellierungsansätze miteinander abgesprochen um eine Doppelbearbeitung zu vermeiden. Als Ergebnis verfolge die Universität Kassel Ansätze des Transfer Learning und Active Learning, welche untersuchen in wie fern sich bereits fertig generierte Modelle auf neue Probleme übertragen lassen. Die JADE HS hingegen fokussierte sich auf das generelle erstellen und optimieren von Modellen.

Kapitel 2

Eingehende Darstellung

2.1 Projektbeschreibung und Motivation

Im Projekt VORAUS PV wurde untersucht, ob und wie moderne Datenanalyseverfahren (maschinelles Lernen) das Erkennen und Vorhersagen von Ausfällen von Photovoltaik (PV)-Wechselrichtern ermöglichen. Kann aus den Daten, die ein Wechselrichter aufnimmt und ständig in die Datenbank des Herstellers einspeist (Internet of Things), vorher erkannt werden, dass ein Ausfall bevorsteht, so kann der Wechselrichter rechtzeitig repariert oder ausgetauscht werden, bevor der tatsächliche Ausfall entsteht.

Die zu erwartenden Lebensdauer von PV-Anlagen und ihrer Komponenten bestimmt maßgeblich ihre Wirtschaftlichkeit und ist damit ein kritischer Faktor für ihren Einsatz [3]. Die meisten Ausfälle von PV-Anlagen werden durch Ausfälle von Wechselrichtern verursacht [4]. Die Ursachen für den Ausfall dieser Anlagen liegen typischerweise in kritischen Bauelementen, wie z.B. Elektrolytkondensatoren oder Halbleiterschaltern im Bereich der Leistungselektronik, im Zusammenspiel mit Umgebungsbedingungen oder spezifischem Nutzungsverhalten [4].

Ziel des Projektes war es anhand der vorhandenen Daten eines Wechselrichterherstellers Algorithmen zu entwickeln, die in der Lage sind einzelne Ausfallereignisse in der Zukunft vorherzusagen. Das Projekt VORAUS PV ermöglichte erstmals eine systematische Auswertung einer sehr großen Menge von nicht standardisierten Daten aus mehreren Jahren Betrieb unterschiedlicher PV-Wechselrichter in weltweitem Einsatz. Der Kooperationspartner, die Firma SMA Solar Technology AG (SMA), stellte dazu einen umfangreichen Datensatz aus dem Betrieb von weltweit installierten Wechselrichtern zur Verfügung. Die Firma SMA mit Sitz in Niestetal bei Kassel ist einer der weltweit führenden Hersteller für Wechselrichter für PV-Anlagen. SMA installiert und wartet die Anlagen und sammelt dabei weltweit kontinuierlich Daten über den Betrieb. Aufgabe dieses Projektes ist die Nutzung dieses Datensatzes zur Entwicklung und Evaluation von Vorhersagealgorithmen für Ausfälle in komplexen leistungselektronischen Systemen in der Photovoltaik.

Der Fokus liegt dabei auf der Analyse und Klassifizierung von Ausfallverhalten von PV-Wechselrichtern, um eine stabile Versorgungssicherheit in stark PV-durchdrungenen Stromnetzen zu garantieren. Der plötzliche Ausfall eines leistungsstarken Wechselrichters verursacht in der Regel signi-

fikante Störungen im angeschlossenen Verteilnetz z.B. durch Frequenzstörungen [5]. Durch Ausfallprognosen für leistungsstarke Wechselrichter kann der Hersteller dem Verteilnetzbetreiber einen zusätzlichen Service bieten.

Aus Sicht eines PV-Anlagen-Herstellers ermöglicht die Vorhersage von Ausfällen komplexer elektronischer Geräte in B2B und B2C Märkten neue Geschäftsmodelle, wie zum Beispiel die spezifische Verlängerung von Gewährleistung oder das Angebot neuer Services an Verteilnetzbetreiber. Weiterhin besteht die Möglichkeit der Optimierung von Gerätedesign oder -wartung. Zudem kann durch gezielten Einsatz sowohl die Kundenzufriedenheit, als auch die Unternehmensplanung verbessert werden. Vorhersagen von bevorstehenden Ausfällen ermöglichen eine rechtzeitige vorausschauende Wartung, Reparatur oder einen Austausch von Wechselrichtern. Somit kann die Anzahl der Stunden, die eine PV-Anlage kontinuierlich Strom liefert sowie ihre über die Lebensdauer berechnete Gesamtleistung und damit ihre Wirtschaftlichkeit erhöht werden. Das Ziel des VORAUS PV Projekts war, durch den Einsatz moderner Analyseverfahren, hierzu einen grundlegenden Beitrag leisten.

2.1.1 Stand von Wissenschaft und Technik

Das Projekt VORAUS PV hat sich an Ansätzen zum Condition Monitoring und Predictive Maintenance von leistungselektronischen Bauteilen orientiert. Während in vorherigen Arbeiten [6, 7] nur das datenbasierte, automatische Erkennen von Ausfällen für PV-Anlagen untersucht wurde, war in VORAUS PV die Entwicklung von Vorhersagealgorithmen für diese das Ziel. Da die Mehrheit aller Ausfälle von PV-Anlagen tatsächlich Ausfälle von Wechselrichtern sind [4], lag der Fokus des Projektes auf der Vorhersage eben jener. Häufig wurden die vorgeschlagenen Methoden zudem nur an wenigen, einzelnen (experimentellen) PV-Anlagen getestet und eine umfangreichere Evaluierung an einem größeren Datensatz fehlte. Eine weitere Besonderheit an dem Projekt war ferner die Art der untersuchten Zentral-Wechselrichter, welche aufgrund ihrer Größe und Leistungsumsetzung im Megawatt-Bereich deutlich komplexer sind, als hinlängliche handelsübliche Kilowatt-Wechselrichter.

Moderne Datenanalyseverfahren erlauben Vorhersagen von seltenen und extremen Ereignissen, durch das Erkennen von Vorläuferstrukturen in Datensätzen [8, 9, 10]. Diese Verfahren sind für die Vorhersagen von kritischen Verbindungen in Stromnetzen [10], starken Windböen in Windparks [9], El-Niño-Ereignissen [11], Molekülbewegungen [12], sowie plötzlichen Umschwüngen in Systemen mit nichtlinearer Dynamik [13] erfolgreich angewandt worden. Ebenfalls sind auch rekursive, künstliche neuronale Netzwerke (Reservoir Computing) [14] erfolgreich zur Vorhersage der Trajektorien von hochdimensionalen, chaotischen Systemen [15] im Einsatz.

Im Rahmen der Evaluation und bei der Entwicklung neuer Dienstleistungen nimmt die Co-Simulation eine zentrale Rolle ein. In [16] wird der aktuelle Stand von Co-Simulation für Energiesysteme sowie ein Ausblick auf künftige Herausforderungen beschrieben. Es sind bisher nur wenige Ansätze zur Kopplung von Co-Simulationsumgebungen mit Komponenten, die auf Konzepten des maschinellen Lernens basieren, untersucht worden. In [17] werden dazu z.B. Support Vector Machines verwendet und in [18] werden Surrogatenmodelle simulativ weiterentwickelt. In VORAUS PV verfolgte die Co-Simulation den Zweck, Daten für die Evaluation der Ausfallvorhersagealgorithmen zu generieren. Hierbei konnte auf Arbeiten aus dem Projekt „Smart Nord - Intelligente Netze in Norddeutschland“ aufgebaut werden [19]. Vor dem Projekt hat Co-Simulation bereits bei der

Analyse der Integration von Informations- und Kommunikationstechnik-Systemen (IKT-Systeme) und neuen Dienstleistungen in konventionellen Energiesystemen eine signifikante Rolle gespielt. Zum Beispiel werden in der Literatur Co-Simulationsteststände verwendet, um die Leistung von großflächigen Überwachungs- sowie von Steuer- und Schutzanwendungen in realen Szenarien für Energiesysteme und die Auswirkungen der unterstützenden IKT zu analysieren und zu quantifizieren [20, 21]. Aus der Marktperspektive hat die Co-Simulation die Tür für anspruchsvollere und zukunftsorientiertere Studien, wie zum Wärmeaustausch in Stromnetzen unter Berücksichtigung dessen thermohydraulischer Eigenschaften, geöffnet [22].

2.1.2 Wissenschaftliche und technische Arbeitsziele und angestrebte Innovationen

Das Projekt zeichnete sich durch einen starken Anwendungsbezug aus, so dass Ergebnisse mit hoher Relevanz für die Praxis erwartet wurden. Dazu war eine intensive Kooperation zwischen dem Kooperationspartner und den Antragsstellern geplant. Dies bezog sich sowohl auf das Erheben der Anforderungen, den Austausch von Daten als auch auf eine kontinuierliche Evaluation der Ergebnisse. Konkret wurden folgende Ergebnisse angestrebt:

1. Datenbasierte Bestimmung von Ausfallraten.
2. Algorithmen zur automatisierten Erkennung und Klassifizierung des Ausfalls von PV-Wechselrichtern gemäß vordefinierter Ausfallklassen.
3. Informationen über die zu erwartende Häufigkeit verschiedener Ausfälle.
4. Algorithmen zur Vorhersage von Ausfällen mit Hilfe von maschinellen Lernverfahren.
5. Evaluierung der Ausfallprognosen im Hinblick auf ihre Vorhersagegenauigkeit für verschiedene Wechselrichter und Schadensklassen.
6. Verständnis der durch die Datenanalyse identifizierten Zusammenhänge zwischen Ausfällen und internen sowie externen Einflussgrößen des Wechselrichters, wie zum Beispiel die Anbindung an das Verteilnetz, das Wetter und des Standorts durch Co-Simulationen.
7. Prognosen für die Auswirkungen von Ausfälle von PV-Wechselrichtern auf die Netzwerkkumgebung und auf Einspeiseprognosen.
8. Neue softwaregestützte Services, die PV-Hersteller den Netzbetreibern anbieten können.

2.1.3 Zeit- und Arbeitsplan

Im Rahmen des Projekts hat die Jade Hochschule einen wissenschaftlichen Mitarbeiter beschäftigt, der den Teilbereich „Co-Simulation“ des Projekts bearbeitet hat. Zudem wurde das Projekt zeitweise durch eine studentische Hilfskraft gestützt.

Durch die Corona-Pandemie kam es zu wesentlichen Arbeitseinschränkungen, welche auch die Zusammenarbeit und den Ergebnisaustausch im Projekt beeinflussten und erschwerten. Unter anderem konnten beispielsweise die geplanten Aufenthalte des wissenschaftlichen Mitarbeiters auf dem

SMA Werksgelände nicht stattfinden. Für die meisten geplanten Präsenzveranstaltungen wurden online Ersatzformate gefunden. Dennoch führten die Verzögerungen und sowie inhaltliche Herausforderungen und Anpassungen (siehe den weiteren Verlauf des Kapitels) zu einem Mehrbedarf an Bearbeitungszeit. Es wurde sich deshalb für eine Verlängerung der Projektlaufzeit um ein halbes Jahr eingesetzt, dessen auch stattgegeben wurde.

Die im Teilprojekt anfallenden Arbeiten wurde in sechs Sprints eingeteilt, welche aufeinander aufbauende Ziele verfolgten. Im Folgenden sind die ursprünglichen Ziele mit den jeweiligen Änderungen dargestellt. Eine Gegenüberstellung des geplanten zeitlichen Ablaufs und des tatsächlichen Ablaufs inklusive Verlängerung des Projekts ist in Abbildung 2.1 dargestellt.

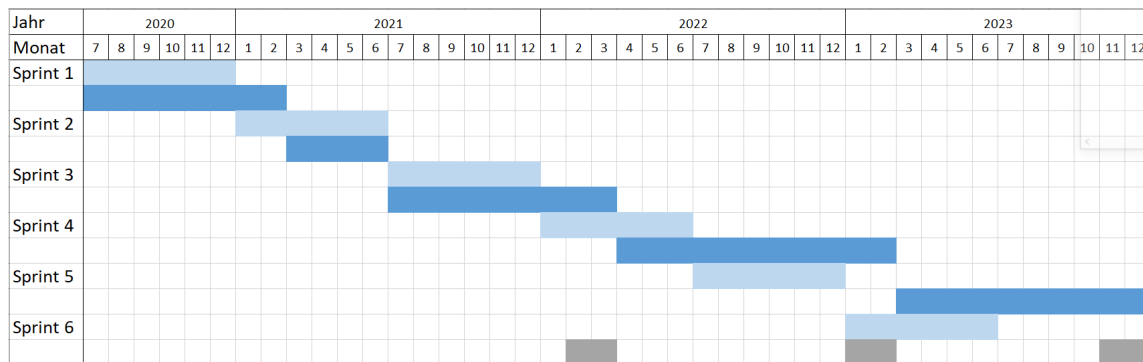


Abbildung 2.1: Gegenüberstellung des geplanten (hell) und tatsächlichen (dunkel) zeitlichen Ablaufs der Sprints. Die Inhalte des sechsten Sprints wurden auf andere Sprints aufgeteilt und sind nur noch zur besseren Nachvollziehbarkeit dargestellt (grau).

Sprint 1: Einarbeitung: Diese erste Arbeitsphase sollte in sehr enger Abstimmung mit SMA erfolgen. Geplant war unter anderem ein mehrwöchiger Aufenthalt des wissenschaftlichen Mitarbeiters bei SMA in Kassel, um dort Kontakte zu knüpfen und das nötige Know-How über den von SMA bereitgestellten Datensatz, Zentral-Wechselrichter und ihre Simulation zu erwerben.

Parallel dazu sollten Anwendungsfälle für neue Dienstleistungen auf Basis der Ausfallvorhersage entwickelt werden. Hierzu sollten basierend auf der Methodik aus dem IEC/PAS 62559 „Use Case Methodology“ standardkonforme Anwendungsfallbeschreibungen entwickelt werden, als Basis für die technische Systemspezifikationen der Simulationen.

Aufgrund der Corona-Pandemie musste die Einarbeitung des Mitarbeiters vor Ort bei SMA jedoch entfallen, was zu Verzögerungen beim Aufbau des nötigen Know-How führte.

Sprint 2: Modellanalyse: Auf Basis der Anwendungsfälle wurde in diesem Sprint abgeleitet, welche technischen Systeme und Modelle in der Simulation benötigt werden. Für nicht vorhandene Modelle wurde die Entwicklung geplant.

Im Laufe des Sprints hat sich heraus gestellt, dass die angedachte physik-basierte Modellierung nicht durchführbar ist. Als Grund hierfür ist insbesondere die große Komplexität der zu modellierenden Zentral-Wechselrichter an zu führen. Ohne Detailwissen über den Aufbau und die internen

Regelabläufe des Wechselrichters, war mit keiner ausreichenden Qualität der Simulation zu rechnen. Das benötigte Detailwissen konnte jedoch durch den pandemiebedingten Wegfall des Aufenthalts bei SMA nicht in der zur Verfügung stehenden Zeit aufgebaut werden. Zusätzlich wäre eine anschließende Parametrierung der Modelle nur unzureichend möglich gewesen, da zum Beispiel die internen Regelparameter den Betriebsgeheimnis von SMA unterliegen. Nach Absprache mit den Projektpartnern wurde deshalb entschieden statt der physik-basierten Modellierung eine datengetriebene Modellierung auf Basis des von SMA bereitgestellten Datensatzes zu verfolgen.

Sprint 3: Szenarienspezifikation In Sprint 3 wurden die im vorherigen Sprint identifizierten Modelle implementiert. Anschließend wurden Schnittstellen implementiert, um die Modelle mit der Co-Simulation zu verbinden. Zur Co-Simulation wurde der mosaik-Framework verwendet werden. Anschließend sollte die von der HAW entwickelte Komponente zur Ausfallvorhersage ebenfalls angebunden werden. Abschließend sollte eine Parametrierung aller verwendeten Modelle vorgenommen werden.

Aufgrund pandemiebedingter Umstände kam es auch im Teilprojekt der HAW zu Verzögerungen. Die Anbindung der Ausfallvorhersage wurde daher auf Sprint 4 verschoben. Da datengetriebene anstatt physik-basierte Modelle verwendet wurden, entfiel die Parametrierung der Modelle. An dessen Stelle wurde eine iterative Optimierung der verwendeten Modelle durchgeführt.

Sprint 4: Simulation I Nachdem alle Simulatoren inklusive der Modelle über die Co-Simulation gekoppelt sind, wurden in Sprint 4 erste Szenarien zur Simulation kreiert und simuliert. Insbesondere sollte hier untersucht werden inwiefern eine Vorhersage des Verlaufs der zu analysierenden Kenngrößen auch die Vorhersage des Ausfallereignisses beeinflussen könnte.

Bei der Simulation der ersten Szenarien stellten sich Leistungsprobleme der Co-Simulation heraus. Gegenüber des ursprünglichen Plans kam es daher in Sprint 4 in sofern zu Änderungen, dass die ursprünglich in Sprint 5 geplante Analyse des Echtzeitverhaltens der Co-Simulation in Sprint 4 vorgezogen wurde. Im Gegenzug wurde die Implementierung von Zeitreihen in Sprint 5 verschoben.

Sprint 5: Simulation II Auf Basis der vorherigen Ergebnisse wurden in Sprint 5 verschiedene Modelle und Ansätze miteinander verglichen und bewertet. Zusätzlich sollte analysiert werden, ob eine es möglich ist die Ausfallvorhersage in Echtzeit zu berechnen oder ob es nötig ist auf vorberechnete Daten zurück zu greifen.

Gegenüber des ursprünglichen Plans wurde die Analyse des Echtzeitverhaltens bereits in Sprint 4 durchgeführt (siehe oben). Im Gegenzug wurde die Implementierung von Zeitreihen in Sprint 5 durchgeführt. Zudem wurde dem sechsten Sprint vorgegriffen und die Kombination verschiedener Methoden getestet.

Sprint 6: Publikationen Im letzten Sprint des Projekts sollte getestet werden inwiefern eine Kombination der zuvor getesteten Methoden zu einer Verbesserung der Vorhersagequalität führen könnte. Zudem sollte in Sprint 6 die nötige Dokumentation und Veröffentlichung der Ergebnisse vorgenommen werden.

Im Zuge der Projektarbeit hat sich herausgestellt, das der sechste Sprint so wie er geplant war nicht zielführend ist. Anstatt die Ergebnisse am Ende des Projekts zu Dokumentieren und

Veröffentlichen, wurde dies sukzessive im Laufe des Projekts immer dann getan, wenn relevante Ergebnisse vorlagen. Der Sechste Sprint wurde daher aufgelöst und die Inhalte auf andere Sprints verteilt.

2.2 Ausführlicher Arbeitsbericht

Die unternommenen Arbeitsschritte im Teilprojekt lassen sich zielführend unter den Überschriften standardkonforme Anwendungsfälle, Datensatz, Modellierung und Simulation zusammen fassen. Im Folgenden wird auf die jeweiligen Einzelaspekte eingegangen und die relevantesten Ergebnisse präsentiert.

2.2.1 Standardkonforme Anwendungsfälle

Im Zuge der Entwicklung IEC/PAS 62559 konformer Anwendungsfälle wurden zunächst diverse mögliche Anwendungsfälle gesammelt und zusammengefasst. Nach Rücksprache mit den Projektpartnern und unter Berücksichtigung von, von SMA eingebrachter, Business-Cases entstanden 12 mögliche Anwendungsfälle, welche in Form von beschreibenden Einseitern ausgearbeitet wurden.

1. PV-Portfolio Zusammenstellung
2. Degradationsanalyse von PV-Modulen
3. Degradationsanalyse von PV-betriebenen Batteriespeichern
4. Kurzzeitvorhersage von Wechselrichterausfällen
5. Langzeitvorhersage von Wechselrichterausfällen
6. Erstellen von Wolkenkarten basierend auf PV-Anlagendaten
7. Zuverlässigkeitsanalyse von Wechselrichtern basierend auf statischen (meta) Daten
8. Automatisierte Standortanalyse basierend auf PV-Daten
9. Soft-Sensor Integration
10. Fehlererkennung durch Leistungsanalyse von Wechselrichter-Flottendaten
11. Integration großer PV-Anlagen in SCADA-Leitsysteme
12. Zustandsanalyse der Peripherie basierend auf Wechselrichterdaten

Das Ziel der Einseiter war es die Anwendungsfälle möglichst kurz und prägnant aber dennoch verständlich zum umreißen, um einen besseren thematischen Austausch und eine verbesserte Informationsweitergabe zu erreichen. Im Folgenden ist der Einseiter der PV-Portfolio Zusammenstellung dargestellt.

Name	Generation portfolio
Target in context	The user wants to place generation portfolios on the market. These portfolios consist of multiple units, aggregated in groups that assure predefined generation/demand profiles.
Actors	User, Weather Forecast, PV Model, Inverter Model, aggregation Module, demand profile, HMI
Trigger	User creates a request to the HMI
Requirements	<ul style="list-style-type: none"> • The inverter model and PV-model are adapted to their physical counterparts • All components are in normal operation • The weather forecast is correct
Essential steps (main success scenario)	<ul style="list-style-type: none"> • With the weather forecast data, a generation profile per inverter is calculated by the PV model • The generation profile is used in the inverter model to calculate the expected generation profile. • In the Aggregation Module, several expected generation profiles are aggregated in portfolios to meet predefined demand profiles. • The portfolios are provided to the user
Possible extensions	<ul style="list-style-type: none"> • The system also calculates the expected standard deviation and shows it to the user. • The system warns the user if a demand profile cannot be fulfilled. <ul style="list-style-type: none"> – the weather forecast or models are too inaccurate – generation is unable to meet demand

Im weiteren Verlauf wurde entschieden die drei Anwendungsfälle „Kurzzeitvorhersage von Wechselrichter ausfällen“, „Soft-Sensor Integration“ und „Zuverlässigkeitsanalyse von Wechselrichtern basierend auf statischen Daten“ weiter zu IEC/PAS 62559 standardkonformen Anwendungsfallbeschreibungen auszuarbeiten. Diese wurden anschließend den Projektpartnern zur Verfügung gestellt.

Ein beispielhaft aus der Beschreibung des Anwendungsfalls „Kurzzeitvorhersage von Wechselrichter ausfällen“ entnommenes UML Diagramm ist in Abbildung 2.2 dargestellt. Es beschreibt die Kommunikationsverbindungen zwischen den einzelnen Akteuren des Anwendungsfalls. Diese können dabei sowohl menschlicher als auch technischer Natur sein. Im Anhang: Inverter Failure Detection Use Case (Auszug) befindet sich zudem ein Auszug aus der Anwendungsfallbeschreibung nach IEC/PAS 62559.

2. Diagrams of Use Case

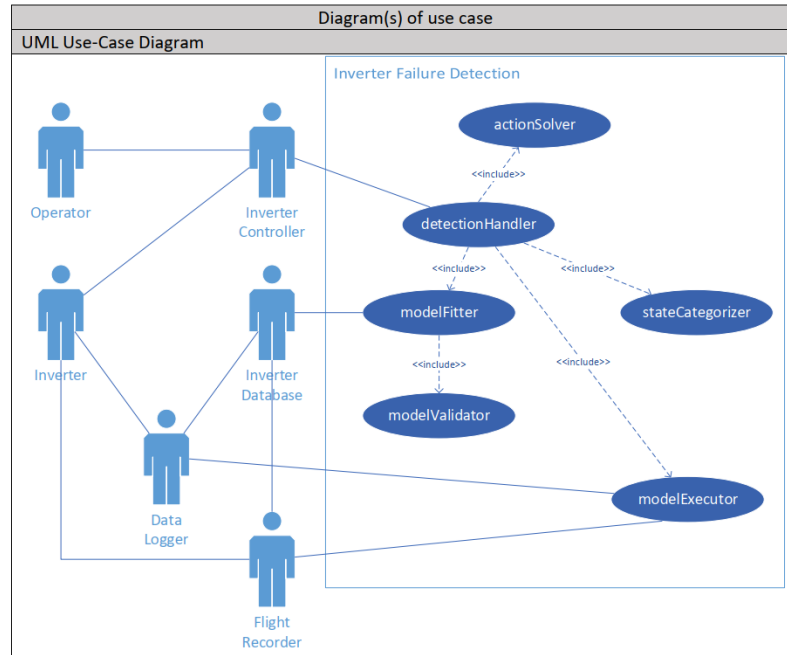


Abbildung 2.2: Entnommenes UML Diagramm aus der standardisieren Anwendungsfallbeschreibung des Anwendungsfalls „Kurzzeitvorhersage von Wechselrichterausfällen“.

Als mögliche Soft-Sensoren für den Anwendungsfall „Soft-Sensor Integration“ haben sich die Schätzung der Netzimpedanz, eine Inselnetzerkennung oder eine Temperaturschätzung des im Wechselrichter verbauten IGBT Halbleiters herausgestellt. Da die Temperaturschätzung ebenso für den „Kurzzeitvorhersage von Wechselrichterausfällen“-Anwendungsfall in Betracht kam, wurde nach Absprache mit den Projektpartnern beschlossen sich bei der Modellierung und Simulation auf die Berechnung der Temperatur der im Wechselrichter verbauten Halbleiter zur Leistungsumsetzung zu fokussieren. Gestützt wurde dies ebenfalls durch von SMA intern durchgeführte und anderweitige externe Forschungsergebnisse, welche die Halbleitertemperatur als validen Indikator für die Degradation des Wechselrichters nahe legen.

2.2.2 Datensatz

Um die Nutzung firmeninterner und der potenziell der Geheimhaltung unterliegender Daten zu ermöglichen, wurden zwischen den Projektpartnern und SMA eine Geheimhaltungsvereinbarung (non-disclosure-agreement, NDA) ausgearbeitet und unterzeichnet. Im Zuge dessen wurde von SMA ein umfangreicher Datensatz zur Verfügung gestellt. Dieser beinhaltete Sensordaten von 166 baugleichen Hochleistungswechselrichtern, verteilt über den Globus. Die ≈ 150 Kenngrößen eines jeden Wechselrichters wurden über mindestens 100 Tage aufgezeichnet. Bei 19 der Wechselrichter kam es

im Zeitraum zu einem Fehler, welcher zu einem zeitlich begrenzten Totalausfall führte. Insgesamt umfasste der Datensatz ungefähr 4 Milliarden Datenpunkte, unterteilt in circa 20000 Dateien.

Nach Erhalt des Datensatzes wurde zunächst eine Pipeline zum Einlesen und Verarbeiten der Dateien aufgesetzt. Anschließend wurden in Kooperation mit der HAW die Daten tiefgreifend analysiert. Hieraus ergab sich insbesondere, dass obwohl die Wechselrichter baugleich sind, sie nicht dem gleichen operativem Muster unterliegen. So verwendet z.B. nicht jeder Wechselrichter die gleichen Betriebspunkte (siehe Abbildung 2.3). Zudem ließ sich ein klarer Unterschied zwischen Tag und Nacht erkennen. Am Tag fährt der Wechselrichter spezifizier Betriebspunkte an, je nach Leistung. In der Nacht hingegen ist der Wechselrichter abgeschaltet und das Sensordaten-Profil wird im Wesentlichen von Umgebungsvariablen beeinflusst, wie der Außentemperatur.

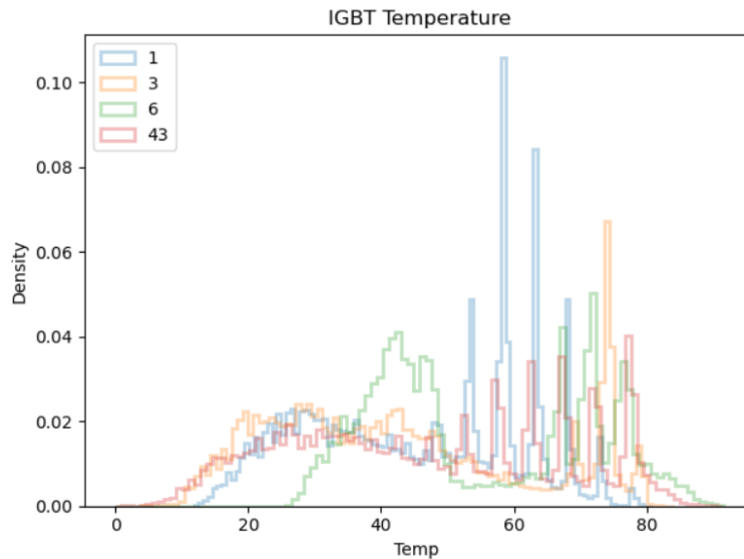


Abbildung 2.3: Verteilungen der IGBT Temperatur verschiedener baugleicher Wechselrichter.

Durch die Verwendung des von der HAW im Rahmen des Projekts aufgebauten Rechenclusters, konnte die HAW die festgestellten operativen Muster tiefgreifender Analysieren und so herausarbeiten, dass sich im Groben vier Gruppen feststellen lassen. Diese in Abbildung 2.4 dargestellten Gruppen stehen in Abhängigkeit von der relativen Luftfeuchtigkeit und der Temperatur des Wechselrichters. Zudem ist auffällig. Die Ergebnisse der Analyse wurde im Projekt geteilt, wodurch diese direkten Einfluss auf die angestrebte Modellierung hatten. Als Grund für die Bildung der Gruppen ist eine Aufteilung der Wechselrichter nach Klimazone am naheliegendsten. Da die Position des Wechselrichters jedoch ein Kundenspezifisches Datum ist, durfte uns SMA hierüber keine Auskunft geben. Ebenso wurde ein Rekonstruktion auf Basis der Sonneneinstrahlung durch den Geheimhaltungsvertrag untersagt, wodurch diese Informationen nicht genutzt werden konnten.

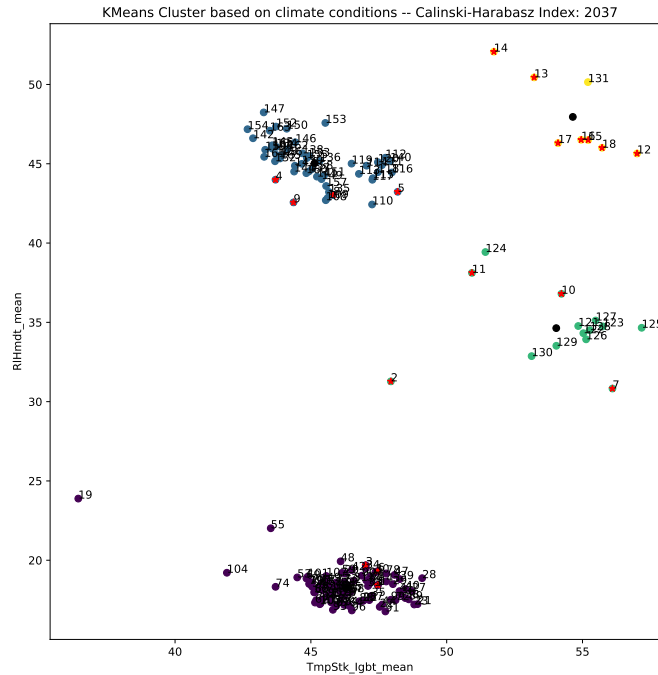


Abbildung 2.4: Gruppierung (clustering) der Wechselrichter in Abhängigkeit von Luftfeuchtigkeit und Halbleitertemperatur.

2.2.3 Modellierung

Zum Durchführen einer Anomalieerkennung wird neben den Messdaten des Systems ein Modell benötigt, mit welchem sich das Verhalten des Systems vergleichen lässt (siehe Abbildung 2.5). Dieses Modell muss das Verhalten des Systems in normalem/ungestörten und somit fehlerfreiem Zustand möglichst exakt widerspiegeln, andernfalls ist es der folgenden Anomalieerkennung nicht möglich ein fehlerhaftes Verhalten des Wechselrichters auch als solches zu erkennen.

Zur Abbildung des Systems wurden zunächst aus dem angestrebten Anwendungsfall „Kurzzeitvorhersage von Wechselrichterausfällen“ die zu modellierenden Systeme entnommen. Eine schematische Übersicht, welche den Aufbau der Wechselrichter und die im Datensatz befindlichen Sensoren zeigt ist in Abbildung 2.6 dargestellt. Im weiteren Verlauf ergab eine detaillierten Betrachtung des Systems unter Einbezug der Partner, dass die bisher angedachte physikalische Modellierung des Wechselrichters mit den vorliegenden Daten nicht möglich ist. Dies lag insbesondere an der Komplexität des Zentral-Wechselrichtersystems inklusive interner Regelung. Um anschließend Si-

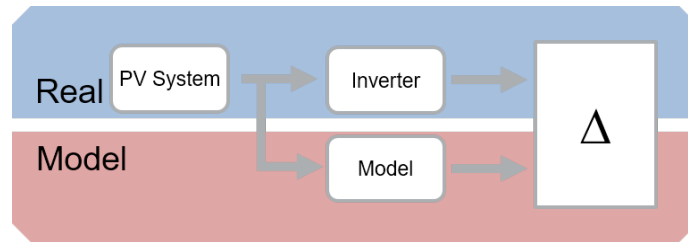


Abbildung 2.5: Schematische Darstellung einer Anomalieerkennung.

Simulationen zur Anomalieerkennung durchführen zu können, wäre eine exakte Parametrierung des Modells erforderlich, andernfalls könnte das Modell nicht das fehlerfreie Verhalten des Wechselrichters abbilden und es wäre mit keiner ausreichenden Qualität der Simulation zu rechnen. Es wurde daher im Konsortium entschieden den Fokus auf eine datengetriebene Modellierung mittels maschinellen Lernens zu legen.

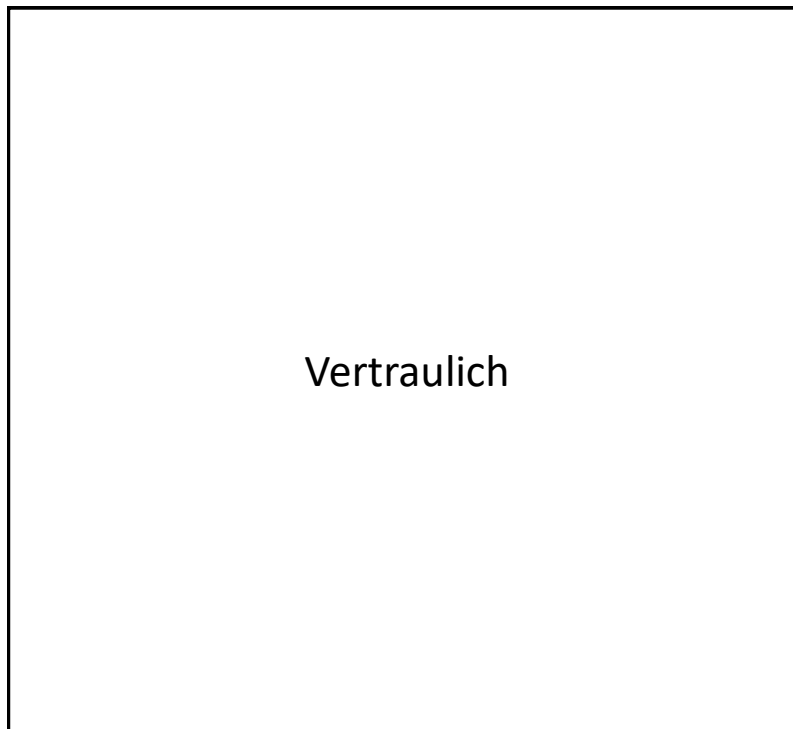


Abbildung 2.6: Schematische Darstellung des Aufbaus des Zentral-Wechselrichters.

In Folge dessen wurden Tests mit verschiedenen Modell-Typen und Teilen des von SMA bereitgestellten Datensatzes durchgeführt (siehe Abbildung 2.7). Hierbei erwiesen sich Modelle auf Basis von neuronalen Netzen als besonders performant. Wohingegen ein mehr statistischer Ansatz wie Prophet [23] sich nicht als geeignet herausstellte. Support Vector Modelle ergaben zwar akzeptable Ergebnisse, jedoch skalieren diese quadratisch bis kubisch mit der Anzahl der Sample; ein unterschiedener negativer Faktor bei der Größe des Datensatzes. Als Ergebnis aus den Tests wurde eine Umsetzung mit einem einfach Multilayer-Perzeptron (MLP) feedforward-Netz (im Folgenden nur noch MLP) und einem rekurrenten Netz in Form eines Long Short-Term Memory (LSTM)-Netzes (im Folgenden nur noch LSTM) beschlossen.

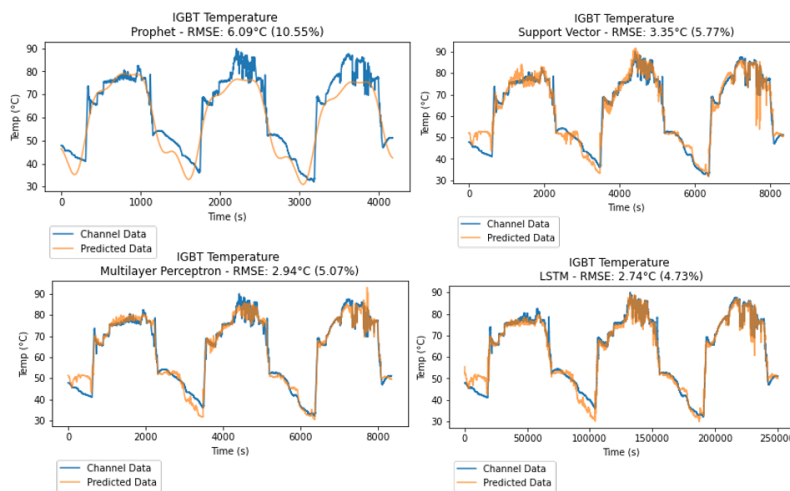


Abbildung 2.7: Vergleich vier verschiedener datengetriebener Modelltypen.

Die Größe des Datensatzes stellte sich dabei als eine besondere Herausforderung heraus. Um die Größe zu reduzieren wurde eine Methode entwickelt, welche es erlaubt ähnliche Datenpunkte zu identifizieren und zu reduzieren [1].

Eine weitere Herausforderung war die Diversität des Datensatzes, die unter anderem aus den verschiedenen Operationspunkten der Wechselrichter und der Ansässigkeit in verschiedenen Ländern hervorging. Es wurden daher sowohl ein one-to-one, als auch one-to-many und one-to-all Ansatz verfolgt. Alle Ansätze wurde im Laufe des Projekt mehrfach evaluiert und iterativ verbessert.

Im one-to-one Ansatz wurde für jeden Wechselrichter ein eigenes Modell trainiert. Die Erfolgsgarantie ist hierbei sehr hoch, jedoch ist der reale Nutzen gering, da zum Training der Anomalieerkennung der Wechselrichter bereits für eine lange Zeit Daten erzeugt haben muss. Wünschenswert wäre ein Modell, welches ab der Installation des Wechselrichtern funktioniert. Dies wurde mit dem one-to-many und dem one-to-all Ansatz verfolgt.

Für das one-to-many Modell wurden die Wechselrichter in Gruppen großer Ähnlichkeit unterteilt, zum Beispiel anhand des Temperaturprofils der Außentemperatur. Anschließend wurde für die Gruppe ein Modell trainiert, mit dem Kalkül, dass das Modell für neue/unbekannte Wechselrichter

der selben Gruppe ebenfalls gute Ergebnisse erzielt. Ist so der Ort eines neuen Wechselrichters bekannt, ließe sich über das Profil der Außentemperatur von Beginn an das richtige Modell verwenden und es müssten nicht erst Daten erfasst werden.

Im one-to-all Ansatz wurde zwei Modelle (MLP und LSTM) für alle Wechselrichter trainiert, unabhängig von etwaigen Ähnlichkeiten.

Sämtliche Modelle wurden mit pyTorch erstellt, einer auf Python basierenden öffentlichen Software Bibliothek für maschinelles Lernen [24]. Für die erste Modellgeneration wurden um den Rechenaufwand in Maßen zu halten lediglich die Kenngrößen verwendet, welche aufgrund des Aufbaus des Wechselrichters als physikalisch Sinnvoll erschienen. Zudem wurde nicht die Daten aller zur Verfügung stehenden fehlerfreien Wechselrichter verwendet.

Bei der Evaluation der ersten Generation der Modelle wurden, unabhängig vom Modell-Ansatz sehr große Fehler bei An- und Abschaltprozessen des Wechselrichters festgestellt. In Folge dessen wurde eine Analyse zur Wichtigkeit der verwendet Kenngrößen in Abhängigkeit von der Tageszeit durchgeführt. Hierbei stellte sich heraus, dass einzelne Kenngrößen besondere Wichtigkeit bei Schaltprozessen haben (siehe Abbildung 2.8). Verwendet wurde die Methode der Permutation Feature Importance.

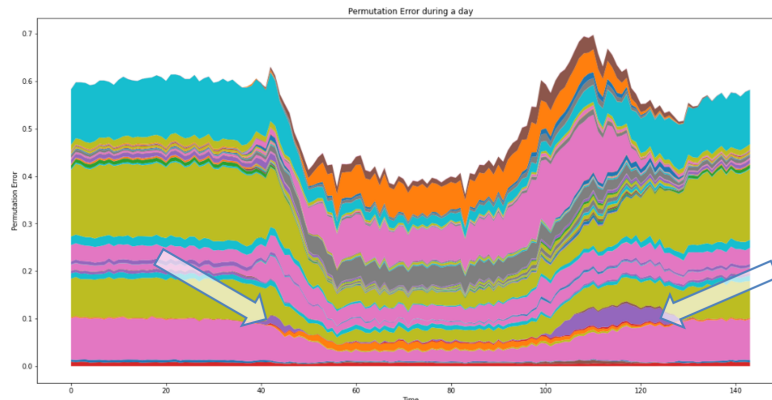


Abbildung 2.8: Ergebnis der Analyse des Einflusses einzelner Kenngrößen in Abhängigkeit von der Tageszeit mittels Permutation Feature Importance.

Für die nächste Iteration der Modelle wurde daher eine Kenngrößenauswahl (Feature Selection) durchgeführt um sämtliche wichtigen Kenngrößen mit einzubinden. Diese ergab unter anderem, dass eine Verwendung aller circa 150 Kenngrößen nicht zielführend ist, da sich eine vergleichbare Modellqualität unter Verwendung von lediglich den 30 relevantesten Kenngrößen erzielen lässt, bei geringerem Rechenaufwand. Hierbei musste jedoch auf physikalische Plausibilität geachtet werden, da nicht jede Kenngröße wirklich einen kausalen Zusammenhang bietet, auch wenn sie mathematisch korreliert.

Die Evaluation der zweiten Modellgeneration ergab eine generell gestiegene Modellqualität. Weiterhin erzielten jedoch die one-to-one Modelle bessere Ergebnisse als die one-to-many Modelle, welche wiederum die one-to-all Modelle übertrafen. Durch das gute Abschneiden der one-to-one

Modelle konnte jedoch bewiesen werden, dass es möglich ist sämtliche Wechselrichter durch Modelle abzubilden und die benötigten Daten hierfür vorhanden sind.

Für die dritte Modellgeneration wurde daher eine ausführliche Hyperparameteroptimierung durchgeführt. Hierfür wurde der State-of-the-Art Framework Raytune verwendet [25]. Optimiert wurden die Batch-Größe, Schichtgröße, Tiefe des Netzes und der Dropout. Die Optimierung ergab für die Modelle klare zu präferierende Hyperparameter, welche im Folgenden verwendet wurden. In Abbildung 2.9 ist so zum Beispiel zu sehen, dass Modelle mit einer Schichttiefe von 1 (*hidden_layer 1*) generell keine gute Modellqualität erzielen. Dies deutet ebenfalls auf die Komplexität des zu modellierenden Systems hin, da tiefere Modelle auch immer potentiell mehr Leistung bieten können. Andere Parameter hingegen, wie zum Beispiel die Batchsize hatten keinen nennenswerten Einfluss. Da die Batchsize jedoch großen Einfluss auf die Trainingsgeschwindigkeit des Modells hat, bestätigte uns die Analyse, dass wir große Batchsizes verwenden konnten um die Trainingszeit zu verringern.

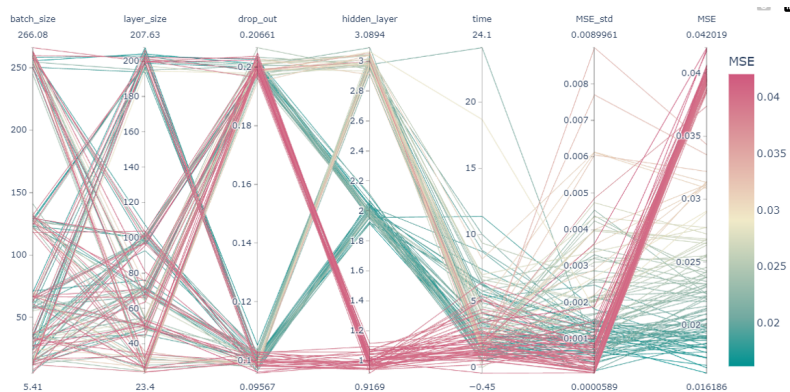
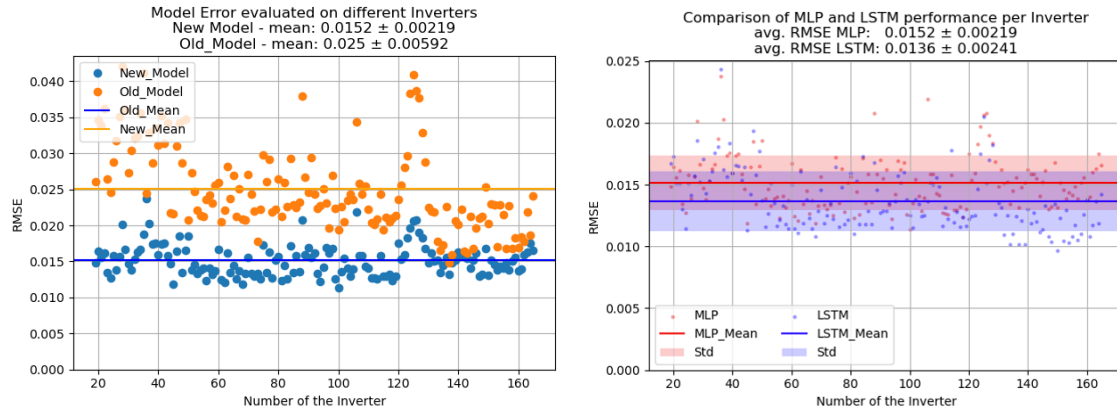


Abbildung 2.9: Ergebnis der Hyperparameteroptimierung eines der one-to-many MLP-Modelle.

Zusätzlich wurde der Sampling Ansatz für die beiden one-to-all Modelle überarbeitet. Aufgrund der Größe des Datensatzes wurden der Einfachheit halber die Daten bisher sequenziell pro Wechselrichter eingelesen. Die vorangegangene Evaluation ergab jedoch, dass die Modelle tendenziell für die zu Beginn des Trainings verwendeten Wechselrichter besser funktionierten als für die zu Ende des Trainings verwendeten Wechselrichter. Dies legte nahe, dass die Modelle während des Trainings zu einem lokalen anstatt globalem Minimum konvergierten.

Die vorgenommenen Änderungen führten zu einer deutlichen Verbesserung der Modellqualität. So verringerte sich zum Beispiel der Fehler des one-to-all Multilayer-Perzeptron (MLP) Modells der dritten Iteration um 40% im Vergleich zur zweiten Iteration (siehe Abbildung 2.10a). Auch verringerte sich, aufgrund der optimalen Hyperparameter und der damit einhergehenden angepassten Größe der Modelle, ebenfalls der Abstand zwischen den one-to-many und den one-to-all Modellen.

Ein Vergleich zwischen den MLP Modellen und Long Short-Term Memory (LSTM) Modellen zeigte zudem, dass es beiden Modelltypen möglich ist den Wechselrichter abzubilden, jedoch ist das LSTM ungefähr 10% genauer (siehe Abbildung 2.10b). Eine detailliertere Analyse zeigte dies-



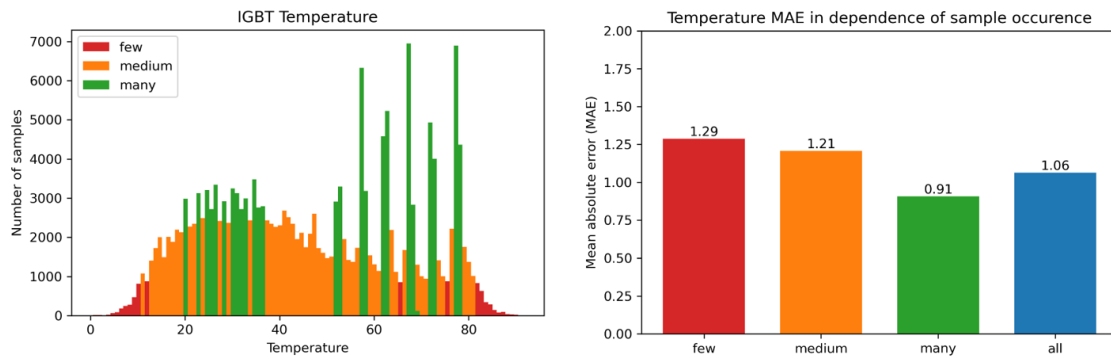
(a) Vergleich des Fehlers der MLP one-to-all Modelle der zweiten und dritten Modellgeneration. (b) Vergleich Fehlers der MLP und LSTM one-to-all Modelle.

Abbildung 2.10: Vergleich des Fehlers verschiedener Modelle und Modell-Generationen.

bezüglich, dass zwar im Großen und ganzen beide Modelltypen dem Verlauf der Wechselrichters folgen können, der Unterschied indes insbesondere zum Tragen kommt, wenn der Wechselrichter abrupten Schwankungen unterliegt. Dies kann zum Beispiel an einem wolkigen Tagen der Fall sein, sowie bei Schaltvorgängen. Die Einbußen des MLP sind hierbei insbesondere damit zu erklären, dass das MLP nur Daten des aktuellen Zeitpunkts als Eingangswerte bekommt, das LSTM hingegen mit Zeitreihen arbeitet. Die Temperatur des IGBT hat durch die Masse des Halbleiters ein inhärentes dynamisches thermisches Verhalten, wodurch die Temperatur der Leistung hinterherläuft. Ein solches Verhalten lässt sich durch Zeitreihen besser abbilden. Der Unterschied lässt sich ebenfalls bei den one-to-many und one-to-one Modellen beobachten und wird zudem unter Anderem in der zweiten Veröffentlichung diskutiert [2].

Des weiteren zeigt die Evaluation, dass die Modelle weiterhin für einige Wechselrichter nur lokale Minima erreichen und daher signifikant schlechtere Ergebnisse ausgeben als im Mittel. Dies könnten unter Umständen Wechselrichter sein, dessen operatives Profil alleinstehend ist und nicht mit anderen Profilen überein stimmt. Hieraus könnte sich eine Verzerrung (Bias) der Netze ergeben, resultierend aus dem Problem von ungleich-repräsentierten Daten („Imbalanced Data“). Dies rührt daher, dass es für das Modell von Vorteil ist sich auf häufig vorkommende Datenpunkte zu spezialisieren. Eine Verringerung des Fehlers eines X-fach vorkommenden Datenpunkts spielt ebenso X-fach in die Verringerung des gesamten Fehlers des Modell hinein. Anders herum hat die Vernachlässigung eines einzigartigen Datenpunkts ebenfalls auch nur 1-fachen Einfluss auf den Modellfehler. Im generellen gilt daher, dass zum Training datengetriebener Modelle alle verschiedenen relevanten Datenpunkte gleich-häufig vorkommen sollten. In der in Kapitel 2.2.2 dargestellten Gruppierung der Wechselrichter ist zu sehen, dass nicht alle Gruppen über gleich viele Wechselrichter verfügen, daher kann es zu einer Verzerrung kommen. Bei der Analyse der Data Imbalance konnte jedoch eine Verzerrung in Abhängigkeit von der gezeigten Gruppierung nicht bestätigt werden. Dies muss jedoch nicht heißen,

dass keine Verzerrung vorliegt, sondern lediglich, dass die Verzerrung ihre Ursache nicht in der ausgewählten Gruppierung hat. Vielmehr viel bei der Analyse nicht nur eine mögliche Verzerrung über die Wechselrichter-Profile hin auf, sondern viel mehr ebenfalls eine innerhalb eines jeden einzelnen Wechselrichters. Abbildung 2.11a zeigt die Verteilung der IGBT Temperatur eines Wechselrichters in Abhängigkeit von der Häufigkeit der Temperatur. Hierdurch lassen sich die Temperaturwerte in die Gruppen selten/durchschnittlich/häufig unterteilen. In Abbildung 2.11b ist zudem der Modellierungsfehler der Temperatur in Abhängigkeit der drei Gruppen dargestellt. Zu sehen ist, dass der Modellfehler, obwohl es sich um den selben Wechselrichter handelt, in Abhängigkeit von der Häufigkeit der Datenpunkte um bis zu 30% schwankt. Um dies zu umgehen, wurden in den vergangenen Jahren Methoden vorgestellt, welche z.B. auf dem over-/undersampling [26] oder auf der Gewichtung [27] von Datenpunkten basieren. Beide weisen jedoch in so fern Lücken auf, dass in der Methode rein die Zielwerte des Modells (in unserem Fall die Temperatur) betrachtet werden, jedoch kein kausaler Zusammenhang mit den Ursachen für die etwaige Temperatursteigerung hergestellt wird. Da die Wechselrichter einer solchen Größe, wie sie im Projekt modelliert werden, jedoch zu den kritischen Infrastrukturen (KRITIS) für die Energieerzeugung zu zählen sind, ist eine Berücksichtigung der Ursachen unserer Ansicht nach essentiell. Aufgrund der Komplexität konnte dieses Problem im Rahmen des Projektes nicht weiter voran getrieben werden, findet jedoch im angestrebten Promotionsvorhaben Anklang.



(a) Verteilung der IGBT Temperatur in Abhängigkeit von der Häufigkeit der Datenpunkte. (b) Fehler der modellierten IGBT Temperatur in Abhängigkeit der Häufigkeit der Datenpunkte.

Abbildung 2.11: Analyse eines one-to-one Modells hinsichtlich der Modellqualität in Abhängigkeit von der Häufigkeit der Datenpunkte.

Gegenstand der zweiten Veröffentlichung ist ebenfalls die Verwendung von Transformer-basierten neuronalen Netzen zur Modellierung von Kenngrößen mit dynamischen Verhalten.

Transformer sind eine erst 2017 vorgestellte neue Art Neuronaler Netze, welche Ursprünglich zur Verarbeitung natürlicher Sprache (NLP) entwickelt wurden. Im NLP Bereich bieten Transformer massive Verbesserungen gegenüber der zuvor verwendeten LSTMs-Modellen. Während der Laufzeit des Projekts erweiterte sich das Anwendungsspektrum über die Verarbeitung natürlicher Sprache

hinaus und es erschienen erste Architekturen welche Regressionsmodelle behandelten. Da die meisten Veröffentlichungen jedoch auf die Vorhersage langer Zeitreihen fußen, wurde beschlossen die Anwendbarkeit zur Modellierung dynamischer Systeme zu untersuchen. Entgegen der anfänglich angenommenen Vermutung stellte sich der Time-Series-Transformer [28] in der Gegenüberstellung jedoch als weniger performant als MLP und LSTM heraus. Die Ergebnisse wurden auf der DACH Energy Informatics Conference 2023 vorgestellt [2]. Für die auf die Modellierung folgende Simulation wurde daher aufgrund der Ergebnisse der Veröffentlichung auf eine Verwendung des Time-Series-Transformers verzichtet.

2.2.4 Simulation

Das Ziel der Simulation war es zu evaluieren ob die Ausfälle von Wechselrichtern in Testszenarien vorgesagt werden können, um eine Beschädigung des Wechselrichters zu verhindern oder, falls dies schon eingetreten ist, ein Kaskadieren des Fehlers zu verhindern. Zur Simulation wurden die von der Jade Hochschule entwickelten Modelle über eine Co-Simulation mit der Anomalieerkennung der HAW verbunden. Als Co-Simulations-Framework wurde Mosaik [29] verwendet.

Insgesamt verfügt der Wechselrichter über drei Leistungsbrücken, welche wiederum jeweils drei Halbleiter zur Wechselrichtung des Stroms haben (einen für jede Phase). Jeder der Halbleiter ist mit zwei Temperatursensoren ausgestattet, unterhalb und oberhalb des Halbleiters (siehe Abbildung 2.6). Insgesamt sind auf diese Weise 18 Temperatursensoren pro Wechselrichter verbaut dessen Werte ausschlaggebend für eine Fehlererkennung sein können.

Als Datenquellen für die Szenarios wurden die Daten der 19 fehlerbehafteten Wechselrichter des Datensatzes verwendet. Der Simulationszeitraum bezog sich dabei jeweils auf 60 Tage vor Auftreten und 30 Tage nach Auftreten des Fehlers. Um einen Vergleich zu nicht-fehlerbehafteten Wechselrichtern zu erhalten wurden zudem 19 weitere fehlerfreie Wechselrichter zufällig ausgewählt. Zusätzlich wurde im Mosaik Szenario ein Datensammler implementiert, welcher die Daten aller Modelle aggregiert, visualisiert und speichert. Aufgrund der Länge (Anzahl der Zeitschritte) der Szenarios war es von Nöten diese zwischenzeitlich pausieren zu können. Da vorab existierende Visualisierungen keine Möglichkeit baten auf das Szenario zurück zu wirken, wurde eine eigene Visualisierung implementiert in Form einer auf React¹ basierendem web-basierenden Benutzeroberfläche. Diese ist in Abbildung 2.12 dargestellt und ermöglichte sowohl das Starten/Stoppen/Pausen der Simulation als auch das Skalieren der Auflösung um Ergebnisse besser erkennen zu können.

Nach der Implementation eines ersten Testszenarios mit allen 18 Temperaturwerten stellten sich Defizite bezüglich der Simulationsgeschwindigkeit heraus. So erzielte die Simulation mit ≈ 200 ms pro Simulationsschritt zwar die ursprünglich angestrebte Echtzeitfähigkeit, bei einem Simulationszeitraum von 90 Tagen mussten die Anforderungen jedoch auf 10 bis 100-fache Echtzeit angehoben werden. Da Mosaik ursprünglich für die Simulation von Stromnetzen mit hunderten von Simulatoren gedacht ist, hätte dies auch möglich sein sollen. Eine tiefgreifendere Analyse führe jedoch zu einem ein Fehler in Kombination von Mosaik und PyTorch dazu, dass die Modelle für jeden Zeitschritt neu geladen wurden, was die Rechenzeit erhöhte. Der Fehler wurde an das Entwicklerteam von Mosaik weitergegeben, konnte jedoch für die Berechnung der Ergebnisse nicht mehr rechtzeitig

¹<https://react.dev>

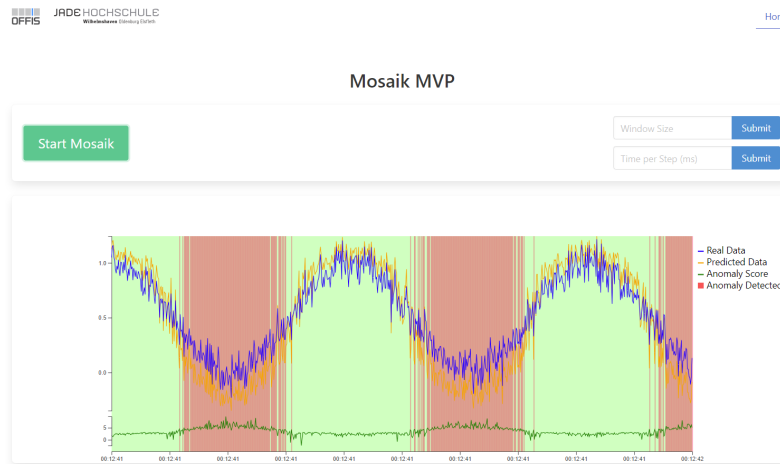


Abbildung 2.12: Implementierte Benutzeroberfläche zum Pausieren und Visualisieren der simulierten Szenarios.

behalten werden. Um die Simulation trotzdem ausführen zu können wurden zweierlei Umgehungen umgesetzt:

- Die Anzahl der Simulatoren wurde von 18 auf 3 verringert (ein Temperatursensor pro Brücke). Diese Annahme ist in sofern valide, da der Aufbau der Sensoren auf den Halbleitern baugleich ist. Um eine Aussage über den ganzen Wechselrichter zu tätigen, müsste die Simulation somit lediglich mehrfach wiederholt werden.
- Die Abtastrate wurde von 1 Hz auf 0.1 Hz verringert. Dies hat in sofern keinen Einfluss auf die Anomalieerkennung, da sich die Temperatur aufgrund der thermischen Kapazität des Halbleiter nur sehr träge ändert. Eine am Wechselrichter anliegende Spannungsspitze hätte somit auf die Temperatur nur einen geringfügigen und geglätteten Einfluss.

Der letztendliche Aufbau der Co-Simulation ist schematisch in Abbildung 2.13 dargestellt.

Insgesamt wurden pro Modell (MLP und LSTM) 38 Szenarien (19 fehlerhafte und 19 fehlerfreie) simuliert. Die Ergebnisse eines einzelnen beispielhaften Szenarios ist in Abbildung 2.14 dargestellt. Zu sehen ist der Verlauf der realen und der modellierten Halbleitertemperatur des Wechselrichters (oben) und die Differenz aus realer und modellierter Temperatur (unten). An dem Verlauf der Temperatur lässt sich ebenfalls erkennen, dass nach Auftreten des Fehlers der Wechselrichter für einige Zeit abgeschaltet war. Auffällig ist, dass vor Auftreten des Fehlers die Abweichung geringer wird, was jedoch auch an äußeren Umständen liegen kann und nicht unbedingt mit dem Ausfall zusammenhängen muss. Die durchschnittliche Differenz der modellierten Temperatur für den gezeigten Wechselrichter liegt bei 1,78°C. Aus der unteren Abbildung lässt sich ebenfalls ablesen, dass trotz der im allgemeinen sehr geringen Abweichung selbst bei Verwendung eines LSTMs immer wieder zu starken Spitzen im Modellfehler kommt. Das Ergebnis der Anomalieerkennung für den selben

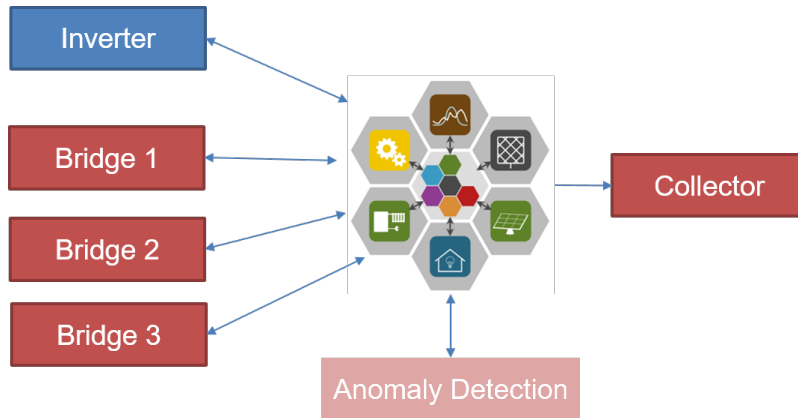


Abbildung 2.13: Schematische Darstellung der Mosaik Co-Simulation.

Wechselrichter ist in Abbildung 2.15a dargestellt. Die Verringerung der Abweichung in den circa 20 Tagen bevor der Fehler auftritt spiegelt sich auch hier in einem flacheren Verlauf der Linie wieder. Auffällig ist, dass nach Auftreten des Fehlers deutlich mehr Anomalien festgestellt wurden. Dies ist dadurch zu begründen, dass sowohl Modell als auch Detektion nur mit in Betrieb befindlichen Wechselrichtern trainiert wurden.

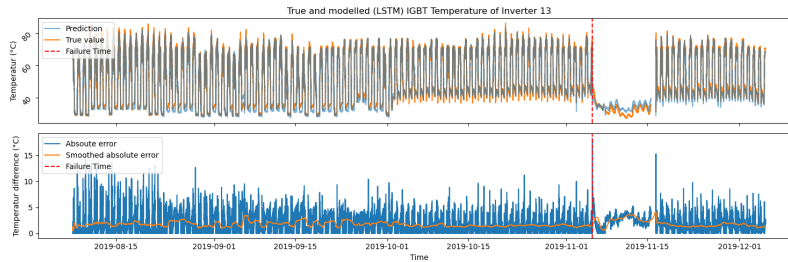
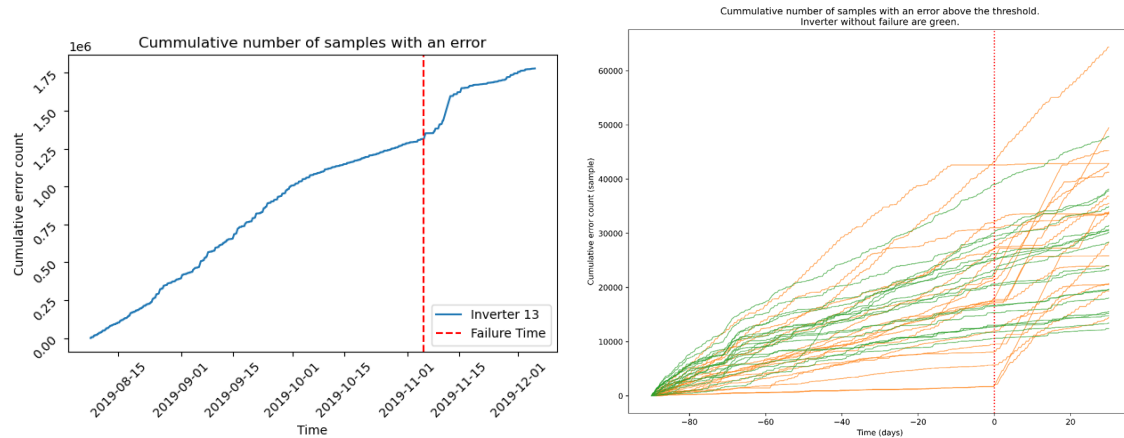


Abbildung 2.14: Vergleich zwischen realen Messwerten und simulierten Ergebnissen eines fehlerhaften Wechselrichters. Verlauf der simulieren und realen Temperatur (oben). Differenz zwischen simulierter und realer Temperatur mit 24h-Glättung (unten).

Die Auswertung der Anomalieerkennung für alle 38 Szenarien des LSTMs ist in Abbildung 2.15b dargestellt. Auffällig ist, dass die fehlerhaften Wechselrichter eine breitere Streuung der Anzahl der als Anomalie eingestufteten Datenpunkte haben. So gibt es einige fehlerhafte Wechselrichter die bis zum Auftreten des Fehlers nahezu gar keine Anomalien aufweisen, andere wiederum detektieren sehr häufig Anomalien. Insgesamt lässt sich jedoch kein signifikanter Unterschied zwischen fehlerhaften und fehlerfreien Wechselrichtern feststellen, welcher zur Vorhersage von Ausfällen genutzt werden könnte.

Die Ergebnisse des MLP-Modells sind nahezu vergleichbar, mit dem Unterschied, dass das MLP-



(a) Ergebnisse eines einzelnen fehlerhaften Wechselrichters. (b) Vergleich der Ergebnisse zwischen fehlerbehafteten (orange) und fehlerfreien (grün) Wechselrichtern.

Abbildung 2.15: Ergebnisse der Anomalieerkennung der simulierten Szenarios mit dem LSTM-Modell. Kumulative Anzahl der Sample welche von der Anomalieerkennung als fehlerhaft erkannt werden.

Modell etwas ungenauer ist als das LSTM. Auf die Darstellung der Ergebnisse des MLP-Modells wurde daher an dieser Stelle verzichtet.

2.3 Veröffentlichungen und Verbreitung von Projektergebnissen

Die wesentlichen Ergebnisse des Projektes wurden bereits veröffentlicht beziehungsweise auf Fachveranstaltungen vorgestellt und stehen somit einem breiten Fachpublikum zu Verfügung:

- J. Wibbeke, P. Teimourzadeh Baboli und S. Rohjans. „Optimal data reduction of training data in machine learning-based modelling: a multidimensional bin packing approach“. In: *Energies* 15.9 (2022), S. 3092
- J. Wibbeke, D. Alves und S. Rohjans. „Estimating time-delayed variables using transformer-based soft sensors“. In: *Energy Informatics* 6.Suppl 1 (2023), S. 16

Die Ergebnisse tragen zur weiteren Stärkung und Profilierung der themenrelevanten Forschungsgebiete an der Jade Hochschule bei. Darüber hinaus wurden die aus der Analyse der Co-Simulation ergangenen Leistungsprobleme des mosaik-Frameworks in Verbindung mit PyTorch an das Entwicklerteam weiter gegeben und beseitigt. Der Framework wird vom OFFIS entwickelt. Durch die Weitergabe des Fehlers wurde sicher gestellt ist, dass sich die Qualität das Frameworks erhöht und zukünftige Forschende nicht auf vergleichbare Probleme stoßen.

Das gestartete Promotionsvorhaben des beschäftigten wissenschaftlichen Mitarbeiters mit dem aktuellen Arbeitstitel „Soft sensors from imbalanced data for safety-critical regression applications“ wird über das Projektende hinaus in Kooperation zwischen der Jade Hochschule, dem OFFIS - Institut für Informatik und der Carl von Ossietzky Universität Oldenburg fortgesetzt. Derzeitig befinden sich zwei weitere Publikationen in Arbeit, welche unter anderen das in Kapitel 2.2.3 identifizierte Thema ungleich verteilter Daten (Imbalanced Data) aufgreifen. Die Veröffentlichungen sind im Laufe des Jahres 2024 geplant. Das Einreichen der Dissertation ist für 2026 angesetzt.

Im Bereich der Ausbildung von Studierenden, bieten die Inhalte des Forschungsprojektes die Möglichkeit zukünftige Lehrveranstaltungen mit entsprechenden praktischen und greifbaren Beispielen anzureichern. Der Themenbereich „Imbalanced Data“ wird zudem in zwei aus dem Projekt heraus motivierten und gegenwärtig andauernden studentischen Abschlussarbeiten bearbeitet. Die gewonnene Expertise fließt darüber hinaus in andere Forschungsprojekte ein.

In der Projektlaufzeit wurden zudem vier potentielle studentische Abschlussarbeiten erarbeitet und ausgeschrieben. Eine Arbeiten mit dem Titel „Performance Vergleich von LSTM und Transformer Modellen: ein PV Anwendungsfall“ wurde Erfolgreich durchgeführt. Die Ergebnisse flossen in die oben genannte Veröffentlichung [2] ein.

2.4 Fazit

Pandemie-bedingt kam es von Projektbeginn an bis in 2022 hinein immer wieder zu Einschränkungen, welche die Arbeit im Projekt erschwerten. So war es zum Beispiel aus Gründen des Arbeitsschutzes untersagt Dienstreisen zu tätigen, wodurch sämtliche bis dahin geplante Präsenztreffen entfallen mussten. Dies hatte großen Einfluss auf das Projekt, da die geplanten mehrwöchigen Aufenthalte auf dem SMA Werksgelände in Kassel nicht stattfinden konnten. Da drüber hinaus gestaltete sich die Zusammenarbeit im Projekt über online Ersatzformate jedoch als gut. Positiv hervorzuheben ist ebenfalls der Austausch mit dem Projekt Digital Twin Solar.

Um den von SMA bereitgestellten Datensatz erhalten und nutzen zu können, musste zum Schutz kundenbezogener Daten ein Geheimhaltungsvertrag ausgearbeitet und unterzeichnet werden. Der Schutz der Kundendaten erforderte zudem zusätzliche Maßnahmen um die Daten vor Fremdzugriff zu schützen. Leider kam es durch den Geheimhaltungsvertrag ebenfalls zu Einschränkungen was die Auswertbarkeit der Daten betrifft. Dies war zwar bedauerlich, jedoch ist ein adäquater Datenschutz als hohes Gut anzusehen und oftmals Grundlage dafür, dass Daten in diesem Umfang überhaupt erst geteilt werden können. Auf der positiven Seite steht jedoch, dass der so geteilte Datensatz, bestehend aus Daten von Zentralwechselrichtern, durchaus einzigartig und wissenschaftlich hoch interessant ist. Aufgrund ihrer systemischen Komplexität sind die Eigenschaften der Zentralwechselrichter nur eingeschränkt auf handelsübliche Wechselrichter übertragbar, was zu neuen, der Literatur nicht entnehmbaren Anforderungen zum Modellieren von Wechselrichtern führt. Gleichzeitig war es aufgrund des großen Umfangs des Datensatzes nicht nötig synthetische Daten zu erzeugen, wodurch eine große Praxisnähe der angewendeten Ansätze sicher gestellt werden konnte.

Auch wenn das Gesamtvorhaben grundsätzlich recht klar definiert war, kam es im Laufe des Projektes immer wieder zu Anpassungen des Arbeitsplans. So wurde z.B. der sechste Sprint aufgelöst und die Inhalte auf die anderen Sprints verteilt. Da drüber hinaus wäre es durchaus möglich gewesen

für einzelne Ziele, wie zum Beispiel die Ausarbeitung der Anwendungsfälle, mehr Zeit in Anspruch zu nehmen. Die Anwendungsfälle hätten in Komplexität und Detailgrad beliebig ausgestaltet werden können. Im Anbetracht des Gesamtziels mussten hier jedoch gezielt Einschnitte hingenommen werden, um die Modellierung nicht aus den Augen zu verlieren. Insgesamt war der Aufwand jedoch gut bemessen. Hervorzuheben ist zudem insbesondere die Flexibilität und Bereitschaft der Partner und beteiligten Mitarbeiter den Fokus von einer physik-basierten Modellierung auf datengetriebene Modellierung zu ändern, aufgrund von besserer Realisierbarkeit. Durch die Verlängerung des Projekts um sechs Monate war es zudem möglich trotz der anfänglichen Verzögerungen alle geplanten Arbeiten durch zu führen.

Das Ziel er Jade HS, das Normalverhalten eines Wechselrichters möglichst detailgetreu zu Modellieren, konnte erreicht werden, auch wenn die Co-Simulation eine Vorhersage des Ausfalls des Wechselrichters leider nicht kohärent treffen konnte.

Das angesetzte Ziel von zwei Publikationen konnte von der Jade HS in sofern erreicht werden, dass zwar die gezielte Anzahl der Veröffentlichungen erreicht wurde, diese jedoch nicht wie ursprünglich geplant in der Domäne der Co-Simulation zu verorten sind, sondern im Bereich Data Science und in der datengetriebenen Modellierung. Positiv hervorzuheben ist zudem, dass der 2022 veröffentlichte Artikel „Optimal data reduction of training data in machine learning-based modeling: a multidimensional bin packing approach“ bereits erste Resonanz in Form von Zitierungen erfährt. Bezüglich des zweiten Artikels „Estimating time-delayed variables using transformer-based soft sensor“ ist dies noch nicht der Fall, jedoch wurde dieser auch erst Ende 2023 veröffentlicht. Aufgrund des positiven Feedbacks nach der Vorstellung auf der Energy Informatics Conference rechnen wir jedoch ebenfalls mit guter Resonanz.

Da drüber hinaus konnte das Projekt erfolgreich zu Qualifizierung wissenschaftlichen Personals genutzt werden. So konnte zum Beispiel im Rahmen des Projekts eine Bachelorarbeit verfasst werden und der Bachelorand nach Abschluss der Arbeit als wissenschaftliche Hilfskraft erste Erfahrungen bei der Mitarbeit an einer wissenschaftlichen Publikation sammeln. Durch den Besuch und die Publikation auf der DACH Energy Informatics Conference konnte der wissenschaftliche Mitarbeiter zudem Einblicke in den aktuellen Diskurs auf einer in der Energieinformatik renommierten Fachkonferenz erhalten.

Durch die Kooperation mit dem OFFIS stellte sich im Laufe des Projekts heraus, dass viele der angewendeten Ansätze eine große Übertragbarkeit auf andere Forschungsprojekte bieten. So konnte zum Beispiel die Abhängigkeit der Modelqualität von der Häufigkeit der Datenpunkte (Imbalanced Data) ebenfalls bei der Modellierung von Wasserstoff-Elektrolyseuren nachgewiesen werden.

Das Ziel der Förderung eine Promotion im Projekt konnte in sofern erreicht werden, dass ein passendes und wissenschaftlich interessantes Promotionsthema gefunden werden konnte. Der, im Projektantrag angedachte, Abschluss der Promotion innerhalb eines Jahres nach Projektende wird jedoch verfehlt werden. Ein Einreichen der Dissertation ist für 2026 geplant. Das Dissertationsvorhaben wird in einer Kooperation aus Jade HS, dem OFFIS und der Carl von Ossietzky Universität Oldenburg weiter vorran getrieben. Für das laufende Jahr 2024 sind zwei weitere Publikationen im Bereich Data Science angestrebt, in denen das Imbalanced Data Problem aufgegriffen und behandelt wird.

Insgesamt war VORAUS PV aus Sicht der Jade HS trotz der genannten Widrigkeiten ein erfolgreiches Projekt, welches einen guten Beitrag zum wissenschaftlichen Diskurs und Qualifizierung

wissenschaftlichen Personals leisten konnte.

Literaturverzeichnis

- [1] J. Wibbeke, P. Teimourzadeh Baboli und S. Rohjans. „Optimal data reduction of training data in machine learning-based modelling: a multidimensional bin packing approach“. In: *Energies* 15.9 (2022), S. 3092.
- [2] J. Wibbeke, D. Alves und S. Rohjans. „Estimating time-delayed variables using transformer-based soft sensors“. In: *Energy Informatics* 6.Suppl 1 (2023), S. 16.
- [3] T. J. Formica, H. A. Khan und M. G. Pecht. „The effect of inverter failures on the return on investment of solar photovoltaic systems“. In: *IEEE Access* 5 (2017), S. 21336–21343.
- [4] G. T. Klise, O. Lavrova und R. L. Gooding. *PV System Component Fault and Failure Compilation and Analysis*. Techn. Ber. Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2018.
- [5] J. Von Appen, M. Braun, T. Stetz, K. Diwold und D. Geibel. „Time in the sun: the challenge of high PV penetration in the German electric grid“. In: *IEEE Power and Energy magazine* 11.2 (2013), S. 55–64.
- [6] S. R. Madeti und S. Singh. „A comprehensive study on different types of faults and detection techniques for solar photovoltaic system“. In: *Solar Energy* 158 (2017), S. 161–185.
- [7] A. M. Pavan, A. Mellit, D. De Pieri und S. A. Kalogirou. „A comparison between BNN and regression polynomial methods for the evaluation of the effect of soiling in large scale photovoltaic plants“. In: *Applied energy* 108 (2013), S. 392–401.
- [8] S. Hallerberg und H. Kantz. „How does the quality of a prediction depend on the magnitude of the events under study?“ In: *Nonlinear Processes in Geophysics* 15.2 (2008), S. 321–331.
- [9] H. Kantz, D. Holstein, M. Ragwitz und N. K. Vitanov. „Markov chain model for turbulent wind speed data“. In: *Physica A: Statistical Mechanics and its Applications* 342.1-2 (2004), S. 315–321.
- [10] D. Witthaut, M. Rohden, X. Zhang, S. Hallerberg und M. Timme. „Critical links and nonlocal rerouting in complex supply networks“. In: *Physical review letters* 116.13 (2016), S. 138701.
- [11] J. Ludescher u. a. „Very early warning of next El Niño“. In: *Proceedings of the National Academy of Sciences* 111.6 (2014), S. 2064–2066.
- [12] S. Hallerberg und A. S. de Wijn. „Understanding and controlling regime switching in molecular diffusion“. In: *Physical Review E* 90.6 (2014), S. 062901.

- [13] X. Zhang, C. Kuehn und S. Hallerberg. „Predictability of critical transitions“. In: *Physical Review E* 92.5 (2015), S. 052905.
- [14] M. Inubushi und K. Yoshimura. „Reservoir computing beyond memory-nonlinearity trade-off“. In: *Scientific reports* 7.1 (2017), S. 10199.
- [15] J. Pathak, B. Hunt, M. Girvan, Z. Lu und E. Ott. „Model-free prediction of large spatio-temporally chaotic systems from data: A reservoir computing approach“. In: *Physical review letters* 120.2 (2018), S. 024102.
- [16] C. Steinbrink u. a. „Future perspectives of co-simulation in the smart grid domain“. In: *2018 IEEE International Energy Conference (ENERGYCON)*. IEEE. 2018, S. 1–6.
- [17] J.-M. Bourinet, F. Deheeger und M. Lemaire. „Assessing small failure probabilities by combined subset simulation and support vector machines“. In: *Structural Safety* 33.6 (2011), S. 343–353.
- [18] Z. Qian, C. C. Seepersad, V. R. Joseph, J. K. Allen und C. Jeff Wu. „Building surrogate models based on detailed and approximate simulations“. In: (2006).
- [19] J. Bremer und M. Sonnenschein. „Sampling the Search Space of Energy Resources for Self-organized, Agent-based Planning of Active Power Provision.“ In: *EnviroInfo*. 2013, S. 214–222.
- [20] D. Babazadeh, M. Chenine, K. Zhu, L. Nordström und A. Al-Hammouri. „A platform for wide area monitoring and control system ICT analysis and development“. In: *2013 IEEE Grenoble Conference*. IEEE. 2013, S. 1–7.
- [21] D. Bhor, K. Angappan und K. M. Sivalingam. „A co-simulation framework for smart grid wide-area monitoring networks“. In: *2014 Sixth International Conference on Communication Systems and Networks (COMSNETS)*. IEEE. 2014, S. 1–8.
- [22] H. Nikula u. a. „Co-simulation of a dynamic process simulator and an event-based control system: Case district heating system“. In: *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*. IEEE. 2014, S. 1–7.
- [23] S. J. Taylor und B. Letham. „Forecasting at scale“. In: *The American Statistician* 72.1 (2018), S. 37–45.
- [24] A. Paszke u. a. „Automatic differentiation in PyTorch“. In: (2017).
- [25] R. Liaw u. a. „Tune: A Research Platform for Distributed Model Selection and Training“. In: *arXiv preprint arXiv:1807.05118* (2018).
- [26] P. Branco, L. Torgo und R. P. Ribeiro. „SMOEN: a pre-processing approach for imbalanced regression“. In: *First international workshop on learning with imbalanced domains: Theory and applications*. PMLR. 2017, S. 36–50.
- [27] M. Steininger, K. Kobs, P. Davidson, A. Krause und A. Hotho. „Density-based weighting for imbalanced regression“. In: *Machine Learning* 110 (2021), S. 2187–2211.
- [28] G. Zerveas, S. Jayaraman, D. Patel, A. Bhamidipaty und C. Eickhoff. „A transformer-based framework for multivariate time series representation learning“. In: *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*. 2021, S. 2114–2124.

- [29] C. Steinbrink u. a. „CPES Testing with mosaik: Co-Simulation Planning, Execution and Analysis“. In: *Applied Sciences* 9.5 (2019). ISSN: 2076-3417. DOI: 10.3390/app9050923. URL: <https://www.mdpi.com/2076-3417/9/5/923>.

Anhang: Inverter Failure Detection Use Case (Auszug)

Im Folgenden ist ein Auszug aus dem Inverter Failure Detection Anwendungsfall dargestellt.

Content

Abstract (1-Page)	2
1. Description of the Use Case	3
1.1. Name of the Use Case	3
1.2. Version Management	3
1.3. Scope and Objectives of Use Case	3
1.4. Narrative of Use Case	4
1.5. Key Performance Indicators (KPI)	6
1.6. Use Case Conditions	6
1.7. Further Information to the Use Case for Classification/Mapping	7
1.8. General Remarks	7
2. Diagrams of Use Case	8
3. Technical Details	10
3.1. Actors	10
3.2. References	11
4. Step by Step Analysis of the Use Case	12
4.1. Overview of Scenarios	12
4.2. Steps – Scenarios	13
5. Information Exchanged	17
6. Requirements	17
7. Common Terms and Definitions	18
8. Custom Information	18
Literature	18
Open Questions	18
Notes	19

1. Description of the Use Case

1.1. Name of the Use Case

Use case identification		
ID	Area Domain(s)/Zone(s)	Name of the use case
UC#1	DER / Field	Inverter Failure Detection (short term)

1.2. Version Management

Version management				
Version no.	Date	Name of author(s)	Changes	Approval status
0.0	17.11.2020	Wibbeke	Initialization of document	-
1.0	05.01.2021	Wibbeke	First version of the document. (Section 6 Requirements not included).	-
1.1	18.01.2021	Wibbeke	Revision of the document. Diagrams and Scenarios adapted	-
1.2	22.01.2021	Wibbeke	Literature section added. Revision after meeting with Gabriel and Dominik	-
1.3	29.01.2021	Wibbeke	Slight changes over the week, proof-reading done.	-
1.4	19.05.2021	Wibbeke	Included minor changes from the Use-Case Meeting yesterday. Use Case considered Done.	yes

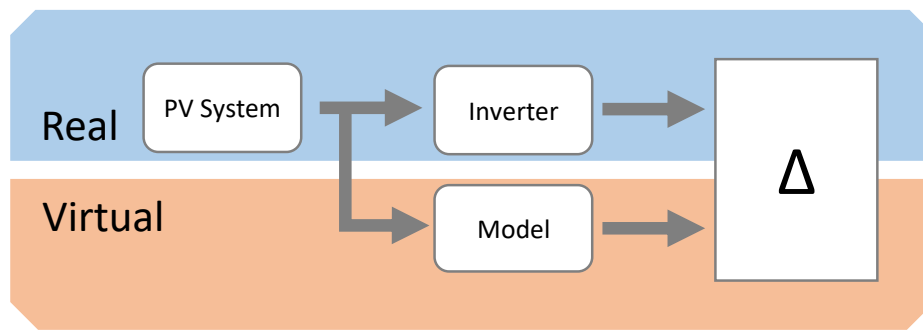
1.3. Scope and Objectives of Use Case

Scope and objectives of use case	
Scope	<p>The use case shall provide the following functionality to inverters: During operation, the function monitors the health status of the inverter components by processing real time sensor data of the inverter parameter. The goal is to detect and take actions against imminent faults, before they can damage the inverter. Thus, if an imminent harmful condition is detected, the function may issue the curbing or an emergency shutdown of the inverter. If a persisting fault is detected, the function shall classify the fault and rise an alarm to the operator. In both scenarios, the curbing/shutdown is intended to suppress the spread of harmful conditions to previously unaffected components, preventing a cascading effect.</p> <p>The function is designed only for large-scale inverters with sufficient real-time sensing capacities.</p> <p>The use case does not cover actions of the Inverter Controller.</p>
Objective(s)	<ul style="list-style-type: none"> • Inverter monitoring: The aim of the use case is to monitor the state of the inverter and detect faulty component states • Inverter component protection: If an inverter fault is imminent, the function is able to shut down the inverter, before damaging conditions can spread to unaffected components. • Error source limitation: When a faulty state was detected, the sources and included components are to be identified

	<ul style="list-style-type: none"> • Operator notification: When a fault is detected or a shutdown is issued, the operator is notified
Related business case(s)	<ul style="list-style-type: none"> • Improving the monitoring of inverters • Improving the reliability of inverters • Downtime cost saving • More efficient maintenance • Preventive Maintenance

1.4. Narrative of Use Case

Narrative of use case	
Short description	
<p>For the protection of large-scale inverters, it should be possible for the user to add a function that monitors the status of the inverter. Special focus is on the detection of imminent or already existing errors within the inverter. After classifying the errors, the affected components shall be determined and the user shall be informed. In the event of a major fault, the function shall also be able to perform an emergency shutdown of the inverter to protect adjacent systems/components from cascading effects.</p>	
Complete description	
<p>When constructing PV systems, the investment in the required inverters is a significant cost factor. Whereas if the inverter fails, no power can be generate. The inverter therefore has a big impact on the profitability of the plant. In order to minimize the downtime of these inverters and to reduce maintenance costs, a mechanism is to be created which allows the health status of the inverter components to be monitored and the inverter to be switched off in the event of occurring irregularities.</p> <p>The system is focused on large inverters, where acquisition costs exceed maintenance costs and it therefore is economically more feasible to locate and maintain faulty components than to simply replace the inverter.</p> <p>Procedure:</p> <p>The operator of the system can switch on the failure detection function via the Inverter Controller. After switching on, it accesses the database of the inverter. In the database, information about the inverter is gathered regarding preferences, models, metadata and historical data. The preferences and historical sensor data gathered during operation of the inverter are used to derive the parameters and properties of the inverter to create a virtual model by fitting one of the models in the database to the inverter. The goal of the model is to represent the inverter as accurately as possible.</p> <p>After verifying the model, it is connected to the real-time sensor data of the inverter. The sensor data is requested from the Data Logger and Flight recorder of the inverter. It includes input values (e.g. DC voltage) as well as output values (e.g. AC voltage). The input values applied to the inverter are now also fed into the model of the inverter. This allows a comparison between the real output values at the inverter and the theoretical output values calculated by the model.</p>	



Since the model was generated using the data of a healthy, functioning inverter, it is to be assumed that in case of a deviation of the calculated output to the measured output, the inverter deviates from the optimal status and functional problems are present.

After comparing the values, the function categorizes the status of the inverter into one of three health categories, which are defined as follows:

- Healthy
 - The inverter is in a good status and no fault is expected.
- Imminent fault
 - The inverter is in a good state, but there are indications of an impending fault.
- Persisting fault
 - A fault has been identified

In addition to the categorization of the health status, the function should also be able to evaluate where in the inverter the fault probably lies. For this, the respective affected components are to be determined.

After categorization, each category also triggers an action scheme:

- Healthy
 - No particular action is taken, monitoring continuous
 - Notify the user about the status
- Imminent fault
 - Determine which components are likely to be affected
 - Calculate the failure occurrence probability
 - Decide about the shutdown/curbing of the inverter
 - Notify the user about the status
- Persisting fault
 - Determine which components are affected
 - Determine the propagation probability of the fault
 - Determine the severity of the error
 - Decide about the shutdown/curbing of the inverter
 - Notify the use about the status

The emergency shutdown or curbing of the inverter is intended to prevent cascading effects from being triggered by the fault, resulting in a complete loss of the inverter. The decision about and initiation of the action must therefore always be made quickly enough to prevent the fault from occurring and/or spreading to surrounding components and systems.

1.5. Key Performance Indicators (KPI)

Key performance indicators			
ID	Name	Description	Reference to mentioned use case objectives
KPI_1	Number of issued shutdowns	Monitors the average number of issued inverter curbing/shutdowns by the function per year over the lifetime (shutdown/year).	Inverter component protection
KPI_2	Number of operator notification	Monitors the average number of operator notifications issued by the function per year over the lifetime (notifications/year).	Operator notification
KPI_3	In time detection quotient	The function shall be able to detect imminent faults/faulty states before the fault develops into a fatal one. The indicator is calculated by the quotient of the number of incidents where a fault was detected and the number of incidents where a inverter component is damaged (detected/damaged)	Inverter monitoring
KPI_4	Number of affected components	After an action is issued to the inverter the number of affected components is evaluated (components/shutdown)	Error source limitation
KPI_4	False alarm rate	After the analysis of a shut down through the maintenance team, the team decides whether the shutdown was issued correctly or not. False Alarm / Total alarms	Inverter Monitoring

1.6. Use Case Conditions

Use case conditions
Assumptions
There is a correlation between sensor data and faults.
The real time sensor data stream is of sufficient variety and sampling rate to detect faults
The real time sensor data stream is of sufficient variety and sampling rate to determine the fault type
The historical data in the database is of sufficient volume, variety and variability to generate a model capable of monitoring the inverter.
The database contains model blueprints, which are suitable for the detection purpose.
The historical data in the database is up to date.
The function only has to send an emergency shutdown command to the Inverter Controller in order to shut down the inverter.
The function only has to send a notification command to the Inverter Controller in order to notify the operator.
The function only has to send a curbing command to the Inverter Controller in order to curb the inverter components.
The function can receive time stamped real time measurement data from the Flight Recorder by command.
The function can receive time stamped real time measurement data from Data Logger by command.
Prerequisites