

Implementation of a network cluster model for large-scale renewable power estimation and effective energy utilization

Von der Fakultät für Maschinenbau, Elektro- und Energiesysteme
der Brandenburgischen Technischen Universität Cottbus-Senftenberg
zur Erlangung des akademischen Grades eines
Doktors der Ingenieurwissenschaften (Dr.-Ing.)

genehmigte Dissertation

vorgelegt von

M.Sc.

Yang LI

geboren am 23.09.1989 in Anhui, China

Vorsitzender: Prof. Dr.-Ing. Georg Möhlenkamp

Gutachter: Prof. Dr.-Ing. Harald Schwarz

Gutachter: Dr. hab. inz. Przemysław Janik

Tag der mündlichen Prüfung: 25.07.2023

DOI: <https://doi.org/10.26127/BTUOpen-6456>

Vorwort:

Diese Dissertation wurde am Lehrstuhl Energieverteilung und Hochspannungstechnik der Brandenburgischen Technischen Universität Cottbus-Senftenberg verfasst. Ohne die zahlreiche fachliche und persönliche Unterstützung von Kollegen wäre die Erstellung dieser Arbeit nicht möglich gewesen. Kein Erfolg ist eine Einzelleistung. Daher möchte ich allen beteiligten Unterstützern meinen aufrichtigen Dank aussprechen.

Mein besonderer Dank gilt Prof. Dr.-Ing. Harald Schwarz, dem Leiter des Lehrstuhls Energieverteilung und Hochspannungstechnik, für sein Vertrauen, seine kontinuierliche Unterstützung und Betreuung.

Ich möchte mich besonders bei Dr. hab. inz. Przemysław Janik für die fachlichen Diskussionen, Betreuung, und interessanten Fragestellungen zur Methodik und zu den Ergebnissen.

Ein weiterer besonderer Dank geht an Dr.-Ing. Klaus Pfeiffer für die persönlichen Fachgespräche und die vielen hilfreichen Kommentare, insbesondere während der Promotionszeit.

Ein herzliches Dankeschön gilt all meinen Kollegen und Freunden, die mir in guten wie in schwierigen Zeiten zur Seite standen. Ihre Kameradschaft hat die Forschungs- und Studienreise angenehmer und weniger entmutigend gemacht. Meiner Familie, insbesondere meiner Frau Ying und meiner Tochter Julia, danke ich für ihre Liebe, ihre Unterstützung und ihr Vertrauen in mich – sie waren meine Stützen. Ihre ständige Ermutigung und ihr Verständnis während dieser schwierigen Zeit bedeuten mir sehr viel.

Cottbus, im August 2023

Yang LI

Abstract

With the large-scale expansion of decentralized power generation from renewable energy resources (RES) and reduction of fossil fuel-based power generation, this leads to fundamental structural changes in the power supply system. The transition is characterized by a shift from central to decentral, from directional to bi-directional. As the share of renewable power systems in the electricity grid increases, the inherent uncertainty of renewable energies poses challenges to the grid system's stability.

To achieve increased efficiency, decarbonisation, decentralization, and digitisation in energy sector, a flexible and resilient energy system is required. Electrical energy data and non-electrical energy data are generated from various sources, including the supervisory control and data acquisition system (SCADA), geographic information system (GIS), and weather information system. Based on these data, different spatial and temporal models of energy systems and their databases are created and connected. This thesis contributes to the analysis of regional power performance. To this end, an efficient simplified grid-oriented network cluster model is presented to account for the regional allocation of RES power systems. The model clustering includes the grid topology, temporal and spatial resolution, and the structure of electrical and non-electrical data. This approach contributes to methodological aspects of energy system modelling and analysis.

In addition, based on the proposed cluster model, this thesis presents the physical profile-based and data-driven methods to estimate the regional large-scale photovoltaic (PV) and wind power generations. In order to adapt to the spatial-temporal heterogeneity of the regional renewable power generations and to improve the local power utilization rate, this thesis conducts empirical analyses of the regional energy storage and combined PV and wind power systems using cable pooling with shared grid connection.

In summary, this thesis presents a flexible modelling approach that takes into account the regional allocation of renewable power systems and enables a more accurate estimation of regional power generation. The proposed model and analytical methods can help the efficiency and resiliency of the energy system and support the transition towards renewable energies.

Kurzfassung

Mit dem großflächigen Ausbau der dezentralen Stromerzeugung aus erneuerbaren Energiequellen (RES) und der Reduzierung der fossilbasierten Stromerzeugung kommt es zu grundlegenden strukturellen Veränderungen im Stromversorgungssystem. Der Übergang ist durch einen Wandel von zentral zu dezentral und von direktional zu bi-direktional gekennzeichnet. Mit zunehmendem Anteil erneuerbarer Energiesysteme im Stromnetz stellen die inhärente Unsicherheit dieser erneuerbaren Energien Herausforderungen für die Stabilität des Netzes dar.

Um eine erhöhte Effizienz, Dekarbonisierung, Dezentralisierung und Digitalisierung im Energiesektor zu erreichen, ist ein flexibles und robustes Energiesystem erforderlich. Elektrische Energie- und Nichtelektrische Energiemessdaten werden aus verschiedenen Quellen generiert, einschließlich des Supervisory Control and Data Acquisition Systems (SCADA), des Geographic Information Systems (GIS) und des Wetterinformationssystems. Basierend auf diesen Daten werden verschiedene räumliche und zeitliche Modelle von Energiesystemen und deren Datenbanken erstellt und miteinander verbunden. Diese Arbeit trägt zur Analyse der regionalen Leistung bei. Zu diesem Zweck wird ein effizientes, vereinfachtes netzorientiertes Netzcluster Modell vorgestellt, um die regionale Zuordnung von RES-Energiesystemen zu berücksichtigen. Das Modell-Clustering umfasst die Netztopologie, zeitliche und räumliche Auflösung sowie die Struktur der elektrischen und nichtelektrischen Daten. Dieser Ansatz trägt zu methodischen Aspekten der Energiesystemmodellierung und -analyse bei.

Darüber hinaus präsentiert diese Arbeit auf der Grundlage des vorgeschlagenen Clustermodells physikbasierte und datengetriebene Methoden zur Schätzung der regionalen großflächigen Photovoltaik- (PV) und Windstromerzeugung. Um sich an die räumlich-zeitliche Heterogenität der regionalen erneuerbaren Stromerzeugung anzupassen und die lokale Stromnutzungsgrade zu verbessern, führt diese Arbeit empirische Analyse der regionalen Energiespeicherung und kombinierte PV- und Windenergiesysteme unter Verwendung von Kabel-Pooling mit gemeinsam genutzten Netzverbindungen durch.

Zusammenfassend präsentiert diese Arbeit einen flexiblen Modellierungsansatz, der die regionale Zuordnung von erneuerbaren Energiesystemen berücksichtigt und eine genauere Schätzung der regionalen Stromerzeugung ermöglicht. Das vorgeschlagene Modell und die Analysemethoden können die Effizienz und Robustheit des Energiesystems verbessern und den Übergang zu erneuerbaren Energien unterstützen.

Table of Contents

Abstract	I
Kurzfassung.....	II
Table of Contents	III
List of Figures	VI
List of Tables.....	IX
List of Abbreviations.....	X
List of Symbols	XII
1 Introduction and Motivation.....	1
1.1 Introduction	1
1.2 Problem Analysis and Challenges	3
1.3 Motivation and Goal.....	6
1.4 Structure of the Thesis.....	7
2 Modelling of Regional Network Cluster.....	9
2.1 Overview of Power Supply System.....	9
2.2 Concept of Regional Network Cluster.....	12
2.2.1 Grid-Oriented Cluster Model	12
2.2.2 Cluster Characteristics.....	13
2.3 Data Sources	16
2.3.1 Spatial and Temporal Datasets.....	17
2.3.2 Renewable Energy Master Dataset	19
2.3.3 Clustering of Datasets	22
2.4 Cluster Related Performance	24
2.4.1 Allocation of Decentralized Renewable Generation.....	24
2.4.2 Unbalance between Power Generation and Load	25
2.4.3 Mutual Relations of Cluster Characteristics.....	29
2.5 Concluding Remarks	30

3	Methodology for the Regional Aggregated Power Estimation	32
3.1	Overview of Regional RES Power Estimation.....	32
3.2	Aggregated Wind Power Generation Model	33
3.2.1	Wind Power Output Characteristic	33
3.2.2	Modelling of Regional Aggregated Wind Power.....	34
3.3	Aggregated PV Power Generation Model	39
3.3.1	PV Power Output Characteristic	39
3.3.2	Modelling of Regional Aggregated PV Power	41
3.4	Data Driven Models.....	44
3.4.1	Artificial Intelligence	44
3.4.2	Modelling of the Regional Aggregated Power.....	48
3.5	Evaluation of Predictions.....	50
4	Implementation of Cluster Model in Regional Power Estimation	51
4.1	Simulation Based on Neural Computations.....	51
4.1.1	Experimental Setup and Data Pre-Processing.....	51
4.1.2	Comparison of Different Input Categories.....	58
4.2	Simulation Based on Aggregated Model.....	64
4.2.1	Experimental Setup and Data Pre-processing	64
4.2.2	Regional Wind Power Estimation	67
4.2.3	Regional PV Power Estimation.....	71
4.3	Simulation based on Deep Learning Method	74
4.3.1	Experimental Setup and Data Pre-processing	75
4.3.2	Comparison of Different Models	76
4.4	Concluding Remarks	78
5	Implementation of Cluster Model in Virtual Regional Energy Storage System.....	80
5.1	Overview of Regional Energy Storage System	80
5.2	Methodology for Storage System.....	81
5.3	Model Development and Experimental Setup.....	83

5.4	Simulation and Results	87
5.5	Concluding Remarks	92
6	Implementation of Cluster Model in Cable Pooling Approach.....	93
6.1	Overview of Cable Pooling	93
6.2	Methodology for the System Modelling.....	94
6.2.1	Assessment Approaches	94
6.2.2	System Modelling	97
6.3	Results of Case Studies for Single and Regional RES Models	99
6.4	Concluding Remarks	104
7	Conclusion and Outlook.....	105
7.1	Summary.....	105
7.2	Future Work.....	106
	Bibliography.....	108

List of Figures

Figure 1-1: Number of PV power plants (a) and installed capacity (b) in Germany [data based on [5]].....	2
Figure 1-2 Number of wind power plants (a) and installed capacity (b) in Germany [data based on [5]].....	2
Figure 1-3 Overview of the scenarios framework of the German NEP-2035 [data based on [4]]	2
Figure 1-4: Vertical power flow of 50Hertz-T control area [data based on [5]].....	4
Figure 1-5: Development of curtailment energy and costs within feed-in management [data based on [11,12]].....	5
Figure 2-1: Hierarchical structure of electrical energy system [based on [26]].....	10
Figure 2-2: Regional electrical network topology.....	12
Figure 2-3: Regional aggregated network cluster	13
Figure 2-4: Different scenarios of horizontal power flows	15
Figure 2-5: Two scenarios of the vertical power flow	16
Figure 2-6: OSM data acquisition and processing	18
Figure 2-7: Time change according to UTC+1 time zone	19
Figure 2-8: Preparation of the EEG master data	21
Figure 2-9: Availability period of the EEG plant system.....	22
Figure 2-10: Data structure of network cluster model	23
Figure 2-11: Classification of network clusters (a), allocation of installed PV (b) and WT (b) power systems to the end 2019	24
Figure 2-12: Development of WT installation in different spatial resolution.....	25
Figure 2-13: Definition of percentile energy.....	26
Figure 2-14: Power saldo curves of the network cluster 3 in June (a) and December (b)	27
Figure 2-15: Annual average curves of power saldo.....	28
Figure 2-16: Annual duration profile (a) and frequency distribution (b) of the power saldo ..	29
Figure 2-17: Correlation coefficient of different variables of cluster 3	30
Figure 3-1: A typical power characteristic curve of one single wind turbine generator	34
Figure 3-2: Regional aggregated wind power characteristics	35
Figure 3-3: Procedure of the regional wind power generation calculation based on aggregation model.....	38
Figure 3-4: Influence of solar irradiation and module temperature on the PV characteristic [64]	39

Figure 3-5: Linear dependency of module temperature with air temperature (a), with MPP efficiency (b) by different solar irradianations	42
Figure 3-6: The procedure of the regional PV power generation calculation based on aggregation model.....	43
Figure 3-7: Architecture of ANN model	44
Figure 3-8: Framework of neural network for training, validation and testing.....	45
Figure 3-9: Deep learning model structures of LSTM (a) and stacked LSTM (b)	47
Figure 3-10: Structure of the input features and output target	47
Figure 3-11: CNN-LSTM structure.....	48
Figure 3-12: Procedure of the regional power generation estimation.....	49
Figure 4-1: Regional DSO grids with postal code area (a), regional power generation (b) and regional installed capacity of PV and wind power systems (c).....	52
Figure 4-2: Distribution of regional DWD weather stations.....	53
Figure 4-3: Scatter matrix and correlation matrix of data	55
Figure 4-4: PV and wind prediction models based on different input features.....	56
Figure 4-5: PV power prediction for 1 st week of June (a) and 1 st week of December in 2019 (b)	59
Figure 4-6: Wind power prediction for the first two weeks of March in 2019	59
Figure 4-7: Cumulative distribution of normalized error	61
Figure 4-8. Monthly nMAE of ANN prediction based on different input features	62
Figure 4-9. PV prediction metrics nMAE and MAPE	63
Figure 4-10. Wind prediction metrics nMAE and MAPE	64
Figure 4-11: Mean weather data for each cluster	65
Figure 4-12: The mean installed capacity and hub height for the region.....	65
Figure 4-13: Clustered installed wind capacity for different age groups (status end year 2019)	66
Figure 4-14: Clustered mean hub height for different age groups (status end year 2019).....	66
Figure 4-15: Clustered installed PV capacity for different voltage levels (status end year 2019)	67
Figure 4-16: Modelled clustered wind power generation	68
Figure 4-17: Monthly relative wind peak power output from historical measurements and model calculations.....	68
Figure 4-18: Regional aggregated wind power generation in January 2019.....	69

Figure 4-19: Monthly relative wind peak power output with correction factor (a) and the annual duration curve of regional wind power (b).....	71
Figure 4-20: Regional aggregated PV power generation in January and June 2019	72
Figure 4-21: Monthly relative PV peak power output from historical measurements (a) and model calculations (b)	72
Figure 4-22: Monthly relative PV peak power output with correction (a) and its annual duration curve of regional PV power (b).....	73
Figure 4-23: Wind power estimation based on different neural network models	76
Figure 4-24: Monthly nMAE of deep learning models.....	77
Figure 4-25: Annual nMAE and nRMSE of deep learning models	77
Figure 4-26: Annual duration curve of wind power estimation from deep learning models ...	78
Figure 5-1: Energy storage system in power supply system	82
Figure 5-2: Charging/discharging strategies for a storage system	83
Figure 5-3: Basic configuration of a fuzzy system	84
Figure 5-4 Input / Output membership functions.....	85
Figure 5-5: Profiles of RES generations (a), feedlines power and power saldo for Cluster 3 from 1 st June to 3 rd June (b) of year 2019	87
Figure 5-6: Procedure of the proposed simulation	88
Figure 5-7: Normalized power output profiles.....	89
Figure 5-8: SoC of ESS.....	89
Figure 5-9: Distribution of normalized power ESS	90
Figure 5-10: Distribution of 15-min deviation of normalized power ESS.....	90
Figure 5-11: Power saldo profiles with / without ESS.....	91
Figure 5-12: Distribution of power saldo with / without ESS.....	91
Figure 5-13: Distribution of 15-min deviation of power saldo with / without ESS.....	91
Figure 6-1: Independent grid connection (a) and cable pooling with shared grid connection (b)	94
Figure 6-2: Wind power output performance model.....	98
Figure 6-3: Power curves of combined system with different PV systems	100
Figure 6-4: Regional installed capacity of PV and wind power systems from 2014 to 2019	101
Figure 6-5: Capacity factor from regional measurements for PV and wind power systems .	101
Figure 6-6: Energy curtailment (a) and capacity factor (b) for different capacity ratios	103
Figure 6-7: Stability coefficient for different capacity ratios.....	103

List of Tables

Table 2-1: Voltage levels in Germany [25].....	10
Table 2-2: Example search options with tags in OSM under a specific area.....	17
Table 2-3: Data state of the two data sources to the end of year 2019.....	20
Table 2-4: Installed capacities of PV and WT power systems in clusters	25
Table 2-5: Results of the regional power saldo in the cluster 3	27
Table 3-1: Parameters for aggregated characteristics.....	37
Table 3-2: Typical values of parameters for different module and mounting types [21,65] ...	40
Table 4-1: Number of weather measurement stations.....	53
Table 4-2: PV neural network based on different input category parameters.....	57
Table 4-3. Wind neural network based on different input category parameters	57
Table 4-4: Parameter setup for ANN model	58
Table 4-5: MAE and nMAE for PV ANN Model	60
Table 4-6: MAE and nMAE for Wind ANN Model	60
Table 4-7: Normalized error distribution	61
Table 4-8: Monthly and annual average metrics of the regional wind power estimation for year 2019	70
Table 4-9: Monthly and annual average metrics of the regional PV power estimation for year 2019	74
Table 4-10: Parameters for neural networks	75
Table 5-1: Input fuzzy sets description	85
Table 5-2: Output fuzzy sets description.....	85
Table 5-3: Conditional statements for control systems.....	86
Table 5-4: Assessment items.....	87
Table 6-1: Parameter setting for WT model.....	97
Table 6-2: Parameter setting for PV model.....	98
Table 6-3: Effects of adding different PV systems to one WT system	100

List of Abbreviations

AF	Activation Function
AI	Artificial Intelligence
ANN	Artificial Neural Network
BNetzA	Federal Network Agency
CDC	Climate Data Center
CEST	Central European Summer Time
CET	Central European Time
CF	Capacity Factor
CM	Commission
CNN	Convolutional Neural Network
CT	Category
DENA	Germany Energy Agency
DM	Decommission
DSO	Distribution System Operator
DWD	German Weather Service
e2p	Energy to Power ratio
EEG	Renewable Energy Sources Act
EHV	Extra-High Voltage
EMS	Energy Management System
ESS	Energy Storage System
FIS	Fuzzy Inference System
FLH	Full Load Hour
GIS	Geographic Information System
HV	High Voltage
LCOE	Levelized Cost of Electricity
LSTM	Long Short Term Memory
LV	Low Voltage
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MaStR	Market Master Data Register
MLR	Multi Linear Regression
MPP	Maximum Power Point
MV	Medium Voltage

NC	Network Cluster
NEP	Network Development Plan
NH	Negative High
NL	Negative Low
nMAE	normalized Mean Absolute Error
nRMSE	normalized Root Mean Squared Error
OSM	OpenStreetMap
PCC	Point of Common Coupling
PDF	Probability Density Function
PH	Positive High
PL	Positive Low
PTDF	Power Transfer Distribution Factor
PV	Photovoltaic
RES	Renewable Energy Sources
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
SCADA	Supervisory Control and Data Acquisition System
SLSTM	Stacked LSTM
SoC	State of Charge
STC	Standard Test Condition
Sub	Substation
SVM	Support Vector Machine
TSO	Transmission System Operator
UTC	Coordinated Universal Time
Var	Variance
VPP	Virtual Power Plant
WT	Wind Turbine
ZE	Zero

List of Symbols

c_p	Power coefficient	
C_{stab}	Stability coefficient	%
E_{cap}	Rated energy capacity	Wh
E_{curt}	Curtailement energy	Wh
E_{neg}	Negative percentile energy	Wh
E_{pos}	Positive percentile energy	Wh
f_c	Dynamic correction index	%
f_{curt}	Curtailement energy percentage	%
I_{GS}	Global solar irradiation	W/m ²
I_{STC}	Solar irradiation by STC	W/m ²
$p_{norm,PV}$	Normalized PV power output	%
$p_{norm,wind}$	Normalized wind power output	%
P_{inst}	Installed capacity	W
P_{pred}	Predicted power output	W
P_{true}	Real measured power generation	W
P_{ESS}^{NC}	Power flow of energy storage system in network cluster	W
P_{Gen}^{NC}	Power generation of network cluster	W
$P_{hor_j}^{NC}$	Horizontal power flow from substation j to network cluster	W
$P_{Line_k}^{NC}$	Power flow of feed-line k to the network cluster	W
P_{Load}^{NC}	Power load of network cluster	W
P_{PV}^{NC}	PV power generation of network cluster	W
P_{RL}^{NC}	Residual load of network cluster	W
P_{Saldo}^{NC}	Power saldo of the network cluster	W

P_{Wind}^{NC}	Wind power generation of network cluster	W
P_{ver}^{Sub-j}	Vertical power flow from the substation j	W
t_{curt}	Curtailement time percentage	%
T_{air}	Air temperature	°C
T_{mod}	Solar module temperature	°C
T_{STC}	Solar module temperature by STC	°C
v_{10}	Wind speed at 10m height	m/s
v_{cut-in}	Cut-in wind speed	m/s
$v_{cut-out}$	Cut-out wind speed	m/s
v_h	Wind speed at hub height of h	m/s
v_{rated}	Rated wind speed	m/s
v_{wind}	Wind speed	m/s
α	Temperature coefficient	%/°C
β	Scale parameter	m/s
ε	Normalized prediction error	%
γ	Parameter for temperature rise by solar irradiation	°C*m ² /W
δ	Hellmann exponent	
μ	Location parameter	m/s
η_{inv}	Inverter efficiency	%
η_{MPP}	Normalized MPP efficiency	%
ρ_{air}	Air density	kg/m ³
$\rho_{x,y}$	Pearson correlation coefficient	

1 Introduction and Motivation

1.1 Introduction

Renewable energies contribute a major share of the energy supply in the world [1]. Since 2000, the German Renewable Energy Sources Act (EEG - Erneuerbare-Energien-Gesetz) was launched to encourage the sustainable development of energy supply through renewable energy sources (RES). EEG accelerates indeed the investment in RES in the power supply system. The German installed capacity of RES increases from around 10 GW in 2000 to approx. 132 GW in 2020 [2]. Among the RES, the integration of photovoltaic (PV) and wind turbine (WT) generators in the power grid system is increasing. Figure 1-1 and Figure 1-2 show the capacity situation of the PV and WT power plant systems from the four control zones of the German transmission grid systems. Up to end 2019, approx. 30,000 WT generators with a total installed capacity of about 61 GW and over 1.8 million PV plants with a total installed capacity of 48 GW are connected to the entire German electricity grid networks from low-voltage (LV) to extra-high-voltage (EHV) levels. Obviously, a very high number and capacity of PV and WT power plants are connected to distribution system operators (DSO) with the voltage levels from LV to high-voltage (HV). About 25.5 GW of PV systems has been installed at LV grid, which corresponds to approx. 53 % of the total installed PV capacity. In addition, about 85 % WT power plants are mainly connected to the medium-voltage (MV) to high-voltage (HV) networks. In order to reduce greenhouse gas emissions, the electricity grid is to be increasingly penetrated by renewable energies. In addition to decommissioning German nuclear power plants, the German government has decided to phase out coal-fired power generation in 2019 [3]. The Coal Phase-Out Act stipulates that the last conventional coal-fired power plants (lignite and hard coal) is to be closed down no later than 2038 [3]. Therefore, a considerable part of the power supply must be covered by renewable energies in the near future. To achieve this, the German four transmission system operators (TSOs) have published scenario frameworks for the network development plan 2035 (NEP-2035) [4], which will double the current installed capacities of renewable generation, particularly PV and WT power plants. The three scenarios for the NEP-2035 are showed in the Figure 1-3.

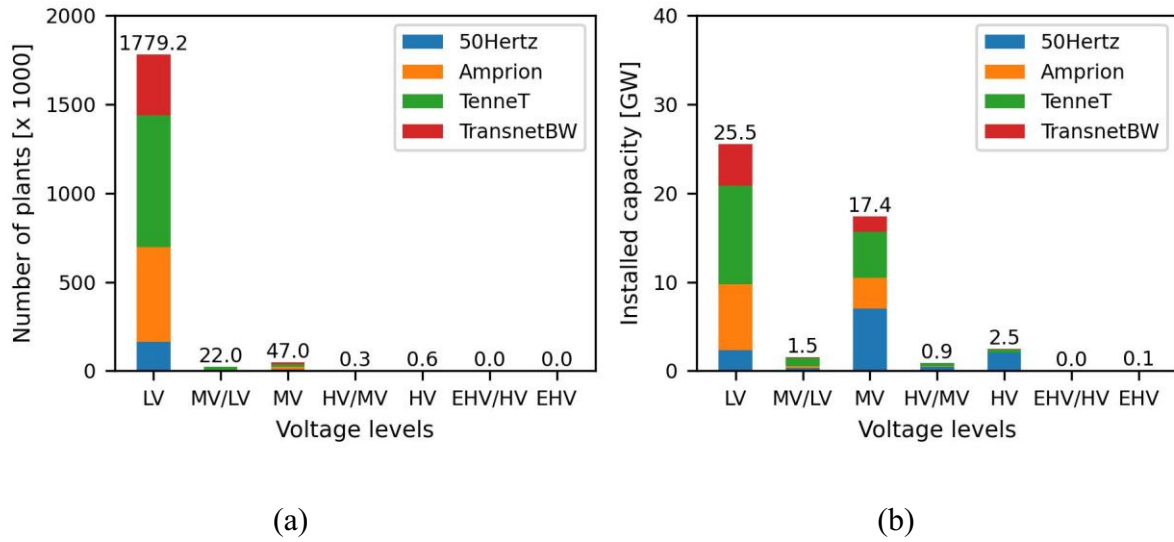


Figure 1-1: Number of PV power plants (a) and installed capacity (b) in Germany [data based on [5]]

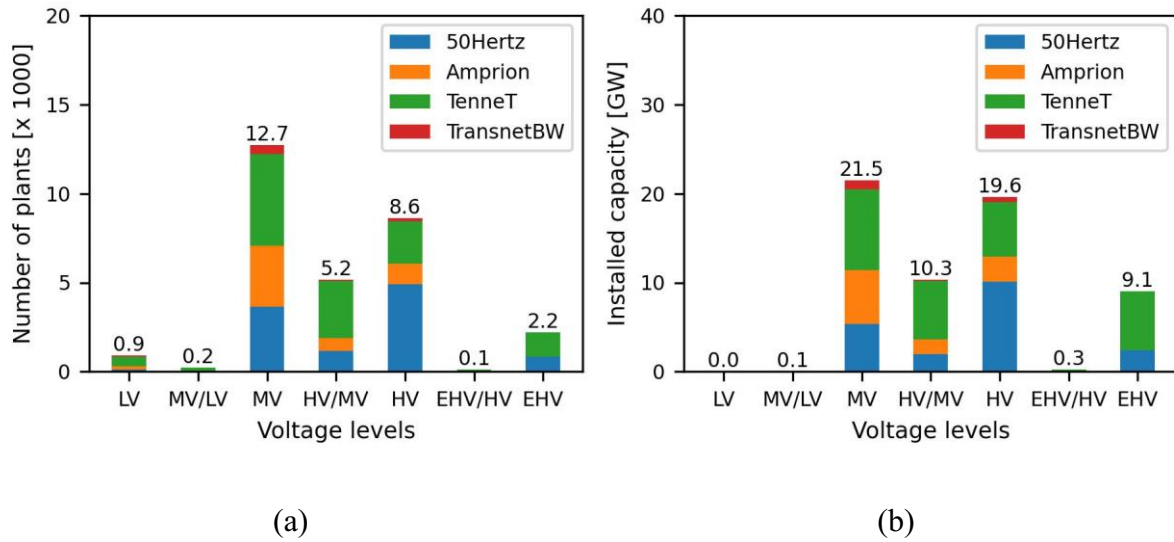


Figure 1-2 Number of wind power plants (a) and installed capacity (b) in Germany [data based on [5]]

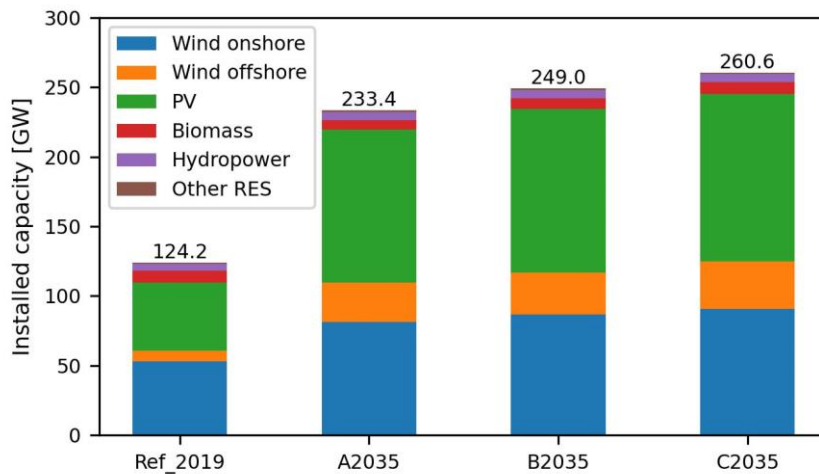


Figure 1-3 Overview of the scenarios framework of the German NEP-2035 [data based on [4]]

In addition to the amount of RES installation capacity, the temporal balance of demand and power generation is also decisive. The electricity load in Germany varies from 35~40 GW (off-peak load) to 80~85 GW (peak load) [2,6]. According to the report [2], the share of RES in gross electricity consumption in Germany has risen continuously since 2000. It rose from 6 % in 2000 to 17 % in 2010 and to approx. 45 % in 2020. In some distribution network areas and some federal states in northeast Germany, the share of renewable energy in gross electricity consumption is high and even over 100 percent [7]. However, this does not mean the full energy supply from RES all time in the power grid system. Because only the feed-in of extreme overproduction of RES electricity for some periods of time, such as days or weeks [6]. In the context of energy transition, the energy supply is no longer provided by the limited number of large power plants but a variety of decentralized power generation systems. The drastic changes of the German electricity energy system with time-fluctuating nature of RES power generation bring new challenges and requirements for the stability and reliability of the power grids and the balance of power between electricity generation and consumption. This is because the electrical power generation such from wind intensity and solar irradiation energy rarely meets the demand for electricity power from time to time.

1.2 Problem Analysis and Challenges

With the massive expansion of decentralized power systems in the past decades in the electricity networks, primarily from PV and WT power systems, the power supply system has been fundamentally changing, from central to de-central, from directional to bi-directional [8]. The necessary further expansion and increasing complexity of power supply system pose new challenges due to the intermittent power generation, energy management and bidirectional power flows.

The vertical power through transformers indicates that how much power is exchanged from different voltage levels. It can be seen from Figure 1-4 that the 50Hertz-T (50Hertz Transmission GmbH), one of the four TSOs in Germany, supplied less and less electricity energy to the connected distribution grid systems in its control area. On the contrary, the amount of negative vertical electricity power increases, which means the feedback power from the connected distribution grids to the upstream transmission grid, whereas in the past the transmission and distribution of electricity power was always from higher voltage level to medium- and low-voltage grids. In addition, the maximum reference power from 50Hertz-T has remained at about 10 GW over the last ten years. At the same time, the maximum vertical

feedback power has continuously increased from 2 GW in 2010 to 7.2 GW in 2020. More attention on the vertical feedback power should be concerned due to electricity overproduction from RES power plants connected to the distribution grid system.

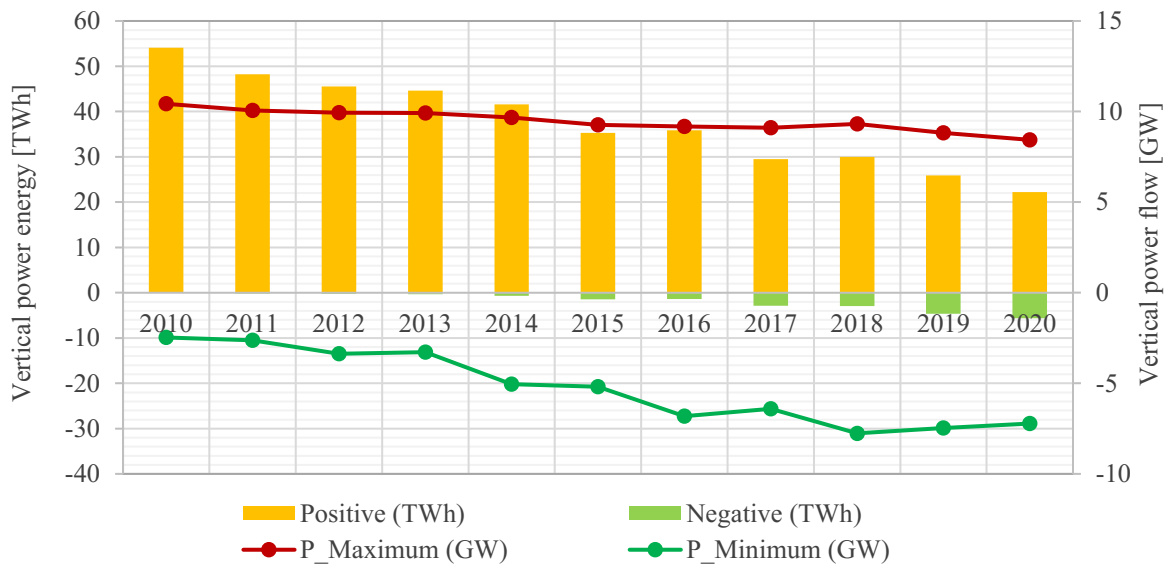


Figure 1-4: Vertical power flow of 50Hertz-T control area [data based on [5]]

Moreover, the inherent uncertainty of the RES energy increases the instability of the power system. Sometimes the grid system cannot absorb so much intermittent RES power at all, because the amount of electricity production exceeds the consumption and the grid capacity. A congestion becomes apparent at certain power system lines and transformers [9]. The (n-1) criterion for electricity grid system is endangered by grid congestion with regard to the high penetration of RES. In order to guarantee the grid components not to be overloaded and maintain the grid security, various security measures must be implemented. Within the framework of feed-in management [10], renewable energy feed-ins are forcibly curtailed when part of the distribution or transmission systems are congested, or the supply security is threatened by the congestion. As can be seen from the Figure 1-5, in recent years the feed-in management volume and the resulting curtailment energy has been increasing rapidly. The associated costs for this measure, which is compensated for the curtailment energy, have also risen continuously. Within the framework of the feed-in management in Germany, approx. 6 million euros were paid for the measure compensation in 2009 and this value rose to around 760 million euros in 2020 for 6,146 GWh curtailed RES energy [11]. Approx. 2.7 % of the curtailed energy are from solar energy source and a high number of 96.7 % are from wind energy source. According to the reports [11,12], a total around 80~85 % curtailment are caused by bottlenecks in transmission grid systems, but up to 74~98 % curtailed energy occur in distribution grid systems. This is because more than 90 % of RES generation are connected to

the electrical distribution systems [13]. Future role of DSOs in the energy transition is characterised by greater autonomy in system operation and responsibility for the overall system [14]. Therefore, the efficient energy use and transport of the decentralized energy remains the central challenge for distribution grid systems.

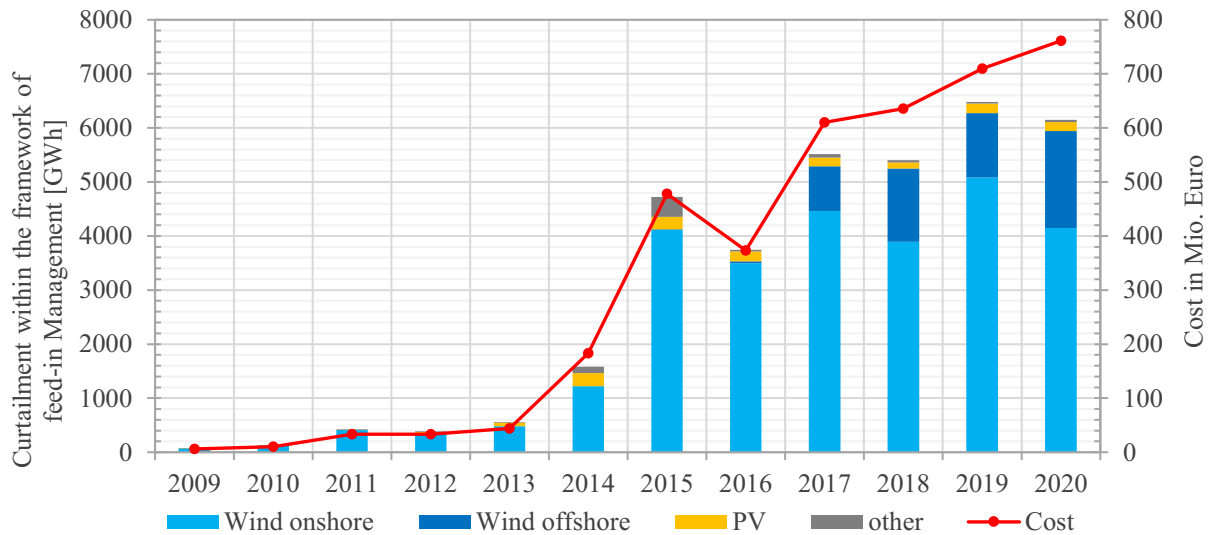


Figure 1-5: Development of curtailment energy and costs within feed-in management [data based on [11,12]]

The steady increase of electricity feed-in power into the grid and the temporal electricity overproduction pose new challenges for the power system. These changes, which are characterized by steadily increasing weather-dependent electricity supplies from renewable energies, lead to increased grid congestion in the electricity transmission and distribution networks. In particular, the power generation profiles of the highly weather-dependent power generations from solar and wind energies are less predictable and harmonic between power generation and consumption [15,16]. Wind turbine generators mainly produce electricity when there is a lot of wind energy, and PV system can only convert solar irradiation to electricity at daytime. In addition, the spatial distribution of the installed renewable energy system is highly dependent on the meteorological and geographical factors [17,18]. While the conventional large power plants are positioned in the near of metropolitan areas, the renewable electricity has to be transported from sparsely populated areas to the electricity consumption centres. The spatial and temporal distribution of the renewable energies is essential for the regional power balance. The assignment of the regional renewable power supply system takes place through inhomogeneous distribution of the RES systems. The power generation from single RES power plant under different operating conditions and montage types are analysed in [19,20], but regional power output and prediction should be based on the large-scale power generation in entire region. The aggregated power output is very important for the development of the

regional network planning and energy management of the power generation and consumption as well as the stability of the power supply system [21]. Therefore, an accurate prediction especially real-time estimation of PV and wind power gradually plays a significant role in reducing power fluctuations in the power supply system and power plant scheduling and dispatches.

With the increasing development of digitization and in particular information and communication technology, energy system is becoming more intelligent. This means that large amounts of energy data are available. Artificial intelligence (AI) has already been used in a variety of areas in recent years [22,23]. DENA (Deutsche Energie-Agentur) published the roadmap for exploiting the application possibilities of AI in the energy sector, how the exciting AI method can be used in the energy systems [24]. The first application is the prediction and thus the different AI methods could be applied in the power estimation.

1.3 Motivation and Goal

As mentioned in the previous Sections 1.1 and 1.2, the main challenges of the regional renewable power generation lie in the uncertainty of time and the inhomogeneity of spatial distribution. In addition, much more renewable energy is curtailed due to the limited grid capacity and unbalance to the temporary electricity consumption. Therefore, the central question remains how to fit the temporal and spatial features for the regional RES power systems.

Regional analysis is very important to describe the difference between power generation and consumption, especially when focusing on weather-related renewable energies. Numerous power generation and load units form a regional aggregator regardless of the detailed grid connections of each grid components in order that a simplified regional network cluster is to be created. The network cluster model is used for the regional power performance analysis. It makes sense for the integration of renewable energies into the power grid system and their effects to be quantitatively evaluated in the case of the regional scale.

In addition, it is still hard to achieve a good performance for a large-scale RES power prediction. In order to obtain the regional large-scale RES power generation, the regional aggregated power curve from solar and wind energy systems should be further added up. Nevertheless, it is not feasible due to large amounts of computation and lack of characteristic data of each RES power plant. Furthermore, the regional variability in weather conditions and the distribution of power generation should be considerable. The question is how the prediction of the regional large-

scale RES power generation can be obtained. For this purpose, appropriate estimation methods for the regional large-scale power generation should be proposed. The complex relationship between weather parameters and RES power generation need to be established.

In order to improve the effective utilization of the local RES power and system stability and to avoid the long distance power transport and curtailment, the solutions of regional energy storage and sharing grid connection in the way of cable pooling are further investigated at the regional scale. In summary, the goal of this thesis is to create a regional network cluster with a suitable temporal and spatial resolution, to implement it on the regional power performance analysis and to provide applications for the regional power estimation and effective energy utilization.

1.4 Structure of the Thesis

Chapter 1 provides an overview of the development of RES in the German grid, and introduces the motivation, goal and the structure of the thesis. It highlights the importance and challenges of the RES development in the near future. This chapter also presents the central research questions that the thesis aims to answer, and explains the need for a spatial and temporal model that can account for the inhomogeneous distribution of RES on a regional scale. The characteristics of the regional grid distribution and power supply should be taken into account.

Chapter 2 focuses on the modelling of the regional network cluster and further the creation of the corresponding spatial and temporal databases. It begins by introducing the hierarchical power supply system and providing a definition of the grid-oriented network cluster model, including its characteristics. The chapter then clarifies: how the reduced network cluster model is redesigned and how the clustering approach is applied to analyse regional power performance. The reduced network structure is focused on the regional aggregated generation and consumption. By using high-resolution power data and the spatial assignment of RES power systems, the historical power difference between generation and consumption is compared and analysed. The transferable databases offers opportunities to combine the spatial and temporal datasets and data visualization.

Chapter 3 starts with a review of current research on the power estimation. The chapter proposes physical aggregated models and different data-driven models based on regional network clusters. Accordingly, the detailed methodological approaches and their implementations are explained. Subsequently, Chapter 4 outlines the experimental setups and simulation results for the regional power estimation based on neural computations and physical statistical approaches. The chapter includes the description and classification of the PV and wind power systems at

the regional scale. Additionally, a detailed account is provided of how the performance of the individual methods is evaluated using different metrics.

Chapter 5 elaborates on the implementation of the cluster model on regional virtual energy storage systems, with a particular focus on reducing power exchanges with external grids. The chapter explains the fundamental principles of the charging/discharging strategies, and provides case studies to illustrate their application.

Chapter 6 expands on the concept of cable pooling with shared grid connections, which enables a more efficient utilization of regional renewable power by using existing infrastructure. This chapter presents outcomes of both single and regional modelling, and assesses the developments from various perspectives. This framework helps identify the potential of regional PV and wind power installation and their impact on the local grid system. The potential synergy effect creates new opportunities for further RES installation and enhances grid stability.

Chapter 7 provides a brief summary of the development and application of the network cluster model. The limitations and potential future research directions are also discussed.

2 Modelling of Regional Network Cluster

This chapter proposes a method for a simplified description of regional power grid model aimed at delivering grid reduction and improving grid performance observability. The creation of the regional network cluster model mainly involves the following steps: first, the power grid infrastructure or topology is assigned to energy supply regions. Next, the database is integrated into the network cluster model. The relevant database includes essential spatial and temporal datasets for each cluster, such as geographical data, weather data, power data, and the current market master data from renewable energies. Finally, the results of RES allocation and cluster-related performance are illustrated and analysed.

The roadmap of this chapter is structured as follows. Section 2.1 gives an overview of the decentralization and conversion of power supply system. Section 2.2 describes the methodology used to create grid network cluster while considering the power characteristics of the network cluster. Section 2.3 introduces different data sources and data structure, including temporal and spatial datasets, as well as RES market master data. Section 2.4 presents the results of the cluster related performance analysis including allocation of renewable energies and the power balance between load and generation. Section 2.5 gives the concluding remarks.

2.1 Overview of Power Supply System

The generation of electrical energy and its consumption are interconnected via electrical grids. With regard to the voltage level of grid system participants (e.g. lines and transformers), the electricity grids in Germany are divided into hierarchically graded voltage levels to enable the transmission and distribution of electrical energy. The hierarchical network structure of the electricity energy system is shown in Figure 2-1 and the main voltage levels are summarised in Table 2-1. The structure of the power supply system is divided into seven levels, including the extra-high voltage (EHV) transmission system, the high voltage (HV) and medium voltage (MV) regional distribution systems, as well the low voltage (LV) local distribution grids [25]. Three further intermediate voltage levels consist of the interconnected transformers and their switchgear panels and bus bars between different voltage levels. The transmission of electrical energy usually takes place in the EHV grids, which absorb the large power generation capacity and transport the electrical energy over long distances across regions. The large power plants (e.g. nuclear power plant, coal-fired power plant, offshore wind parks) are connected to the EHV grid level. In Germany, the transmission grid system is divided into four control areas.

The four transmission system operators (TSOs) are responsible for system stability and security as well as for the transport of electrical energy in their EHV grids.

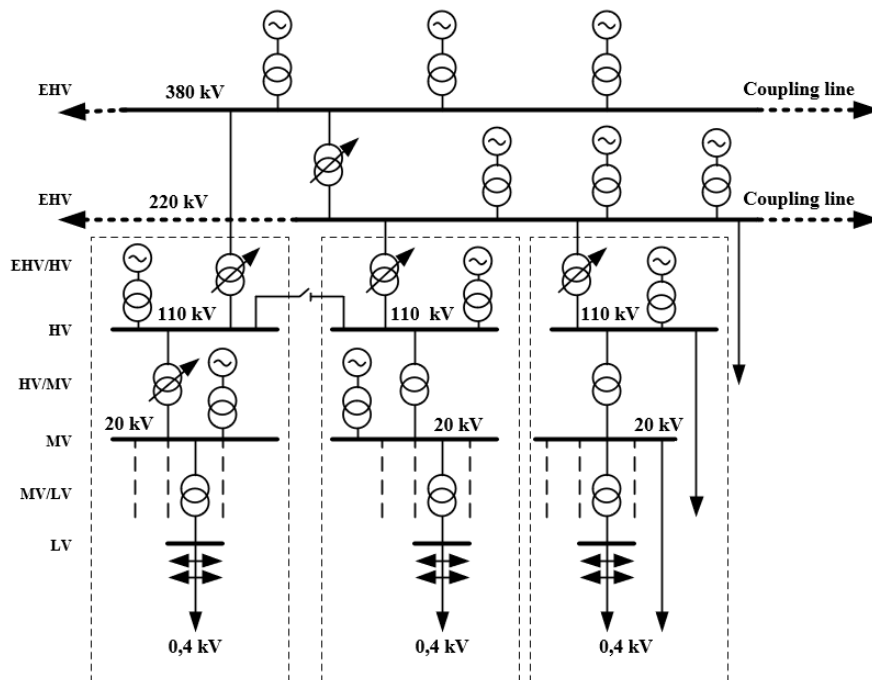


Figure 2-1: Hierarchical structure of electrical energy system [based on [26]]

Table 2-1: Voltage levels in Germany [25]

Name	Abbreviation	Voltage level	Note
Extra-High Voltage	EHV	380 kV	Transmission System Operator
		220 kV	(TSO)
High Voltage	HV	110 kV	
Medium Voltage	MV	20 kV	Distribution System Operator
		10 kV	(DSO)
Low Voltage	LV	230 V / 400 V	

The regional and local distribution networks contribute to deliver electricity to the end consumers. High-voltage grids draw electrical energy from the upstream EHV 220/380 kV transmission grid and pass on the electrical energy at their regional level to the significant grid users and to the medium-voltage distribution grids [27]. Due to the increasing share of renewable energy integration, such as onshore wind farms at HV and MV grid level,

photovoltaic systems at MV and LV grid level, the electrical energy from the distribution grid level can be fed back to the upstream grid level EHV grid.

In order to meet the demand for the constantly growing energy consumption and to reduce the CO₂ emission, more and more electricity from RES is used. The increasing integration of decentralized energy conversion plants into distribution grids leads to an effective application of RES, mostly PV and wind power. With regard to the history of electrical power supply systems, more than a century ago the electricity power supply started with decentralized isolated networks and developed to centralized mixed network with more large power plants [28]. The power supply was in a unidirectional power flow and transported from large central power generation to the end consumer [26,29]. At the beginning 21st century due to energy transition more and more distributed renewable power generation plants are installed and connected to the existing electricity grid system [28]. The grid system has been changing fundamentally compared to the traditional radial distribution grid system. With the trend towards decentralization of energy supply and intelligent grid system the concept of distributed and centralized energy system need to be considered for the future power supply system [28].

Related work to grid reduction method

A large set of power system model studies on a regional scale has received much attention and research in recent years, in order to analyse the mid- and long-term impacts of energy related policies [30,31]. Current energy system models, which use grid reduction with the help of the method of Power Transfer Distribution Factors (PTDFs), focus on the operation points [32]. Urban energy modelling is presented in [30], with the static and dynamic characterization of power generation and energy consumption in one city. A Geographic Information Systems (GIS) based platform for the allocation and optimization of the distributed energy storage system is developed in [30,33]. Small-scale modelling of the current and future energy demand and electricity generation from renewable energies is built in [17,18]. Large-scale integrated energy systems are presented in [34], the perspective of implementation is introduced and many examples of benchmark systems are used. The concept of an integrated energy system with sector couplings [35] and a multi-agent system for hybrid renewable energy sources are introduced in [36–38]. The surplus power generation can be stored or transformed to other energy forms.

2.2 Concept of Regional Network Cluster

2.2.1 Grid-Oriented Cluster Model

In order to reduce the complexity of the overall power grid system and to analyse the effects of the decentralized power generation, a regional grid network model is required. For this purpose, a simplified grid network structure is created at the grid connection points using the following reduction method. Figure 2-2 shows an example of a regional electrical network topology from low-voltage to superordinate medium-voltage and high-voltage levels under distribution grid system. The TSO and DSO grids are connected with each other over the transformers at substations. The region inside the circle is supplied from the two substations through feed-lines. The area between the two substations forms one power supply area, which is called as a grid network cluster (NC) in this work. Within the grid network cluster, there are different power generations and loads connected to various voltage levels. Nevertheless, the physical distribution topologies inside the region is not to be concerned in this work. Without considering the physical distribution topologies, the power generation and load are aggregated within the network cluster.

Figure 2-3 shows the definition of a regional aggregated network cluster, how the reduced network model equivalent to distribution grids can be redesigned. The network cluster composes the individual virtual power plants (VPPs) as well as loads from different voltage levels. In particular, the network cluster in this work focuses on the distribution grid level, into which the decentralised electricity mainly feeds. The electricity generation are from the renewable supply-dependent power plants such as PV and wind power systems and other supply-independent power generation. These components are illustrated in the Figure 2-3.

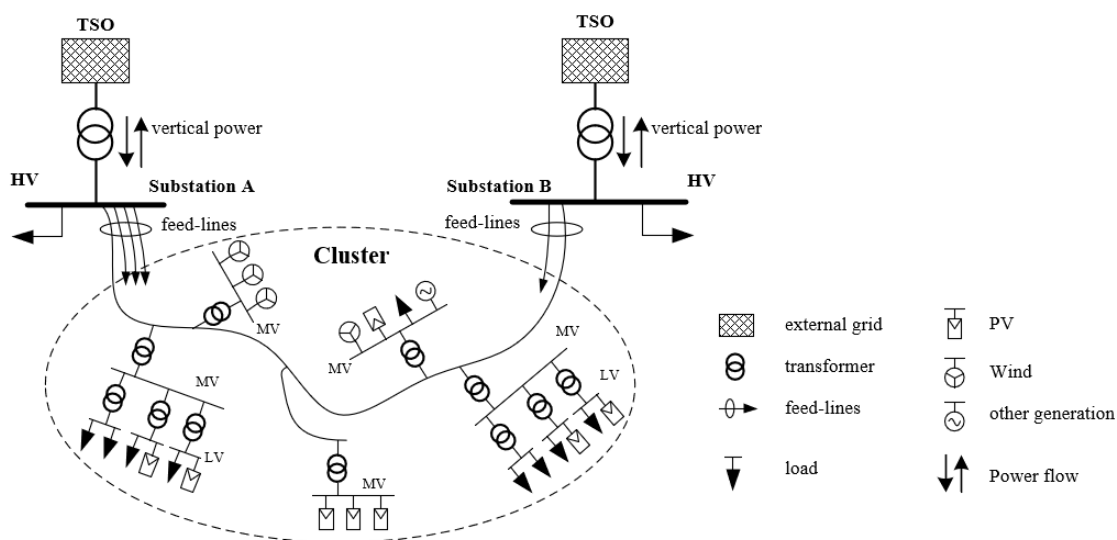


Figure 2-2: Regional electrical network topology

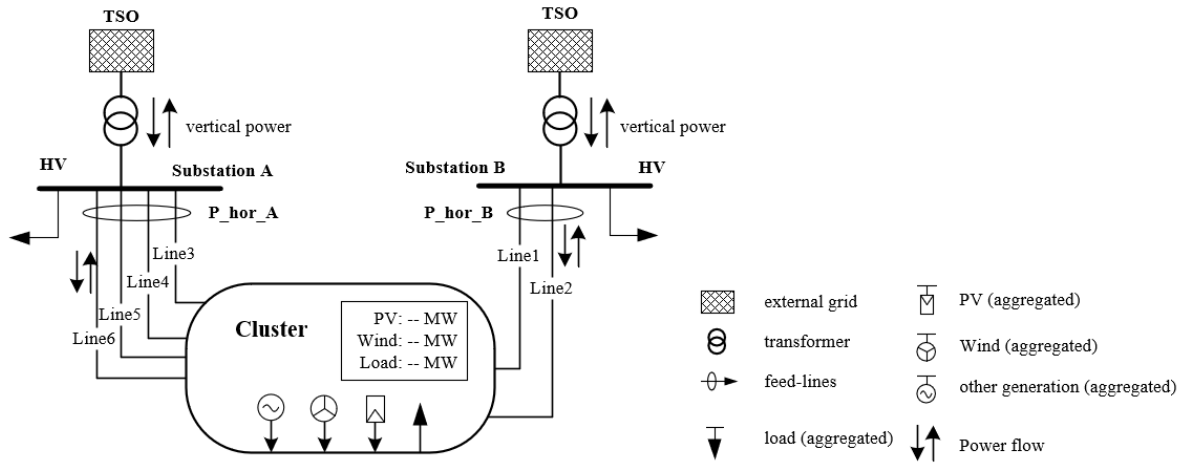


Figure 2-3: Regional aggregated network cluster

2.2.2 Cluster Characteristics

In addition to the local power generation from the distributed RES in the network cluster, as seen from the Figure 2-3, the region is as well supplied from the six feed-lines, which four feed-lines are connected to the Substation A (Sub_A) and two feed-lines are connected to the Substation B (Sub_B). The feed-line power flows can be used to calculate the power balance for the network cluster. The aggregation of power flow of the feed-lines from one substation builds one horizontal power flow for the considered network cluster. This can be formulated by Equation (2-1). Under the perspective of the electricity energy exchange with the external grid systems, the whole horizontal power flows are to be added, then the net power import for the network cluster can be described by Equation (2-2), which is also called as power saldo (P_{Saldo}^{NC}) in this work. It presents how much power is imported or exported from/to the external grids system.

$$P_{hor_j}^{NC} = \sum_{k \in Sub_j} P_{Line_k}^{NC} \quad (2-1)$$

$P_{hor_j}^{NC}$: aggregated horizontal power flow from substation j to network cluster,

$P_{Line_k}^{NC}$: power flow of feed-line k to the network cluster,

$k \in Sub_j$: all feed-lines k to network cluster from the substation j .

$$P_{Saldo}^{NC} = \sum_{all\ hor_j} P_{hor_j}^{NC} = \sum_{all\ Line_k} P_{Line_k}^{NC} \quad (2-2)$$

P_{Saldo}^{NC} : power saldo of network cluster.

Meanwhile, another perspective of the power saldo of one network cluster is focused on the power difference between the regional aggregated power generation and load. As a result, the power saldo can be as well formulated by Equation (2-3). The characterization of generation-load situation for one network cluster can be interpreted by the power saldo. The positive power saldo (equivalent to positive net power import) means that, the regional electricity consumption is more than the regional power generation. The electricity power must be withdrawn from the external grid system to the cluster. While there is too much power generation over the consumption, the surplus power generation is to be exported to the neighbour clusters or contributes to the vertical power flow to the superordinate transmission grid system via transformer substations, which are the grid connection nodes between the TSO and DSO.

In addition, the residual load is the proportion of the electricity load that is independent of the fluctuation energies [39]. For this reason, the residual load depends on the following two factors: the overall power load and the fluctuating RES power generation. The fluctuating power generation at distribution grid level is in particular from PV and wind power. Equation (2-4) gives the definition of the residual load. Therefore, the power saldo can be estimated by the residual load very well. The residual load can be positive or negative. A positive value means that the power required for the residual load must be covered by other power generations or external grid system. Negative residual load results from local regenerative power generation surpluses, which means the power generation from RES exceeds the electricity consumption.

$$P_{Saldo}^{NC} = P_{Load}^{NC} - P_{Gen}^{NC} \quad (2-3)$$

P_{Load}^{NC} : aggregated power load of network cluster,

P_{Gen}^{NC} : aggregated power generation of network cluster.

$$P_{RL}^{NC} = P_{Load}^{NC} - P_{PV}^{NC} - P_{Wind}^{NC} \quad (2-4)$$

P_{RL}^{NC} : residual load of network cluster

$P_{PV}^{NC}, P_{Wind}^{NC}$: aggregated PV and wind power generation respectively

Two aspects on the balanced power are introduced to present the characterization of one network cluster. The grid reduction procedure has been demonstrated and the regional network cluster is created. The reduced grid model is sufficient to capture the major information from the regional aggregated power. Based on the aggregated physical reality in terms of the grid infrastructures the aggregated horizontal power can outline the power balance of the network cluster. For this reason, different scenarios of the horizontal power flows can occur corresponding to the following example network cluster.

As seen in Figure 2-4, scenario 1 shows that the cluster withdraws the electrical energy from both supply substations when there is a lack of power energy for the regional network cluster. In contrast, volatile electricity generation from renewables can meet and even exceed the local electricity load. As a result, the surplus generation should be transported to the external grids (scenario 2). It should be noted that, a third intermediate state is presented in the scenario 3. The power can flow from one substation and to another substation at the case of different allocation of power generation and load sites. By comparing the two horizontal power flow, the result can give the power balance of the network cluster.

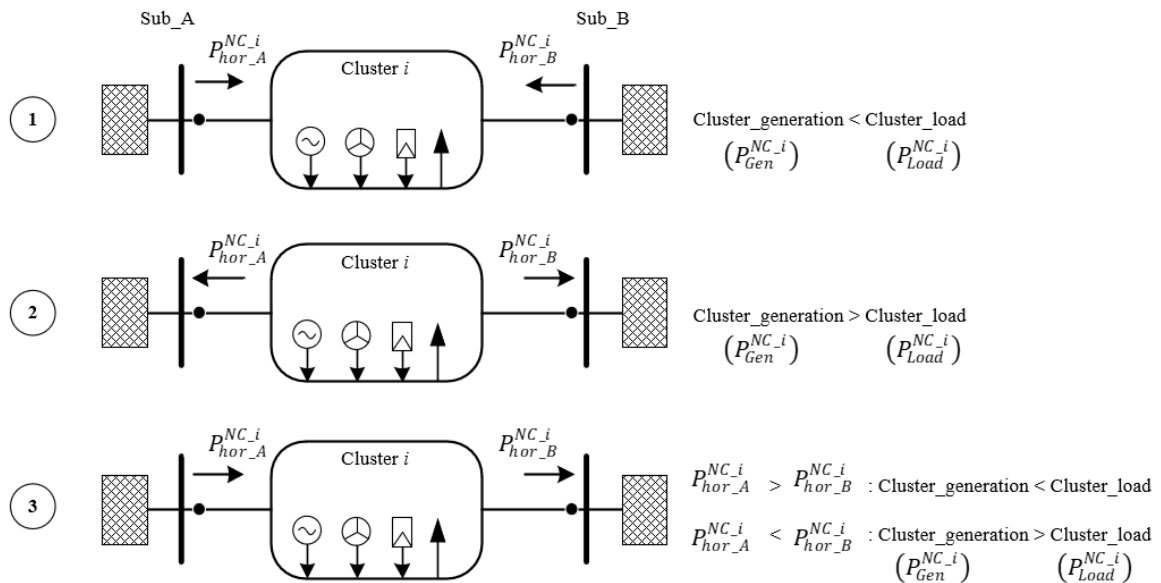


Figure 2-4: Different scenarios of horizontal power flows

As mentioned, vertical power flow refers to the power flow between different voltage levels. In this work, vertical power is based on the generated network cluster, with a special focus on the power flow at transformer substations between the sub-ordinate HV and super-ordinate EHV grid networks. The horizontal feed-lines power flows are aggregated at bus bar and the vertical power can be formulated by the following Equation (2-5). The bidirectional vertical power flow

depends highly on the regional power generation and demand as well switching states and regional grid topologies. Based on specific conditions two variants of the vertical power flow can be showcased in the Figure 2-5. Obviously, the clusters have different horizontal power flows to the substations, but the regional summation of the vertical power flows is fixed when the power saldo of the clusters is determined at a certain moment.

$$P_{ver}^{Sub-j} = \sum_{NC_i \in Sub_j} P_{hor-j}^{NC_i} \quad (2-5)$$

P_{ver}^{Sub-j} : vertical power flow to TSO at the substation j ,

$NC_i \in Sub_j$: network clusters supplied from the substation j .

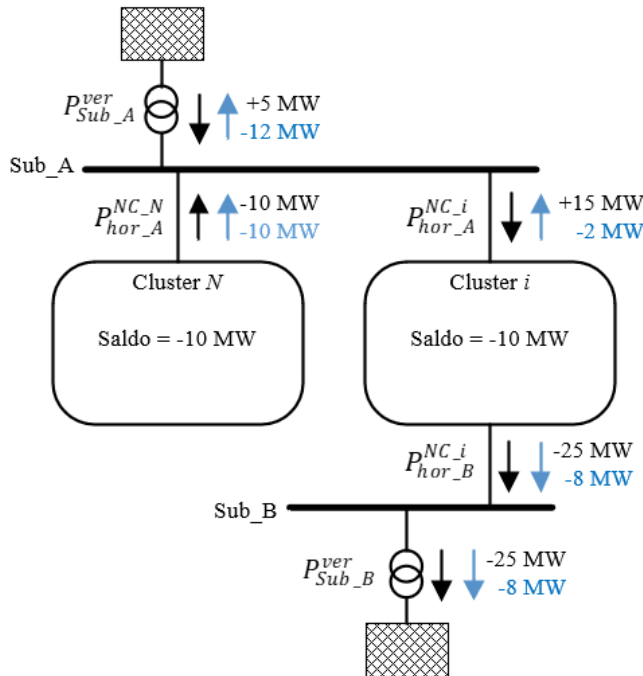


Figure 2-5: Two scenarios of the vertical power flow

2.3 Data Sources

In the following section, different data sources are described. A comprehensive data structure and concept for electrical and non-electrical energy data are carried out. First, the acquisitions of the data are introduced in terms of spatial and temporal datasets. Subsequently, data models are defined and processed. In addition, the development of RES installation master data in postal code spatial scale are created.

2.3.1 Spatial and Temporal Datasets

Geographic Information System and OpenStreetMap

Geo-data describes the geographical information of objects. OpenStreetMap (OSM) is a free project that collects and structures freely usable geo-data for specific locations. In this work, the geographic data from the OSM source are used as spatial model for the regional electricity demand and renewable energy. To query the data in OSM database, the relations with tags are used. A tag consists of two elements, a key and a value. Tag describes the specific characteristics of map elements, which include nodes (e.g. points, stations), ways (e.g. roads, lines) or polygons (e.g. boundary area). The search keys and values are to be compared with each other. "Overpass turbo" is a web-based data collection tool for OSM. Through the interface of the website "<https://overpass-turbo.eu/>" a query can be performed to produce the geo-referenced search results in the form of points, lines and polygons.

Below in Table 2-2 are some tagged search options and their returned map elements. The spatial scale with postal code area is applied in this work. Therefore, the tags with the key "boundary" and value "postal_code" are queried under administrative level of selected federal states. The following administrative level "admin_level = 4" in the Table 2-2 means the boundary of the German federal states. The definition of the administrative levels of Germany can be found in [40].

Table 2-2: Example search options with tags in OSM under a specific area

First tag:	Additional tag:	Type
key = value	key = value	Map element
boundary = administrative	admin_level = 4	polygon
boundary = postal_code	admin_level = 4	polygon
power = line	voltage = 110000	line
power = substation	voltage = 380000	point

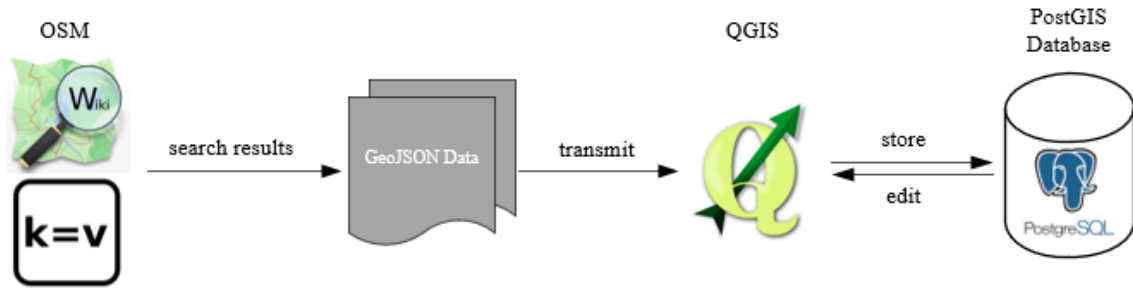


Figure 2-6: OSM data acquisition and processing

The relevant electricity grid topology with power line and substation representation is based on the real static network model, which is obtained from the OSM. Afterward, the qualified results are exported from OSM to *GeoJSON* file, which is an open format to represent geographic data. The spatial data are processed into corresponding shape file with the help of the free software from geographic information system *QGIS*. *PostGIS* database, an extension of a database management system *PostgreSQL* with *QGIS*, is used to manage the related geographic data.

Meteorological data

The fluctuating power generation from RES are significantly dependent on the meteorological condition. Regarding the weather data, they are obtained from the German Weather Service (DWD: Deutscher Wetterdienst), which offers access to CDC portal (Climate Data Center) with a wide range of climate data in different temporal resolution [41]. The real measured weather data are recorded and observed from different weather stations for the global solar irradiation, ground air temperature (2m above ground) and wind speed (10m above ground). In this work, the raw measured weather data in 10-min resolution are used. To match the time scale in 15-min the weather data is resampled based on the linear interpolation. The locations of the weather measurement stations are visualized with geographic coordinates of longitude and latitude in the CDC portal. Each weather station has a unique ID for identification. All information and the location of the weather stations can be found in the station list of DWD [42].

Power data

The power flow of lines and transformers in substation is recorded and transmitted through supervisory control and data acquisition system (SCADA) system. In the frame of Energy Industry Act (EnWG – Energiewirtschaftsgesetz) DSOs are obliged to publish their network

load and power generation for different voltage levels [43]. For case study in this work the data published by one local DSO *E.DIS Netz GmbH* are used and obtained from [44]. In addition, within the framework of the *SMART capital region 2.0* project, the DSO project partners provided the 15-min measured power flows of HV feed-lines, EHV/HV transformers and their regional RES power generations from PV and wind.

One consideration for time series data is the time zone of UTC (Eng. Coordinated Universal Time). Germany follows Central European Time (CET) as standard time, which added one 1 h to UTC (UTC+1) and Central European Summer Time (CEST) as summer time, which added 2 h to UTC (UTC+2). The time change is necessary to ensure the continuous display of the time series. To avoid the time zone inconsistency of the data source, the time is changed to the UTC+1 time zone in this work. The data range of the summer time segment is shifted forward one hour, and the data in the wintertime segment remains unchanged. The processing of the time change of the measurement data is shown schematically in the following Figure 2-7. Additionally, the meteorological data from DWD are assigned with UTC. Therefore, the time zone is as well converted to UTC+1.

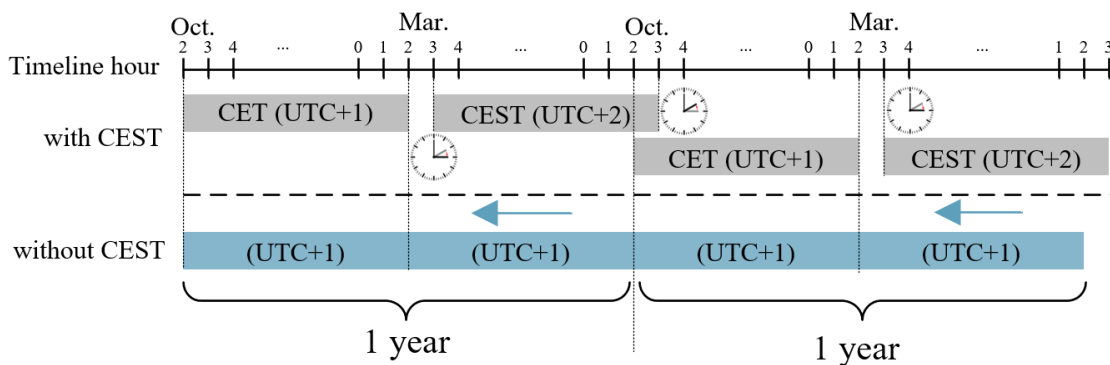


Figure 2-7: Time change according to UTC+1 time zone

2.3.2 Renewable Energy Master Dataset

In order to develop a dynamic model of the EEG system, the updated input data of RES installation are required. The evaluation of regional large-scale renewable power is based on the regional aggregated installed capacity, especially from PV and WT generators. The installed capacity is not constant and will be updated in the term of the year. The renewable energy master datasets for this work are constructed based on two data sources from Transmission System Operator (TSO) and the Federal Network Agency (BNetzA: Bundesnetzagentur). TSOs are obliged to publish the RES power plants statement reported by grid system operators in their

control area. The data are published as annual report on *netztransparenz.de*, an information platform run by German TSOs [5]. In addition, the market master data register (MaStR - Marktstammdatenregister) is the register for the Germany electricity and gas market [45]. The master data on new electricity generation plants must be registered in MaStR, which is managed by BNetzA.

The following Table 2-3 presents the number of installation of PV systems (incl. roof- and ground-mounted PV) and WT generators in Germany from the two data sources. The registered PV and WT generators from MaStR are over 1.8 million and about 30 thousands respectively. TSO published as well similar amount of the RES to the end of 2019. The amount of RES equipment installations indicates that the two datasets does not make a big difference in registered units.

Table 2-3: Data state of the two data sources to the end of year 2019

RES power system	Data source	Data source
	MaStR	TSO
Number of PV systems	1,844,767	1,849,024
Number of on-shore WT systems	28,603	28,504
Number of off-shore WT systems	1,467	1,467
Total number of WT systems	30,070	29,971

The different data sources can fulfil the incomplete information, but the duplicated information has to be eliminated. The dataset from TSOs has the complete feed-in voltage level information but the dataset from BNetzA MaStR lacks much of this information. However, the master data report from TSOs does not contain any technical parameters about the EEG plants and it can only be updated annually. The MaStR enables a significant increase in data quality and simple central registration for many energy management processes. The daily publication of the register data offers more flexibility. For this purpose, both data sources from TSOs and BNetzA are applied in this work. The two datasets are linked via the RES power system index to avoid duplicates and to supplement the missing data. The RES power system index is a 33-digit key that is assigned by the grid system operator for each RES plant for unique identification. The processing of the RES power plant register data is carried out according to the following procedures in Figure 2-8.

In addition, for data protection reasons, most accurate geographic locations of power system devices are not given for all generation systems. The available postal code is actually to be used as the smallest spatial scale data. The postal code is used to determine the location of the regional RES systems. The following data points are concerned from the registered units.

- Energy sources (e.g. wind, solar)
- RES power plant index
- Installed capacity
- Commissioning date (CM date)
- Decommissioning date (DM date)
- Postal code (ZIP)
- Federal state
- Feed-in voltage level
- Grid system operator
- Hub height of WT

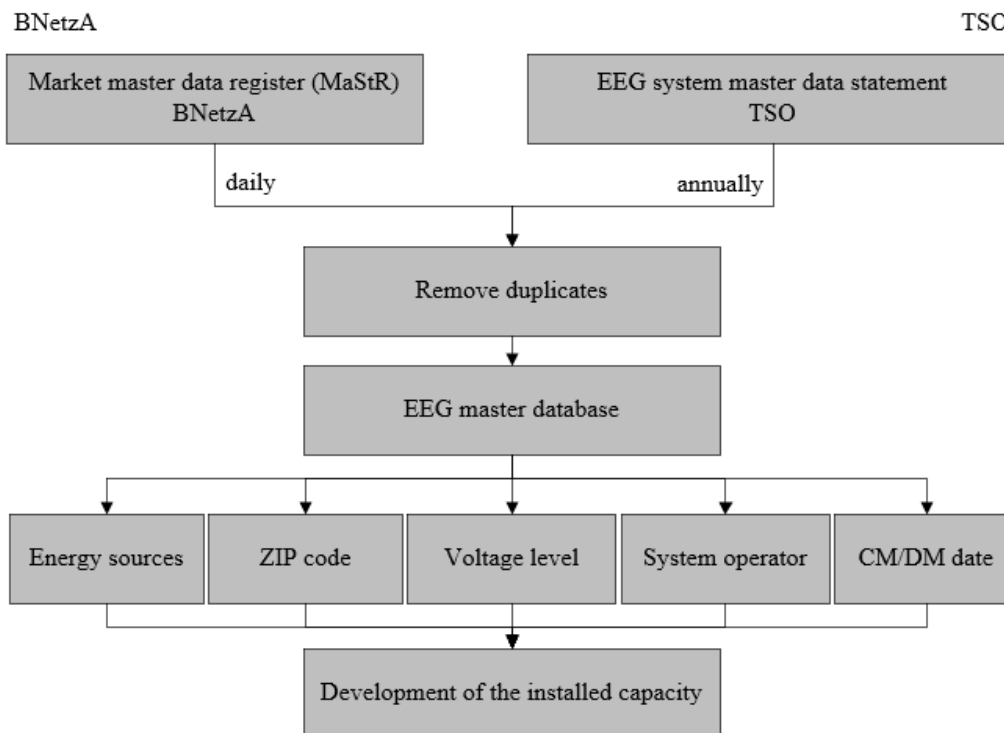


Figure 2-8: Preparation of the EEG master data

The status of the installed capacity for each postal code area will be calculated regarding to the commissioning and decommissioning date of the individual EEG power plants. The commissioning and decommissioning date of the installations are converted into availability

with 0 or 1 (see Figure 2-9). The installed capacity will be then aggregated based on the individual availabilities and capacities of EEG power plants. The installations are filtered by energy source and then created and stored in a cumulative pivot table by zip code and grid system operator. After selecting the multi-index of zip code and energy source, the installed capacities are accumulated.

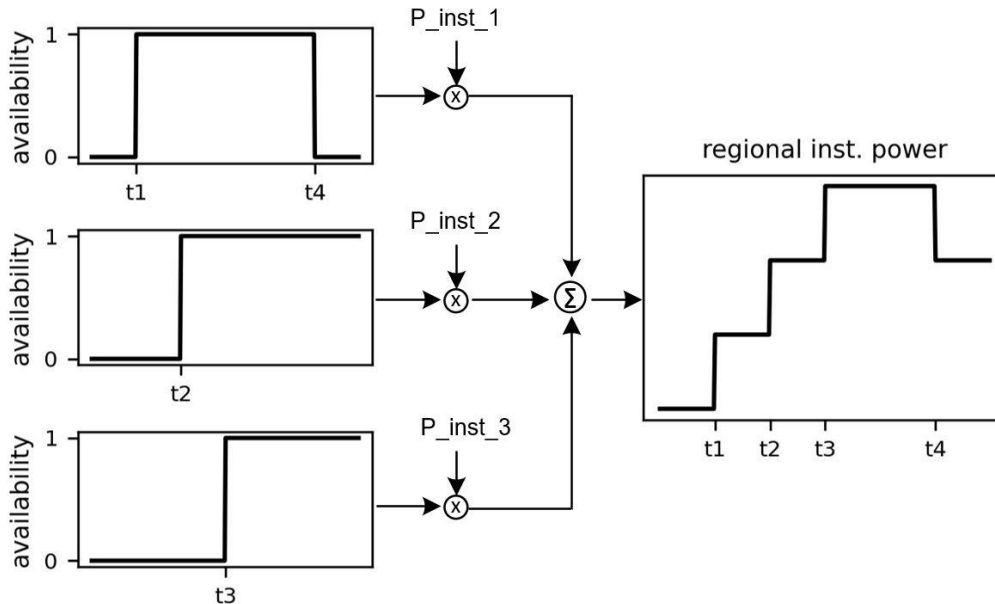


Figure 2-9: Availability period of the EEG plant system

2.3.3 Clustering of Datasets

As shown in Figure 2-10 different resource data are combined with each other through one transfer table, which includes the assignment of postcode area to the network cluster and relevant grid system operator. The *DSO_transfer_table* is derived from the RES market dataset based on the connected system operator of RES power system in each postal code area. The renewable energy forms and their installed capacities are listed according to their location information as postcode and community in the market master dataset. During partitioning the regional power grid into grid cluster the spatial dataset is required. Using the spatial shape file data from OSM HV lines, substations and boundary of postcode area are filtered and assigned to each cluster. The RES market master dataset and temporal measurement datasets are linked to the transfer table, each RES power system is assigned to the respective regional network cluster area and grid system operator. Therefore, the temporal and spatial database for the regional network cluster are created. The regionally installed renewable energies can be accessed in the form of the various spatially aggregated and temporally cumulated installations.

Based on the historical power generation and power flow of the high-voltage feed-lines, the regional clustered generation and load are derived, thus providing the basis for the observability of power generation and load on a spatial scale.

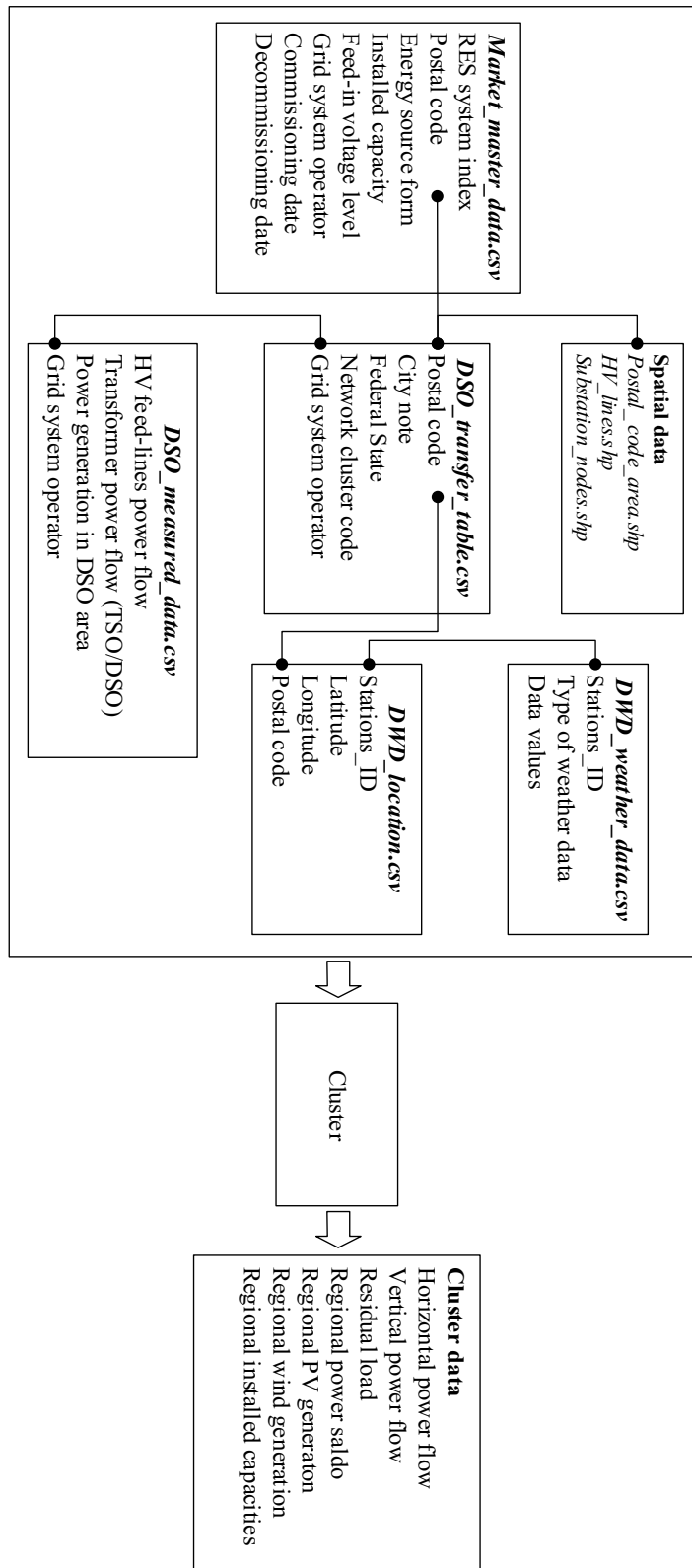


Figure 2-10: Data structure of network cluster model

2.4 Cluster Related Performance

2.4.1 Allocation of Decentralized Renewable Generation

In this thesis, a regional sub-network in West Brandenburg (Germany) is considered as a case study. The spatial electricity supply system consists of a number of regional HV distribution lines and three EHV/HV substations. The study area is partitioned into four clusters. The RES power plant systems connected to the DSO are investigated, whose voltages levels are from LV to HV. The installed RES systems connected to the extra high voltage level are not considered, because the reduced cluster model in this work is based on the regional HV lines and substations, and over 90 % RES power systems are connected to DSOs. The schematic representation of cluster with high-voltage grid lines and substations and the allocation of installed PV and WT power systems up to the end year 2019 are showcased in the Figure 2-11 respectively. Based on the RES power systems allocation, the total installed capacities of PV and WT systems in each cluster area and the ratio between installed WT and PV are calculated and listed in the Table 2-4. It can be seen that network cluster 3 has the highest installed capacities in comparison with other clusters. Over 2600 MW of PV and WT power systems have been installed in this region with four clusters areas. The graphs in the Figure 2-12 represent the development of WT installation in different spatial resolution, from the regional area to the whole Germany (DE). It can be seen that the four graphs have a similar increase trend.

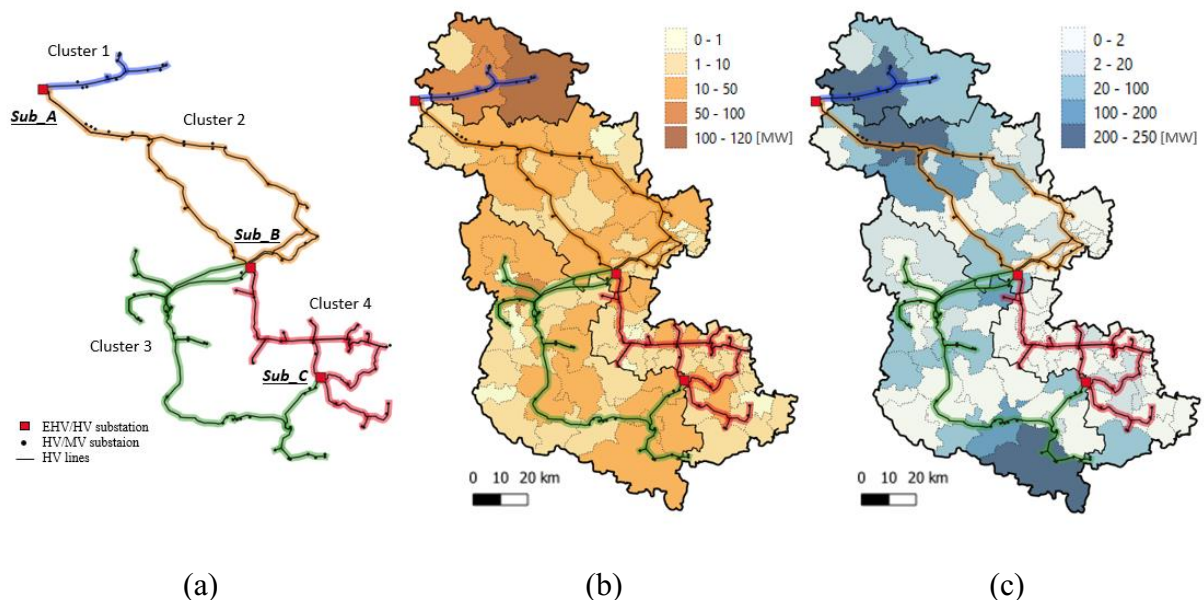
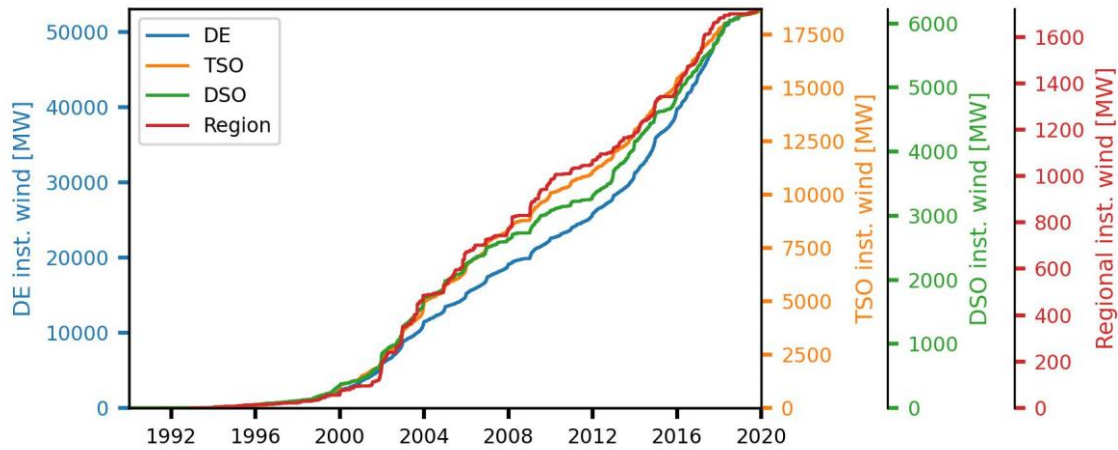


Figure 2-11: Classification of network clusters (a), allocation of installed PV (b) and WT (b) power systems to the end 2019

Table 2-4: Installed capacities of PV and WT power systems in clusters

	PV [MW]	WT [MW]	PV + WT [MW]	Ratio (PV:WT)
Cluster 1	234.0	367.5	601.5	0.64
Cluster 2	191.8	629.7	821.5	0.31
Cluster 3	378.5	629.2	1007.7	0.60
Cluster 4	141.8	95.2	237.0	1.49
Total region	946.1	1721.6	2667.7	0.55

**Figure 2-12: Development of WT installation in different spatial resolution**

2.4.2 Unbalance between Power Generation and Load

By comparing the local power generation and electricity consumption, the good performances on peak-shaving and valley-filling have been highlighted [46]. Percentile values are used to indicate the peak power for the power flow curve of saldo. For example, the power quantile P_1 is the value for which percentile q_1 of all values is smaller than this value. Moreover, the power quantile P_2 is the value for which percentile $(1 - q_2)$ of all values is bigger than this value. Based on the Equations (2-6) and (2-7) the quantity of percentile energy is to be calculated.

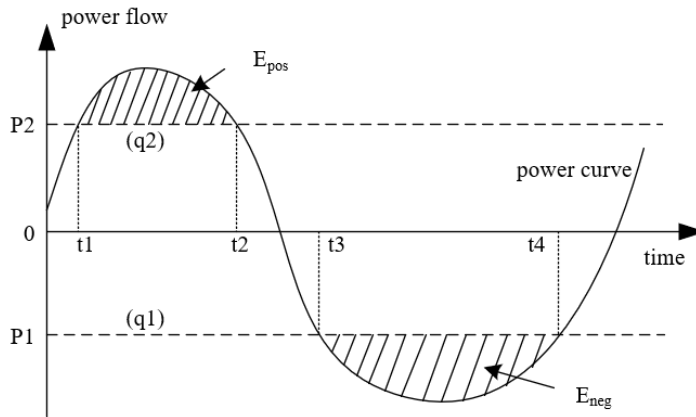


Figure 2-13: Definition of percentile energy

$$E_{pos} = \sum_{t_1}^{t_2} (P(t) - P_2) * \Delta t \quad (2-6)$$

$$E_{neg} = \sum_{t_3}^{t_4} (P(t) - P_1) * \Delta t \quad (2-7)$$

The power saldo, power difference between electricity generation and consumption, of cluster 3 is plotted in the Figure 2-14, where the power over the quantile 75 % (q_{75}) and under the quantile 25 % (q_{25}) are highlighted in colour. It is clear that in June the power saldo fluctuates more frequently than it does in December due to the different contributions of the regional PV and wind power generation. The colour area indicates the peak duration time of the power saldo. Obviously, winter has longer duration time with larger energy amount.

Some results of the regional power balance parameters are listed in the Table 2-5. The resulting demand and supply profiles can be used to investigate storage systems. The reduced positive peak power means the cluster needs less power supply from outside. The negative power saldo means generation surplus and the over-production of electricity can be stored locally in the regional network. Then by lack of power, the storage system can feed energy back into the local power grid. Storage systems connected to the power grid can improve the network operation and reduce the need for network expansion. Due to higher wind power generation in winter, December shows less frequency but higher duration time of peak power. Wind power varies less than solar power. The maximal duration time bigger than quantile q_{75} in June is about 14 hours and in December it continues more than 1 day. The energy requirement E_{pos} in December is almost 3 times than it in June.

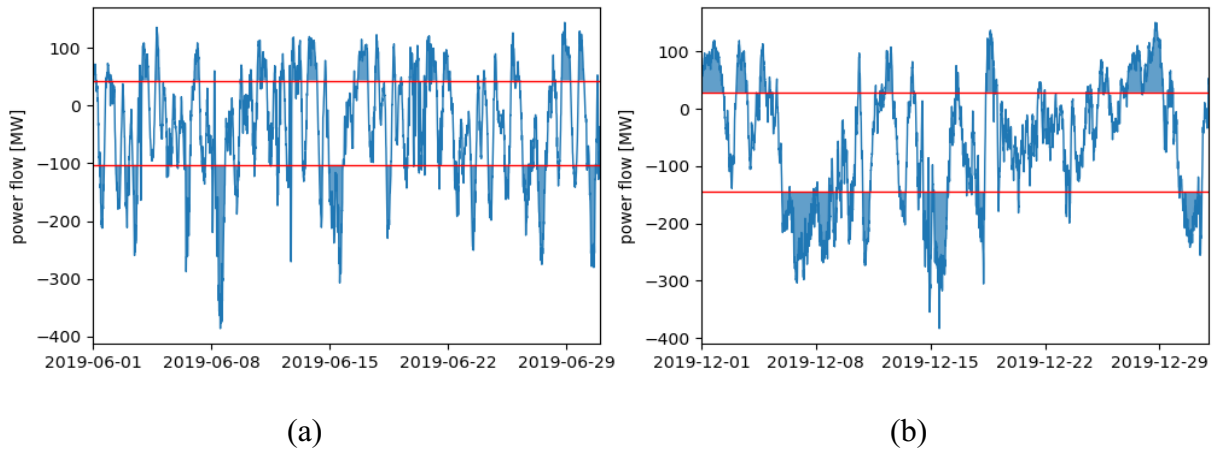


Figure 2-14: Power saldo curves of the network cluster 3 in June (a) and December (b)

Table 2-5: Results of the regional power saldo in the cluster 3

Parameters	June	December
peak pos.[MW]	144.6	150.6
peak neg.[MW]	-386.6	-383.4
E_{pos} max [MWh]	710.5	1920.4
E_{neg} max [MWh]	-2193.4	-2952.1
quantile q ₇₅ [MW]	43.6	28.5
quantile q ₂₅ [MW]	-101.7	-145.2
frequency (> q ₇₅)	57	49
frequency (< q ₂₅)	77	40
max. duration time (> q ₇₅)	14h30min	1d7h45min
max. duration time (< q ₂₅)	19h45min	1d10h15min

Figure 2-15 (a) to (d) illustrate the trend of annual average values of power saldo from different clusters over 24 hour's period. Compared with previous years the power saldo profiles have shifted downwards due to the increased RES installation. The profile of cluster 4 differs from other clusters because of higher electricity consumption and lower installed RES. Cluster 4 power saldo has the minimal power at night from 2:00 to 3:00. In comparison with cluster 4, the other three clusters have an apparent valley power between the daytime from 6:00 to 18:00. Their minimal power saldo values are negative and occur mostly in the period of 12:00 to 13:00

due to high PV power generation. Cluster 2 has a relatively low ratio of installed WT and PV systems (see Table 2-4). As solar power makes a significant contribution during daytime, the power saldo decreases during this period.

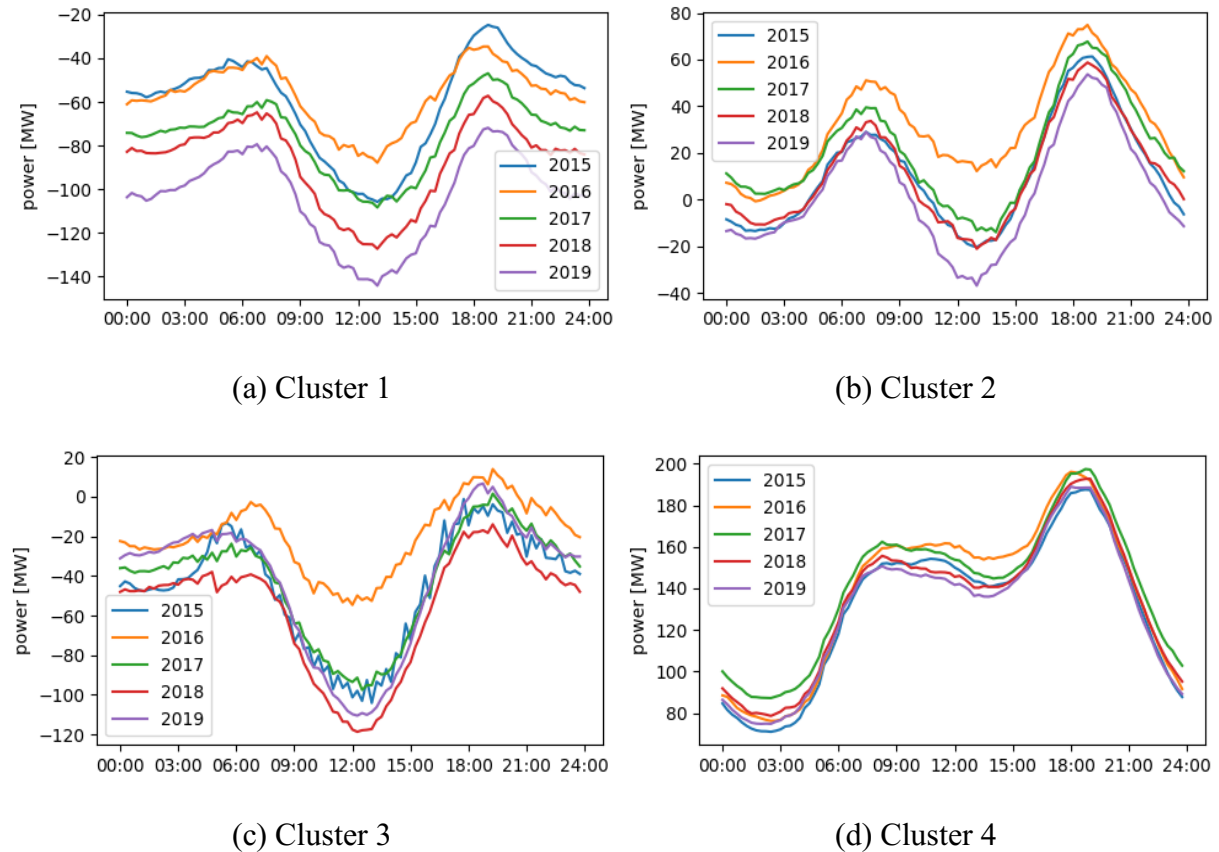


Figure 2-15: Annual average curves of power saldo

In addition, the annual duration curve of power saldo in 2019 and its frequency distribution are illustrated in the Figure 2-16. It can be seen that the maximum power saldo of cluster 4 exceeds 300 MW. A maximum of 100 to 200 MW are needed to compensate for the power shortages in clusters 1 to 3. Due to the higher RES power generation, the power saldos reach a negative value of about -400 MW, while the minimum power saldo of cluster 4 is only about -80 MW. The positive contribution in cluster 1 amounts to 40 GWh and approximately 930 GWh electrical energy surplus is generated because of less power consumers and higher RES installation for the cluster. The ratio between negative and positive energy reaches the value of 24. The energy ratio for cluster 3 amounts to 2.5. For cluster 4 it is almost 0, so most power is imported from external grids.

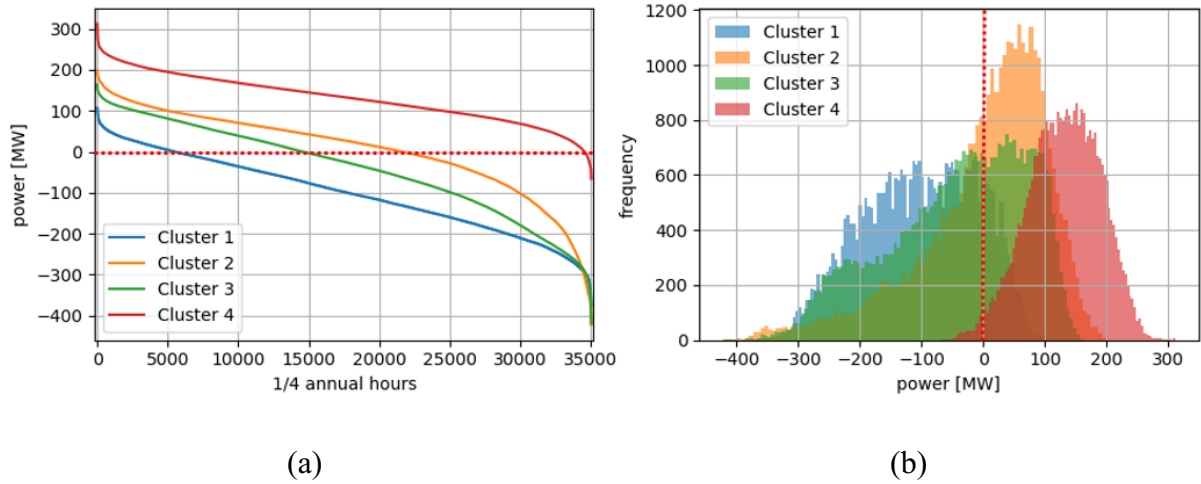


Figure 2-16: Annual duration profile (a) and frequency distribution (b) of the power saldo

2.4.3 Mutual Relations of Cluster Characteristics

The Pearson correlation coefficient of two parameters x and y , namely $\rho_{x,y}$, is a statistical index that describes the strength of the relationship between the two observed variables. It is defined as:

$$\rho_{x,y} = \text{corr}(x, y) = \frac{\sum_i [(x_i - \bar{x}) * (y_i - \bar{y})]}{\sqrt{\sum_i (x_i - \bar{x})^2 * \sum_i (y_i - \bar{y})^2}} \quad (2-8)$$

where x_i , y_i donate the variable values, \bar{x} , \bar{y} donate the mean values. The value of the coefficient ranges from -1 and 1, with the absolute value 1 representing a complete correlation, while the value 0 indicates that there is no correlation between each other.

Figure 2-17 presents the correlation coefficients of the power flow variables, which are derived from the cluster 3 with 15-min temporal resolution data of year 2019. It should be noted that the regional aggregated power generation from PV and wind $P_{PV+Wind}^{NC}$ is more correlated with other variables than the individual power generation P_{PV}^{NC} or P_{Wind}^{NC} . Especially, the aggregated RES power $P_{PV+Wind}^{NC}$ and power saldo P_{Saldo}^{NC} are highly correlated. Furthermore, the two horizontal power flows of the observed cluster $P_{hor_B}^{NC}$ and $P_{hor_C}^{NC}$ have a high correlation coefficient with its power saldo. It is known that, the possible power overproduction from PV and wind is delivered towards outside and fed back to the upstream transmission grid. Therefore, it is clear that the aggregated power $P_{PV+Wind}^{NC}$ has a high correlation coefficient between horizontal power flows and vertical power flows at substations in this region. Moreover, a high positive correlation exists among vertical power flows at different substations, which indicates

similar profile of the vertical power flows. The high-related power parameters can be used in further studies for the forecasting and filling missing measured data of cluster horizontal and vertical power flows.

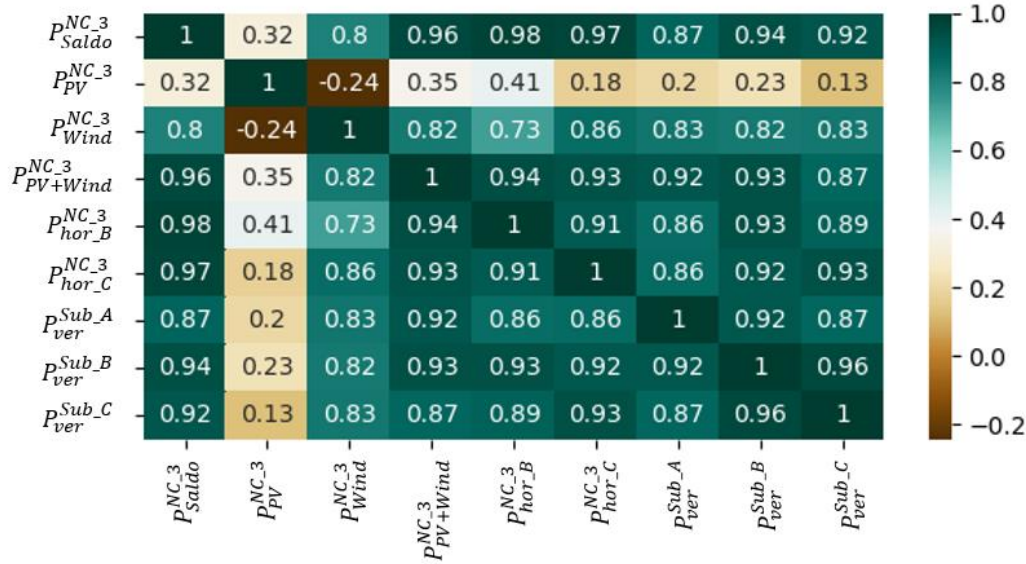


Figure 2-17: Correlation coefficient of different variables of cluster 3

2.5 Concluding Remarks

This chapter presents a method for simplifying regional energy supply systems through a cluster-based approach to a reduced power grid model for performance analysis of local power grid networks. The proposed cluster model refines grid clustering by building a grid-oriented cluster at the high-voltage distribution level. The cluster system enhances the observability, avoids a single physical power grid element, and increases spatial resolution. The proposed cluster system connects OSM spatial data, market master data and weather data. Moreover, this chapter demonstrates the regional allocation of renewable generations. By using the horizontal power flows, the characteristic of cluster energy requirements (consumption and generation) and the extent of variability of power flows are analysed. It has been found that the power saldo between consumption and generation fluctuates more frequently in summer, and winter seasons have a much longer peak duration time. The profile of power saldo differs due to the different contributions from PV and wind. With the high integration of renewable energies, the power saldo decreases gradually in the negative direction. The fluctuations result in a negative power saldo and form the local electricity overproduction, which is transported to neighbour clusters or as vertical power flow toward the superordinate voltage level. Furthermore, the correlations

of the characteristics of power flows are presented and compared. The horizontal, vertical power flows, and power saldo are high correlated.

In the next work, the created spatial and temporal datasets will be further used in the predictions of the regional aggregated PV and wind power. Power saldo is a comprehensive expression of the local power generation and consumption, and it could be used to determine the energy storage method and effective renewable energy utilization at the local scale.

3 Methodology for the Regional Aggregated Power Estimation

This chapter introduces methods about the regional large-scale PV and wind power estimations. Physical process of RES power generation and their characteristic curves address the power estimation based on the local weather measurements. PV and wind power systems are further classified according to the installed capacity, commissioning date, and connection voltage level. With the help of the proposed network cluster model and aggregated regional power curves, the regional power infeed from distributed RES systems to local electricity grids is to be achieved. Additionally, with the development of information technologies, enormous electrical and non-electrical data are collected. Furthermore, in the context of the smart energy management, different data-driven methods are applied using artificial intelligence (AI) to estimate regional renewable energy power. AI means that computers have the cognitive abilities to learn and solve problems like human intelligence through data training and learning process. Without consideration of detailed power system, AI models can capture the sophisticated relationships between inputs and outputs of multiple time series of weather information and power generation. The roadmap of this chapter is structured as follows. Section 3.1 gives an overview of the regional power estimation methods. Section 3.2 defines the proposed aggregated wind power model including wind power output characteristic and modelling procedures. Section 3.3 describes the regional PV power generation based on regional aggregated model. Section 3.4 focuses on the different artificial intelligence methods including artificial neural network, machine learning and deep learning. Different evaluation metrics of prediction results are introduced in Section 3.5.

3.1 Overview of Regional RES Power Estimation

A lot of effort in power estimation, especially for PV and wind power generation, has been carried out by researchers in the recent years. An accurate real-time power prediction is very important to mitigate the instability and enhance efficiency in grid operation. Generally, the prediction methods can be categorized into two approaches, physical approach and data-driven approach [47,48].

Physical approach describes the simulation process of energy conversion. A good power prediction performance can be achieved based on a single wind turbine (WT) model [49]. The regional aggregated power can be calculated by adding the power output from each site [21]. The power output of PV and wind system changes dynamically with their own performance characteristic curves [50,51]. The power output can be predicted based on the physical property

of power plant systems combined with the local weather conditions. However, the physical method based on single virtual power plant (VPP) system is not feasible for the large-scale distributed RES power supply systems. The reason is that, there are no complete technical details about all individual devices, such as orientation or tilt angle for solar modules, or characteristic curves for each PV and wind generation systems [48,52].

Data-driven models based on artificial intelligence (AI) or probabilistic approach can avoid the detailed information of RES power plants [52]. On the contrary, the historical measured power data and meteorological data are essential for model training, validation and testing. Moreover, in order to upscale the regional power generation, the representative power curves from measured reference power plants are used [48,53]. The unknown VPPs power output would be estimated. Spatial interpolation such as inverse distance weighting interpolation method is applied to predict the power generation in the unknown sites without measurement devices [53,54]. Indeed, the reference plants should be carefully selected.

3.2 Aggregated Wind Power Generation Model

3.2.1 Wind Power Output Characteristic

The power output characteristic (i.e. performance curve) of WT describes the relationship between the wind speed at the hub height and the electrical power delivered by the generator. A typical power characteristic curve of a single wind turbine is shown in the Figure 3-1. The wind power performance curve can be divided into four phases according to its operating features [55].

In phase I, the wind speed is below the minimum cut-in speed of a wind turbine. To enable operation of the wind turbine, a minimum wind energy is required to compensate for friction and inertial forces of the turbine, therefore no electrical power is generated in the phase I and the turbine does not start working.

In phase II at a cut-in wind speed of approx. 2 to 4 m/s, the wind turbine is switched on to convert the wind energy into electrical energy. The power output increases with the wind speed. The rated wind power is reached at the turbine-specific rated wind speed at around 12 to 15 m/s. In addition, the power output in the phase II can be simplified as linear relation [56]. Theoretically, the power curve of wind turbines is described by the 3rd power function (cubic function) of the wind speed at hub height. The maximum power that can be extracted from the wind energy is derived from Betz's law [55,57]. The power coefficient c_p defines the

theoretical maximum utilized wind power with the value of 0.593. The WT power output extracted from the wind energy can be described in the following equation:

$$P_{WT} = \frac{1}{2} * c_p * \rho_{air} * A_{rotor} * v_h^3 \quad (3-1)$$

where P_{WT} is wind turbine generator power output, c_p is the power coefficient, ρ_{air} is air density, A_{rotor} is the rotor swept area, v_h donates the wind speed at hub height.

After reaching the rated wind speed, the electrical power output remains at its rated power in phase III. Sustained high wind speed could lead to intense fatigue and extreme loads for the WT generator system. Therefore, the rotor blades are automatically adjusted to the respective wind speed, when the wind speed reaches or exceeds the rated speed. It is intended to limit the power output by means of appropriate different strategies that usually include stall and pitch regulation controls [56], in order not to exceed the rated generator power.

In phase IV, the wind speed is above the secured maximum wind speed value, the wind turbine should be switched off. No electrical power is generated even at extremely high wind speed and the cut-out speed for different wind turbines ranges from approx. 20 to 25 m/s.

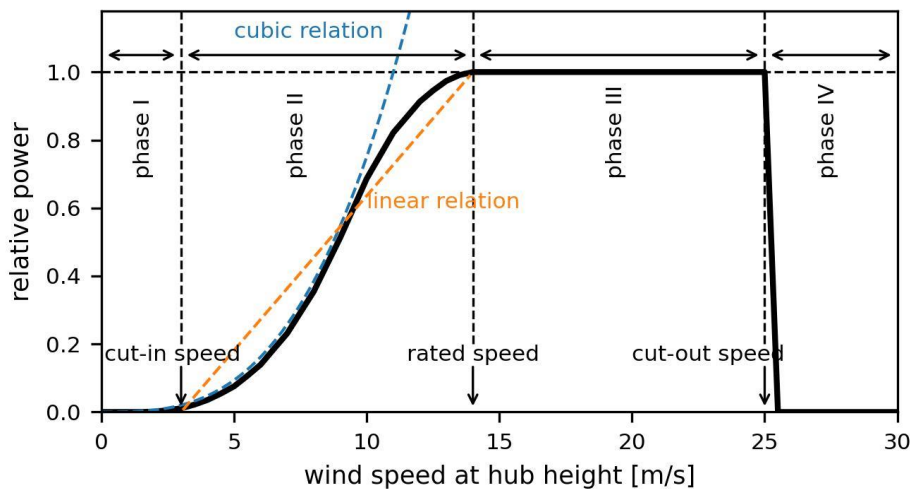


Figure 3-1: A typical power characteristic curve of one single wind turbine generator

3.2.2 Modelling of Regional Aggregated Wind Power

The regional wind power output is actually aggregated from different wind turbine power plants. They have various characteristic curves, inclusive cut-in, cut-out wind speed and slope of power

increase. The power performance curve of one single wind turbine is not sufficient for the regional power output prediction.

Figure 3-2 shows the different characteristic curves of WTs. The power performance curves in phase II vary for different wind turbine manufacturers and types. Furthermore, with the development of the new WT technologies, such as optimizing blade geometry, increasing rotor diameter, improving start-up characteristics, improving electrical efficiency and so on, the so-called low-speed WT can be also applied in areas with less wind [58]. Due to higher power efficiency, it yields high wind power, especially at low and moderate wind speeds. For the typical low-speed WT (i.e. Enercon E115/2500) with a wind speed of 10 m/s, the relative power output can reach over 90 %. Whereas with a standard WT (i.e. Enercon E82/2300), only about 60 % of the rated power can be obtained. As a result, the average full load hours (FLH) and energy yield are lower for standard WTs in comparison with low-speed WTs.

In addition to the consideration of phase II, the switch-off stage in phase IV must also be taken into account. Different cut-out wind speeds and stepwise switch-off processes are illustrated in the Figure 3-2. The different wind turbine characteristic curves for a large-scale region can be smoothed and combined into one regional aggregated characteristic curve.

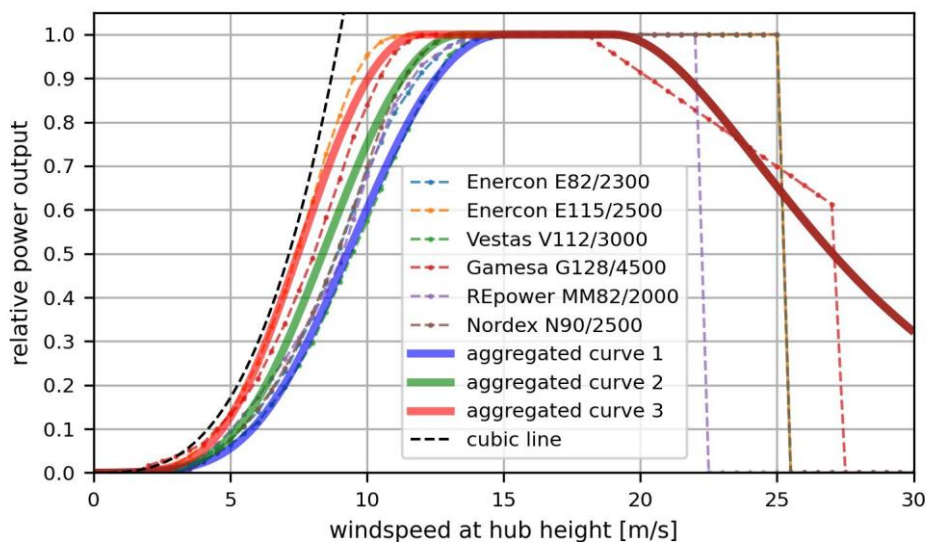


Figure 3-2: Regional aggregated wind power characteristics

Due to the different aging and types of WT, three types of aggregated curves were used to represent WTs installed in different periods. The status of installed capacity until the end of the year 2019 is analysed in this work. Aggregated curve 1 represents the WTs from 15 years ago with relatively low power characteristic (i.e. Enercon E82/2300). Aggregated curve 3 represents the WTs from last 5 years with relatively high power characteristic (i.e. Enercon E115/2500).

A third aggregated curve is generated in between. These three aggregated power curves account for the fact that different technologies were used in different periods. According to these three aggregated characteristic curves, the relative wind power output under the hub height wind speed of 10 m/s are approx. 60 %, 75 % and 90 % respectively.

Gumbel distribution is widely used to determine the distribution of extreme wind values [59–61]. One *Gumbel* probability density function (PDF) is modified and used as a model function for wind farm power curves [62]. In this work, a scaled *Gumbel* distribution function is used to model the regional aggregated wind power characteristic curve. The probability density function of *Gumbel* distribution is defined as follows:

$$g(x; \mu, \beta) = \frac{1}{\beta} * \exp\left(-\frac{x - \mu}{\beta}\right) * \exp\left[-\exp\left(-\frac{x - \mu}{\beta}\right)\right] \quad (3-2)$$

where μ is the location parameter, β is the scale parameter.

Therefore, after setting the values of location and scale parameters of the *Gumbel* distribution function, the maximum value is calculated. The maximum value of the PDF should be scaled to 1 corresponding to the maximum relative power output with a value of 1. Then the scaling process related to maximum value can be described as follows:

$$g_{scaled}(x; \mu, \beta) = \frac{g(x; \mu, \beta)}{g_{max}(x; \mu, \beta)} \quad (3-3)$$

The scaled *Gumbel* distribution function is applied for the phase II and phase IV. The aggregated switch-off phase uses a shifted descending stage of scaled *Gumbel* distribution function. The hybrid distribution function is defined and applied in different intervals:

$$p_{norm,wind}(v_h) = \frac{P_{wind}}{P_{inst,wind}} = \begin{cases} 0 & (\text{for } v_h \text{ in phase I}) \\ g_{scaled}(v_h; \mu, \beta) & (\text{for } v_h \text{ in phase II}) \\ 1 & (\text{for } v_h \text{ in phase III}) \\ g_{scaled}((v_h - \Delta\mu); \mu, \beta) & (\text{for } v_h \text{ in phase IV}) \end{cases} \quad (3-4)$$

where $p_{norm,wind}$ is the normalized power with respect to the installed wind power capacity ($P_{inst,wind}$). $\Delta\mu$ is the shifted value, it can be determined from the cut-out wind speed $v_{cut-out}$ and rated wind speed v_{rated} by the following Equation:

$$\Delta\mu = v_{cut-out} - v_{rated} \quad (3-5)$$

Through hyper-parameter optimization of μ and β for typical power characteristics of standard WT and low-speed WT, the specific parameters for the aggregated power characteristic are set as follows in Table 3-1.

Table 3-1: Parameters for aggregated characteristics

	Parameters [m/s]	Assumption for WTs
Agg. curve 1	$\mu = v_{rated} = 15; \beta = 5.8; v_{cut-in} = 2; v_{cut-out} = 19$	Group 1 (~2005)
Agg. curve 2	$\mu = v_{rated} = 13.5; \beta = 5.2; v_{cut-in} = 2; v_{cut-out} = 19$	Group 2 (2005~2015)
Agg. curve 3	$\mu = v_{rated} = 12; \beta = 4.6; v_{cut-in} = 2; v_{cut-out} = 19$	Group 3 (2015~)

Afterwards, the measured ground wind speed (10 m high) must be transferred into the specific hub height of the wind turbines, as the wind speed is strongly dependent on the height. Using the hybrid distribution function of the Equation (3-4) and parameters setting in the Table 3-1 the regional normalized wind power output is to be calculated. The most common used simplified expression of wind speed at height of h is the Hellmann exponential law [63]. The wind speed at any height can be approximately determined from the measured wind speed at a certain height. It expresses the correlation of the wind speed at two different heights, and it is described as:

$$\frac{v_2}{v_1} = \left(\frac{h_2}{h_1}\right)^\delta \quad (3-6)$$

Therefore, the speed at height h based on the reference wind speed at a height of 10 m is given by:

$$v_h = v_{10} * \left(\frac{h}{10}\right)^\delta \quad (3-7)$$

where δ is Hellmann exponent, h is the target height, v_h and v_{10} are the desired wind speed at height of h and the measured wind speed at height of 10 m respectively. According to the roughness of the ground, different exponent index values are adopted. In this work, an approximate value $\delta = 0.28$ for the exponent would be appropriate for terrain with small obstacles up to 15 m (i.e. forests, settlements and small cities etc.) [63].

In summary, the regional wind power estimation is based on the aggregation model and its process is illustrated in the Figure 3-3. The creation of the regional network cluster model and the allocation of regional wind power are presented in Chapter 2. After the classification of the

WTs into different clusters and aging groups, the clustered wind power is calculated. Accordingly, the regional power is added from each cluster.

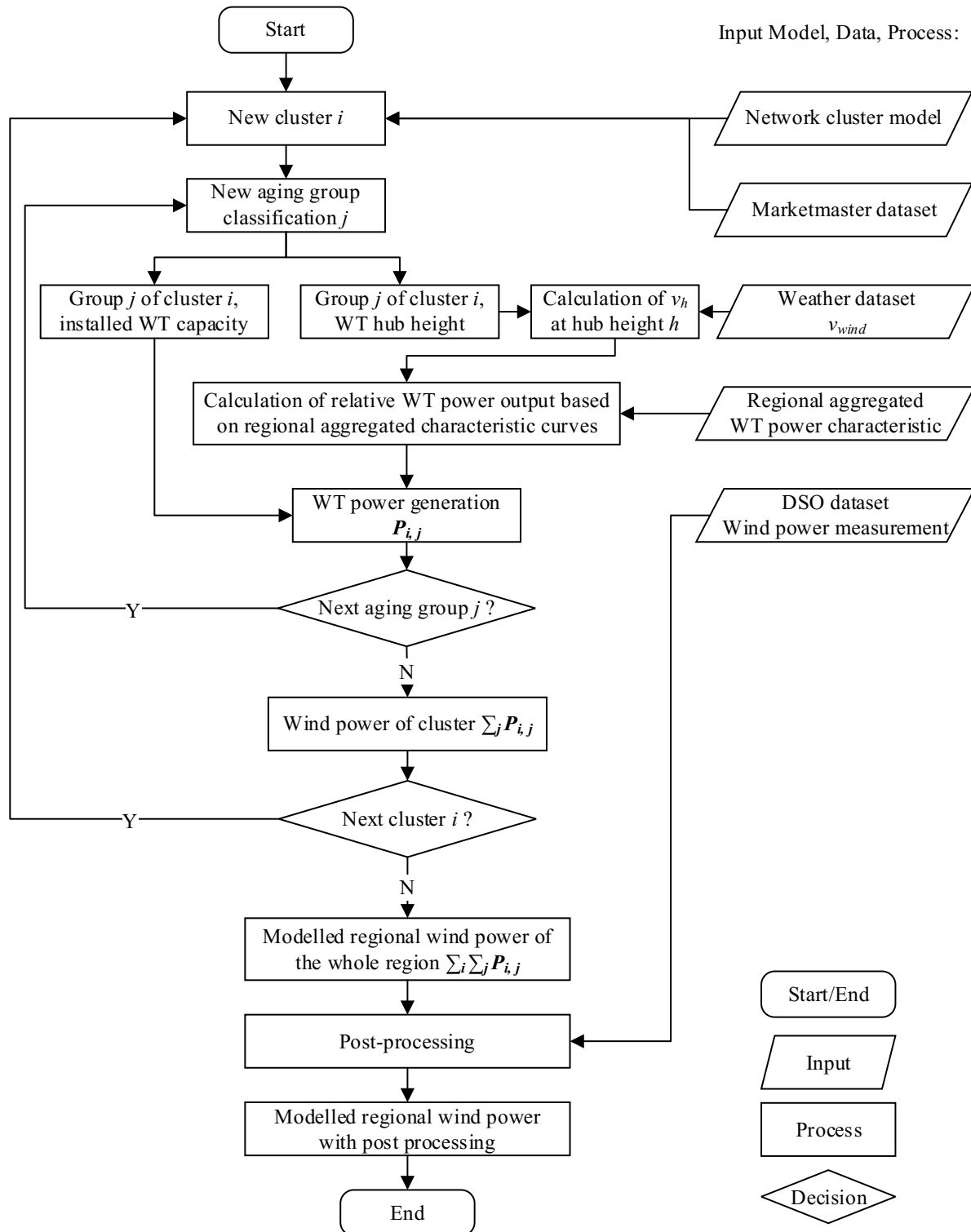


Figure 3-3: Procedure of the regional wind power generation calculation based on aggregation model

3.3 Aggregated PV Power Generation Model

3.3.1 PV Power Output Characteristic

In order to determine the electricity generation from PV systems, according to [21,64] the model is specified on the basis of the two main meteorological influencing factors, solar irradiation and PV module temperature. Figure 3-4 shows the typical current-voltage characteristic of the electrical behaviour of a solar module. The maximum power point (MPP) varies with the module temperature and solar irradiation. It can be seen that the current output is mainly affected by the solar irradiance, while the temperature change has a greater effect on the voltage output of the solar module. The maximum power decreases with decreasing solar irradiation, and at a constant solar irradiation there is a negative correlation between power output and module temperature [64].

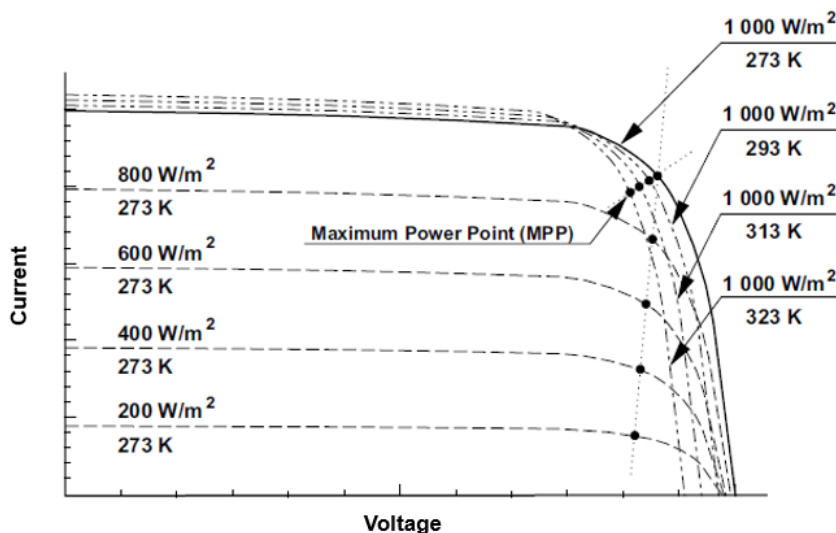


Figure 3-4: Influence of solar irradiation and module temperature on the PV characteristic [64]

In order to be able to independently compare and evaluate the different solar modules, the performance testing of solar modules is generally carried out under Standard Test Conditions (STC). The condition of solar irradiation I_{STC} is 1000 W/m^2 , the spectrum air mass is 1.5, and the temperature of solar module or solar cell T_{STC} is $25 \text{ }^\circ\text{C}$ [64]. The normalized MPP efficiency η_{MPP} of solar module can be easily determined based on the following mathematical Equation (3-8) with actual global solar irradiation I_{GS} and solar module temperature T_{mod} . In addition, the solar irradiation has a positive effect on the module temperature, which can be estimated based on a linear dependency with the Equation (3-9). The normalized PV power output in relative to the installed rated power capacity can be described by the Equation (3-10).

$$\eta_{MPP}(I_{GS}, T_{mod}) = \frac{P_{DC,PV}}{P_{inst,PV}} = \frac{I_{GS}}{I_{STC}} * [1 - \alpha * (T_{mod} - T_{STC})] \quad (3-8)$$

$$T_{mod} = T_{air} + \gamma * I_{GS} \quad (3-9)$$

$$p_{norm,PV}(I_{GS}, T_{mod}) = \frac{P_{AC,PV}}{P_{inst,PV}} = \frac{\eta_{inv} * P_{DC,PV}}{P_{inst,PV}} = \eta_{inv} * \eta_{MPP}(I_{GS}, T_{mod}) \quad (3-10)$$

I_{GS} : Global solar irradiation [W/m²],

I_{STC} : Solar irradiation by STC 1000 W/m²,

T_{STC} : Solar module temperature by STC 25 °C,

T_{mod} : Solar module temperature [°C],

T_{air} : Air temperature [°C],

α : Temperature coefficient [%/°C],

γ : Parameter for temperature rise by solar irradiation [°C*m²/W]

$P_{DC,PV}, P_{AC,PV}$: DC and AC power output of PV system [W],

$P_{inst,PV}$: Rated installation capacity of PV system [W],

$p_{norm,PV}$: Normalized PV power output [%]

η_{inv} : Inverter efficiency [%]

Table 3-2: Typical values of parameters for different module and mounting types [21,65]

Items		Temperature coefficient α [%/°C]	Parameter γ [°C*m ² /W]
Solar modules	Thin-film	0.25	-
	Poly-crystalline	0.44	-
	Mono-crystalline	0.45	-
Mounting type	Rooftop-mounted PV	-	0.056
	Ground-mounted PV	-	0.02

The temperature coefficient α depends on the type of solar module. Three main solar module types are listed in the Table 3-2. The thin-film solar module has a smaller temperature coefficient than the other solar module types and less affected by temperature. The parameter γ is determined by the mounting type of PV system. Due to the better ventilation, the ground-mounted PV system in open space heats up more slowly than the rooftop-mounted PV system. The typical values for rooftop PV systems and ground-mounted PV system are listed in the Table 3-2.

3.3.2 Modelling of Regional Aggregated PV Power

Figure 3-5 shows the characteristic curves of solar module temperature and air temperature as well as solar module temperature and MPP efficiency based on the Equations (3-8) and (3-9). First, from the diagram of Figure 3-5 (a), the PV module temperature is converted from the different solar irradiation and ambient temperature measurements. After calculating the module temperature the percentage of MPP power efficiency can be then determined based on the linear relationship at a specific solar irradiation (see Figure 3-5 (b)). The STC point is also marked in the diagram. The dash line in the diagrams presents an example with air temperature of 20 °C and solar irradiation of 800 W/m², the PV power MPP efficiency is then calculated to be about 76 %.

In order to scale up the regional PV power output, the following steps are presented in the Figure 3-6. Before the MPP efficiency conversion process, two classification of the regional PV systems are carried out due to the voltage level connections. It is assumed that the rooftop-mounted PV system is connected to the low voltage electricity grid and the ground-mounted PV system in open space feeds the electrical power into medium- and high-voltage grids. The normalized power output is derived from the corresponding calculations with cluster air temperature, solar irradiation and inverter efficiency. The regional assignment of the network cluster model and allocation of PV systems are already presented in the Chapter 2. After aggregation of the regional cluster, power output from different voltage levels the regional PV power generation is achieved.

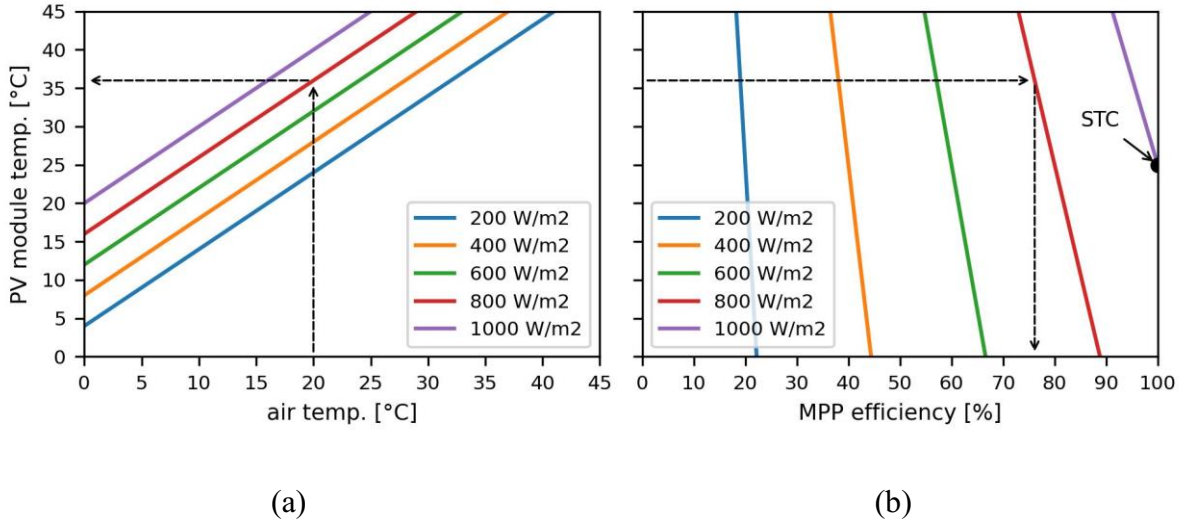


Figure 3-5: Linear dependency of module temperature with air temperature (a), with MPP efficiency (b) by different solar irradiances

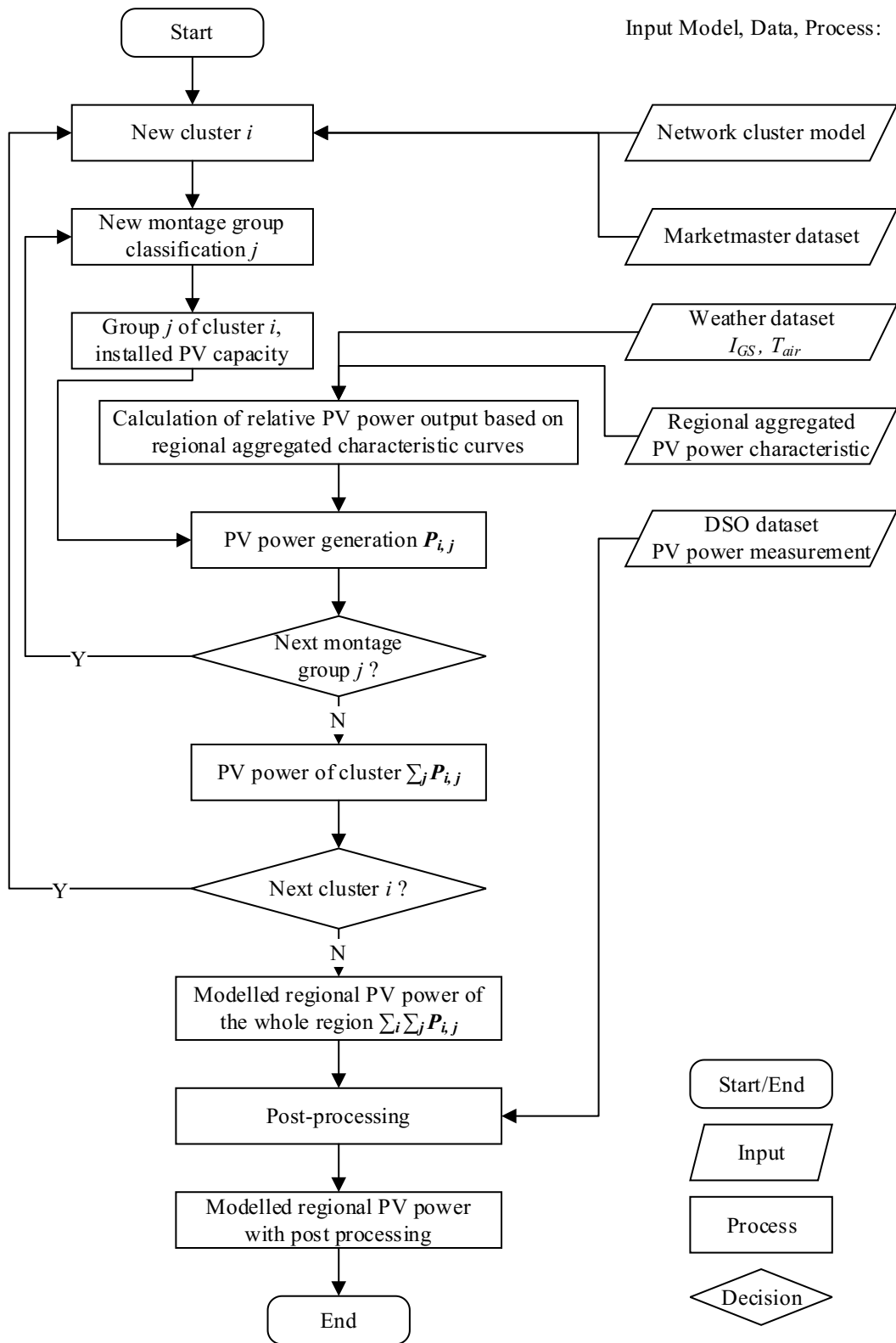


Figure 3-6: The procedure of the regional PV power generation calculation based on aggregation model

3.4 Data Driven Models

3.4.1 Artificial Intelligence

Artificial neural network method

The Artificial Neural Network (ANN) is a powerful and valuable data-driven model that can effectively extract the internal mapping relationships between input features and output target. Basically, the structure of an artificial neural network is divided into three parts: one input layer, one output layer, one or more hidden layers [66]. A typical architecture of ANN with one hidden layer is illustrated in Figure 3-7.

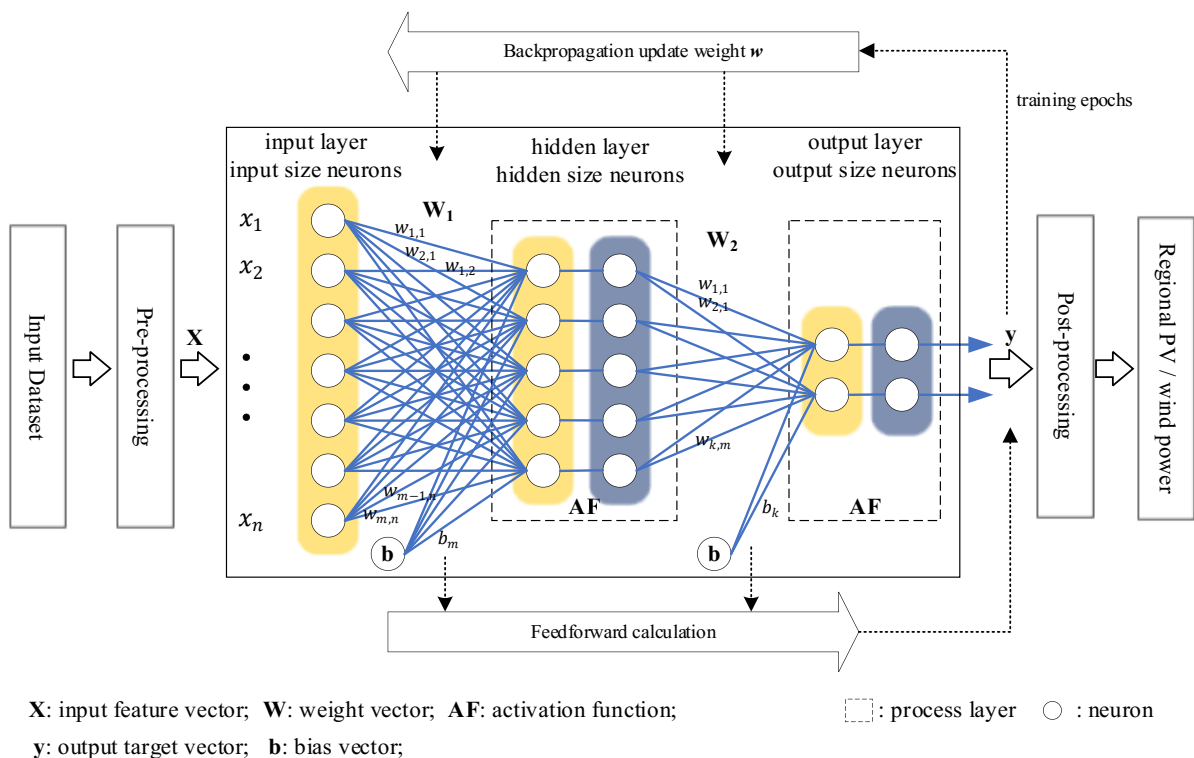


Figure 3-7: Architecture of ANN model

ANN simulates the relationship between the input features \mathbf{X} and the output target \mathbf{y} by combining a large number of computational units called neurons [67]. The description of this relationship is reflected in the weight \mathbf{W} and bias \mathbf{b} factors (i.e. vector). The parameters (weight \mathbf{W} and bias \mathbf{b}) are trained by using backpropagation. The ANN parameters are automatically and iteratively updated and adjusted based on the loss computation (i.e. mean squared error) obtained from the previous epoch (i.e. iteration) to minimize the error rate between the predicted and true target values and to make the model more reliable. The output of one process layer i can be described as following:

$$y_i = AF(w_i * x_i + b_i) \quad (3-11)$$

where x_i, y_i are the input and output of the layer i respectively. AF is an activation function. The purpose of using the activation function is to enable the ANN to handle complex non-linear problems [67]. The most used non-linear AF s are *ReLU* and *sigmoid* function [67].

For the time series data the time consistent model for the input features and output target values are used for the output estimation. The historical available datasets of \mathbf{X} and \mathbf{y} are connected with time stamp and divided into two parts, training, validation datasets. The model is trained with the training dataset and is evaluated on the validation dataset at the end of each training epoch. The prediction of target value based on the new measurement can be carried out by applying the trained model with the valid input features data in the same prediction period.

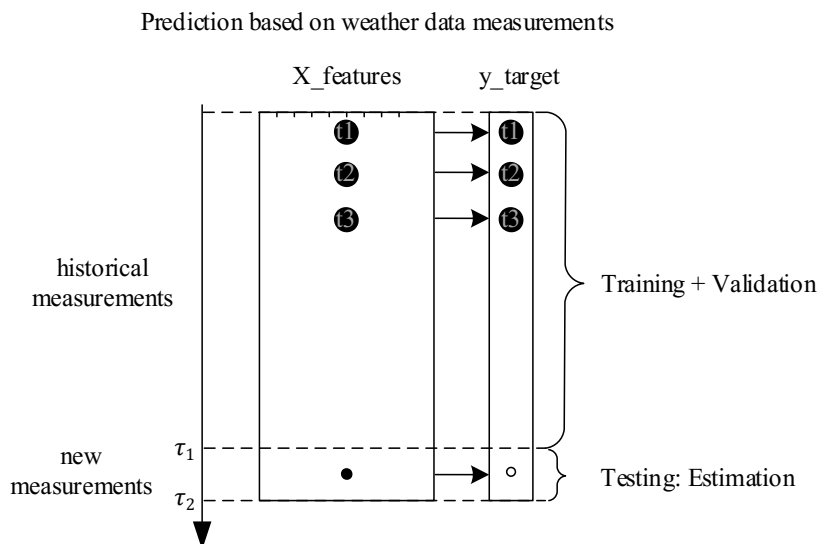


Figure 3-8: Framework of neural network for training, validation and testing

Machine learning is a branch of artificial intelligence. Machine learning algorithms analyse the data automatically to obtain patterns from them and based on these patterns to make predictions about unknown data. Two main types of machine learning statistical algorithms are described below.

Support Vector Machine

Support Vector Machine (SVM) is one of the supervised learning methods used for classification and regression [68]. The main idea of SVM is to determine a function or a hyper-plane (i.e. surface) from one or more feature vectors and attempt to minimize the error on training dataset (deviation to the surface) [68]. In fact, most problems are not linear, and a

hyper-plane, which satisfies such a condition, does not exist at all. For the nonlinear case, associated kernel function is imported to the SVM method, by mapping the data to a high-dimensional space to solve the linearly inseparable problem in the original space [69]. In this work, the default kernel function of Gaussian radial basis function is used for the regression. The theoretical process will not be described here and can be found in [70].

Multi-linear Regression

Multi-linear Regression (MLR) method aims to establish a linear relationship between some independent variables (input features) and a response variable (output target) [71]. MLR attempts to fit the linear model with coefficients to minimize the residual sum of squares between the prediction values and real observed target values. MLR model for the time series data can be explained as following:

$$y(t) = \beta_0 + \beta_1 * X_1(t) + \dots + \beta_k * X_k(t) + \epsilon(t) \quad (3-12)$$

where $y(t)$ is the predicted values, $X_k(t)$ denotes the input features, β_k is the slope coefficient of each input feature, β_0 is a constant and $\epsilon(t)$ is residual error.

With the development of deep learning technology, deep neural networks including recurrent neural network (RNN) and convolutional neural network (CNN) are widely used to deal with complex non-linear problems [72–74]. Compared with the traditional neural network, deep learning method can extract and discover higher-level features from the raw input. An extensive inner structure and numerous intermediate hidden layers are designed for neural networks in deep learning. To compare the regional RES power estimation, three deep learning neural network structures are applied in this work. One is simple Long Short-Term Memory (LSTM) model, one is stacked LSTM (SLSTM) model and the last one is combined CNN-LSTM model.

LSTM and stacked LSTM model

Among neural networks in deep learning, LSTM model is a special kind of recurrent neural network designed to process data with time series features and capture their internal dynamic dependence appropriately. The detailed procedure of the calculation can be found in [75,76]. The most significant contribution of RNN is its memory cell that is feedback loops in the

recurrent layer to store an internal state and then apply it to the next time sept sequence. LSTM can remember and connect the previous information to the presented obtained data [76]. A forget gate in LSTM solves the vanishing and exploding gradient problem in RNN [76].

Figure 3-9 shows the structures of LSTM and stacked LSTM, where the input layer receives the time sequential data, hidden layer computes with LSTM blocks and output layer returns the results. SLSTM model has more hidden layers with LSTM. As shown in the Figure 3-10 the past weather measurements (look back window) can be treated as a whole input. $X_n(t)$ donates the input feature at time t, such as measured wind speed from one measurement site n .

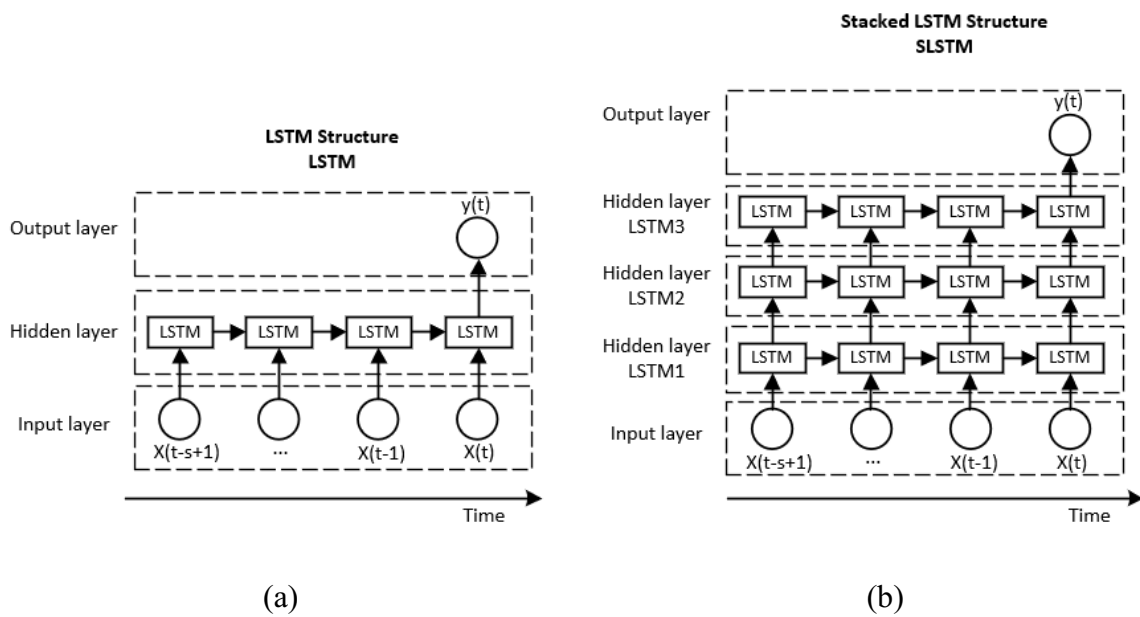


Figure 3-9: Deep learning model structures of LSTM (a) and stacked LSTM (b)

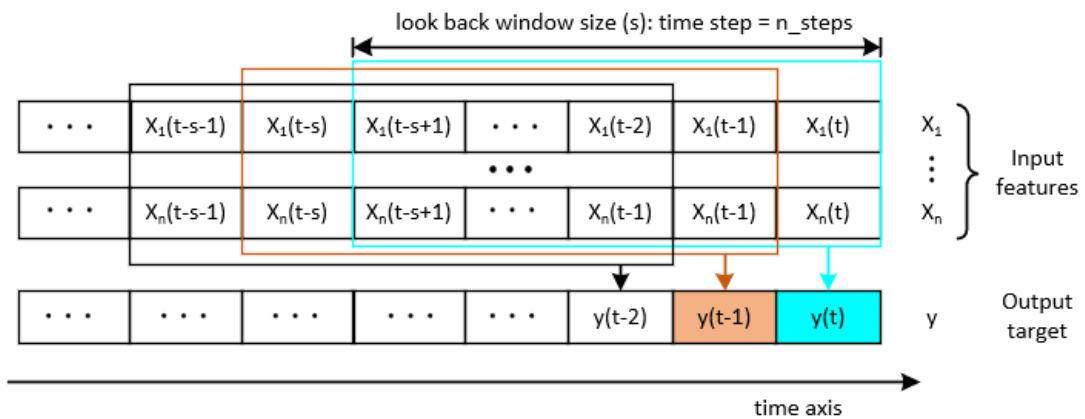


Figure 3-10: Structure of the input features and output target

CNN-LSTM model

The Convolutional Neural Networks (CNN) is one kind of feed-forward neural networks that includes convolutional computation with a deep structure consist of several layers, convolutional layer, pooling layer and flatten layer. The CNN is firstly developed for image recognizing and processing [73]. An essential benefit of CNN is their ability to use convolutional filter and pooling layer to reduce the dimensions of the input data and computational complexity [73]. In this work, a combined CNN-LSTM architecture (Figure 3-11) for regional RES power prediction is implemented. The time sequential data from different input features build one block map as a feature image. The generated 3D data structure feature maps can be received and viewed by CNN model recognition. After a range of CNN treatments, such as convolution, pooling and flatten, the output vector is to be reshaped as input for the LSTM model. The detailed and specific connection structure will not be repeated and can be found in [73,77].

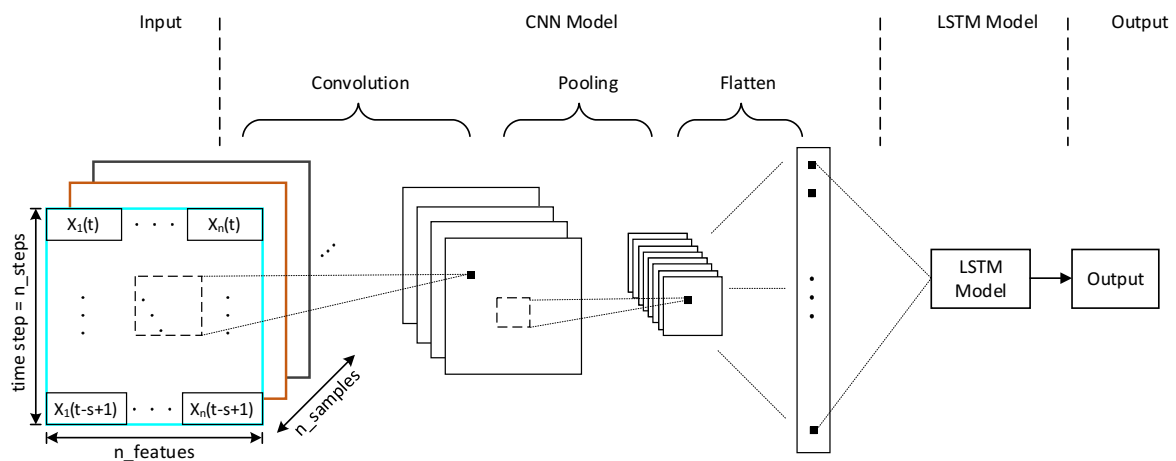


Figure 3-11: CNN-LSTM structure

3.4.2 Modelling of the Regional Aggregated Power

The sense of AI methods is that a large number of parameters (weight and bias) form memories through complex connections to capture the relationship between the input and output features. Utilizing the regional measured weather data as input features and the renewable power generation from solar or wind energy as output target. Different AI model based learning process are considered for the regional PV/Wind power prediction. A large number of regional allocation of PV and WT power systems are identified based on the proposed regional network

cluster. The spatial weather measurement stations provides the regional weather data. The procedure of the power prediction based on the models is illustrated in Figure 3-12.

Before building the model, some data pre-processing should be carried out. The plausibility and availability of the measured data are checked. The missing data and outlier of the measured weather data will be filtered and temporary treated as *NULL* values. Next, for each DWD station its nearest three valid DWD stations will be found. The locations of the DWD stations can be exported from the DWD and their nearest neighbour analysis can be carried out in the Software QGIS [78]. Linear interpolation can fill single missing data very well. However, for a range of missing data it does not work well because the weather fluctuates over time. Due to spatial short distance and the weather situation similarity, it can be understood that the weather data from near weather measurement stations are highly correlated. So the *NULL* values can be filled by using multi-linear regression method [71]. After model training and validation, the best parameters are fitted to establish the mapping relationships.

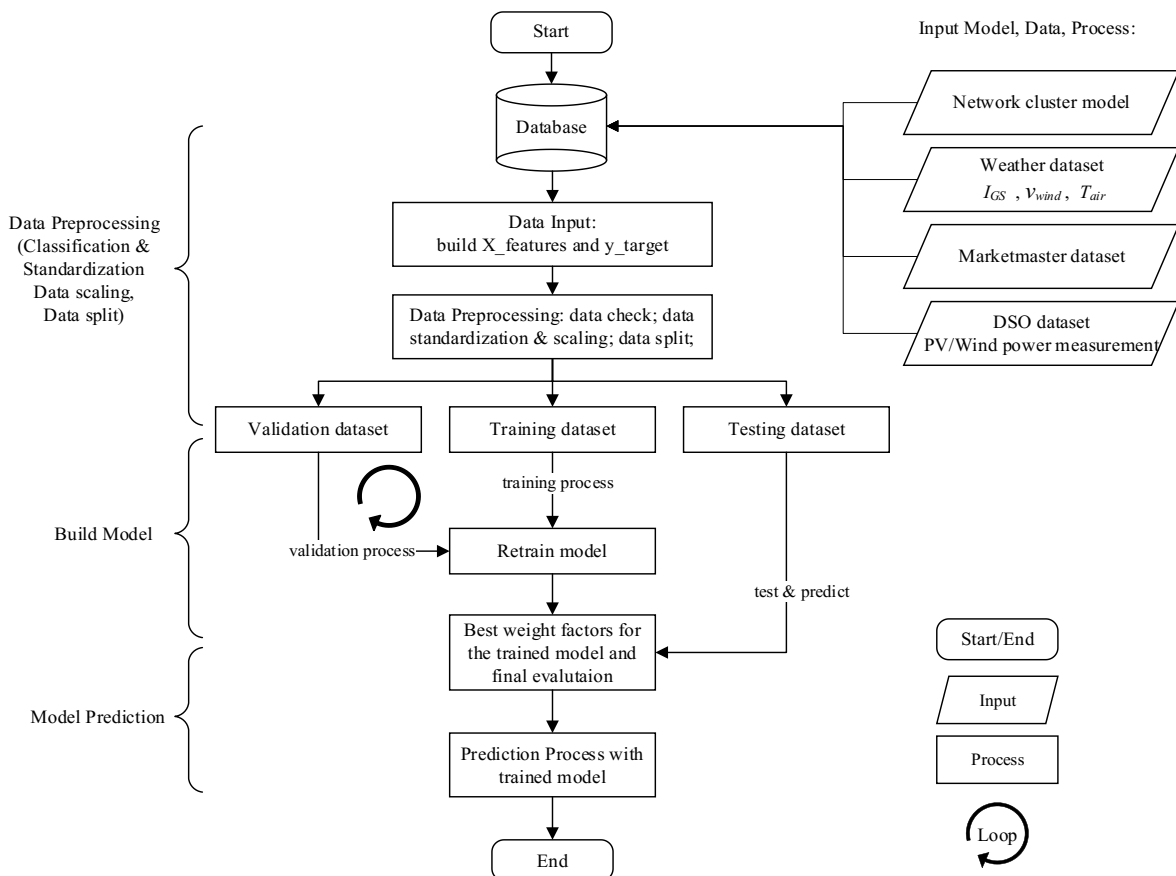


Figure 3-12: Procedure of the regional power generation estimation

3.5 Evaluation of Predictions

This section gives the concerning evaluation of predictions based on the time series of predicted and real measured data. The quality of the regional PV or wind power prediction is measured by the deviation of the predicted value from the actual observed value at each time step. The deviation can provide the information about the over- and underestimation for the prediction results. The corresponding installed capacity of the renewable energy is treated as reference to normalize the deviation result. A common metric of deviation between prediction and real measurement at each time step is called normalized prediction error $\varepsilon(t)$.

$$\varepsilon(t) = \frac{P_{pred}(t) - P_{true}(t)}{P_{inst}(t)} \quad (3-13)$$

where $P_{pred}(t)$ donates the predicted power output. $P_{true}(t)$ donates the true value of power generation from measurements. $P_{inst}(t)$ is the regional installed capacity of the observed RES. In addition, mean absolute error (MAE), normalized mean absolute error (nMAE), root mean squared error (RMSE), normalized root mean squared error (nRMSE) and mean absolute percentage error (MAPE) are the most commonly used indicators for the point prediction [51,71,79]. Especially, nMAE and nRMSE provide the overall average of relative prediction error with respect to the installed capacity. The mathematical expressions are as follows:

$$MAE = \frac{1}{N} \sum_{t=1}^N |P_{pred}(t) - P_{true}(t)| \quad (3-14)$$

$$nMAE = \frac{1}{N} \sum_{t=1}^N \left| \frac{P_{pred}(t) - P_{true}(t)}{P_{inst}(t)} \right| \quad (3-15)$$

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{P_{pred}(t) - P_{true}(t)}{P_{true}(t)} \right| \quad (3-16)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (P_{pred}(t) - P_{true}(t))^2} \quad (3-17)$$

$$nRMSE = \frac{RMSE}{\frac{1}{N} \sum_{t=1}^N P_{inst}(t)} \quad (3-18)$$

where N donates the sample size of time series.

4 Implementation of Cluster Model in Regional Power Estimation

This chapter gives the experimental setup in detail and simulation results based on the methods and modelling procedures presented in Chapter 3. Case studies of regional power estimation based on aggregated models, artificial neural networks and deep learning models are carried out. The weather data are used to model the volatile power generation from solar and wind energies. Different dimensions of the input features are taken into account. The performance of all estimation methods is evaluated based on real measured data.

The roadmap of this chapter is structured as follows. Section 4.1 presents the experimental setup and simulation results based on the different input weather parameters. Section 4.2 gives the calculation results of the regional PV and wind feed-in power estimation based on aggregated characteristic curves. Section 4.3 investigates further the data-driven models for the regional power estimation by using deep learning methods. Section 4.4 gives the concluding remarks.

4.1 Simulation Based on Neural Computations

4.1.1 Experimental Setup and Data Pre-Processing

Data preparing

In this work, a regional sub-network in West Brandenburg (Germany) shown in the Figure 4-1 (a) is taken as an investigation area. The spatial electricity supply system and their allocation of RES power systems are already introduced in the Section 2.4.1. The regional power dataset (PV and wind) is recorded and collected from the local distribution system operator (DSO), which covers the PV and wind power generation from low-voltage to high-voltage power supply systems. Figure 4-1 (b) shows the one-year data for the regional measured PV and wind power generation profiles with 35040 samples (15-min resolution).

For a large-scale energy system, this work focuses on the dynamic development of the regional installed RES. The installed capacity is not constant in the period of the year due to new installation and decommission. In the region, the installed capacity is based on the postal code level, and the new installed RES power generation plants will be updated in the term of the year based on the market master dataset. The evaluation of regional renewable power is based on the regional aggregated installed capacity of PV and WT generators. The development of the regional installed capacity of PV and wind is shown in the Figure 4-1 (c).

The weather data can be obtained from the meteorological service provider. The creation of weather dataset is described in the Section 2.3. The following Figure 4-2 shows the regional 23 observed weather stations with their station ID numbers. Some weather stations provide the measurement for three observed weather parameters, but some stations have only one parameter measurement available. Table 4-1 shows the number of DWD weather stations for each observed weather parameter. The missing data and error are detected and filtered through pre-processing. They are filled through multi-linear regression with data from the neighbouring measurement stations.

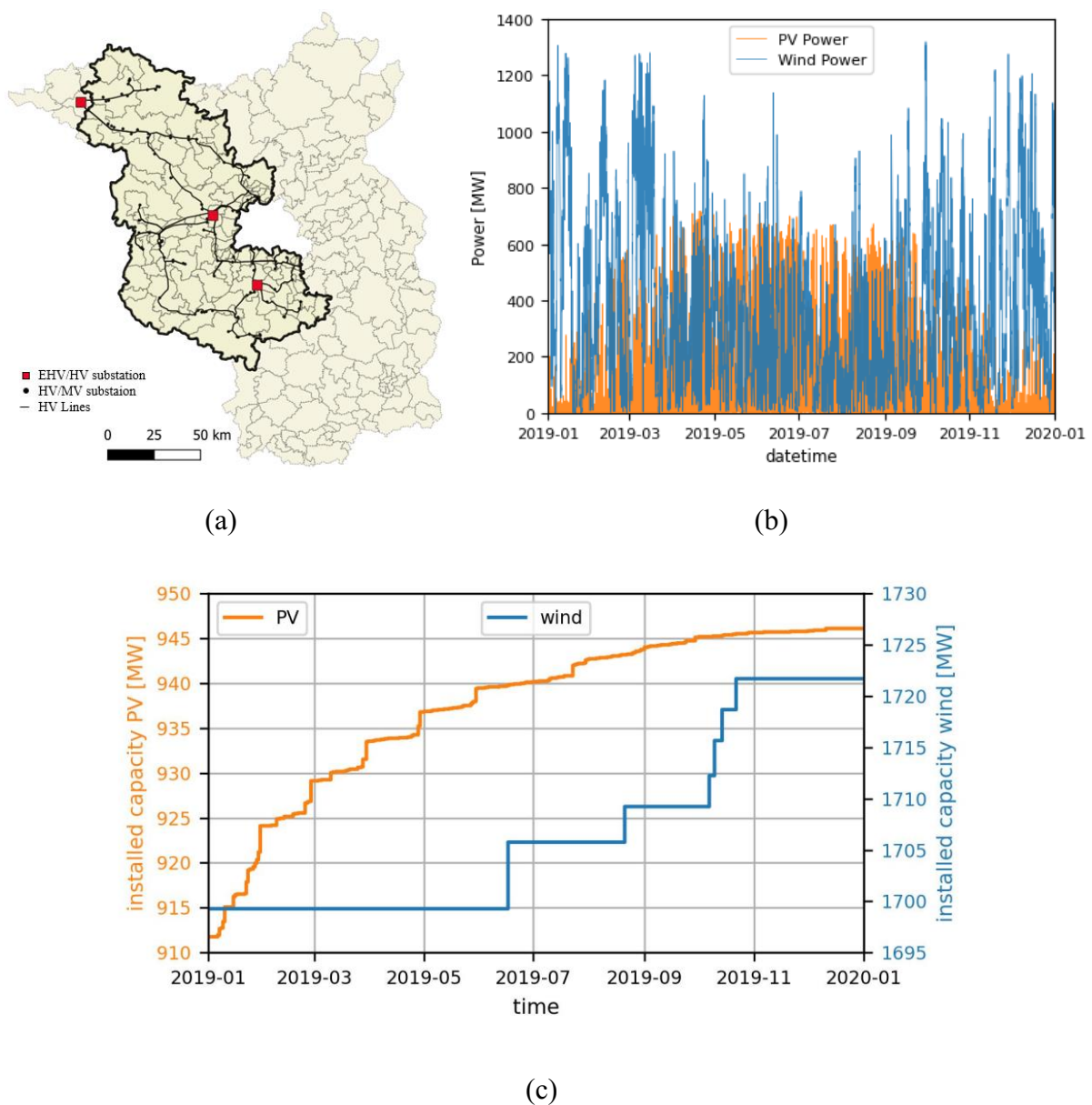


Figure 4-1: Regional DSO grids with postal code area (a), regional power generation (b) and regional installed capacity of PV and wind power systems (c)

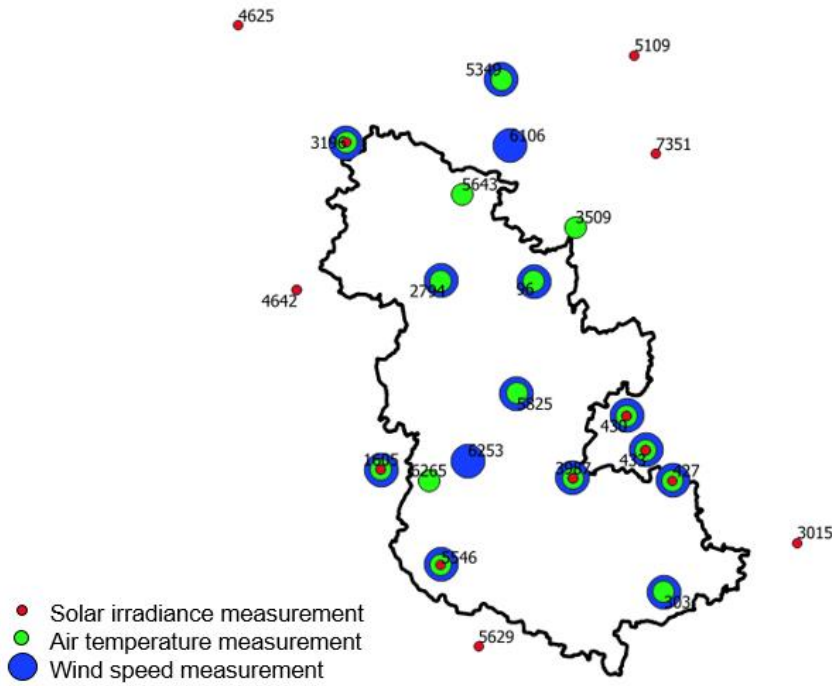


Figure 4-2: Distribution of regional DWD weather stations

Table 4-1: Number of weather measurement stations

Parameters	Number of regional observed DWD stations
Solar irradiance	$N^{GS} = 13$
Air temperature	$N^{air} = 15$
Wind speed	$N^{wind} = 14$
Total observed sites	23

Data Pre-processing

The time resolutions of weather data and the registered RES master data are 10-min and days respectively. Furthermore, the data will be resampled using linear interpolation at 15-min resolution to have the same temporal resolution with the power data. In order to obtain the input features independently, one Min-Max-Scaler is used. The input parameters (features) are valued in the interval range of $[0, 1]$, because different distribution and scale of input data could lead to large weight values and poor training results [71]. The scaled values in a small range are more suitable for training in neural network models. The transformation process for input features is given by:

$$X_{scaled}(t) = \frac{X(t) - X_{min}}{X_{max} - X_{min}} \quad (4-1)$$

where $X(t)$ donates the feature parameters, X_{scaled} donates the scaled features, X_{max} and X_{min} are the maximum and minimum values of the feature parameters respectively.

By taking into account the development of installed capacity in terms of the observed periods, the regional normalized power output can be predicted independent of the installed capacity. Although for some days or months the regional installed capacity differs not so much, but for long periods over years, this must be considered. The normalized power (i.e. $p_{norm,PV}$) is set as output parameter:

$$p_{norm,PV}(t) = \frac{P_{PV}(t)}{P_{inst,PV}(t)} \quad (4-2)$$

Similarly, the wind power is transformed to normalized power as well. The time series of power generation is normalized according to its installed capacity, in order to compare the power output curve independently of the respective scale of the distributed energy resources. For this purpose, the information can be extracted that the higher power generation is due to good weather conditions rather than due to high installed capacity. Lastly, the predicted power output will be inversely transformed with regional installed capacity to power values based on its normalization process.

Set of models and selection of input features

Figure 4-3 represents the scatter matrix with correlation coefficients of data, which include normalized PV power $p_{norm,PV}$, normalized wind power $p_{norm,wind}$, and different meteorological parameters. The correlation matrix is calculated based on Pearson correlation coefficient (see Equation (2-8)) that indicates the strength of the linear dependency between two observed variables. It is obvious that the solar irradiation I_{GS} has the largest linear correlation with PV power output compared to air temperature T_{air} and wind speed v_{wind} parameters. The correlation coefficient ρ between PV power and solar irradiation is about 0.98. Air temperature affects the PV panel temperature, which further affects the conversion efficiency of solar cells (i.e. module) [21]. The main impact of wind speed on PV power generation is reflected in the heat dissipation conditions [21]. It affects ventilation and cooling effects for the PV panels. In addition, wind speed determines the cloud movement, which

affects the solar irradiance and shadows on PV panels. The correlation coefficients of PV power to air temperature and wind speed are 0.512 and 0.145 respectively.

Obviously, the most influential factor for wind power output is wind speed, but they do not reflect a very good linear relationship. The correlation coefficient between them is approx. 0.818. In reality, the relation between wind power output and wind speed is usually more sophisticated and non-linear [80]. Wind power does not have a great linear correlation with air temperature and solar irradiation, but they indirectly affect the wind speed, air density and so on. For this reason, these two weather parameters could be extended as input features for the wind power estimation.

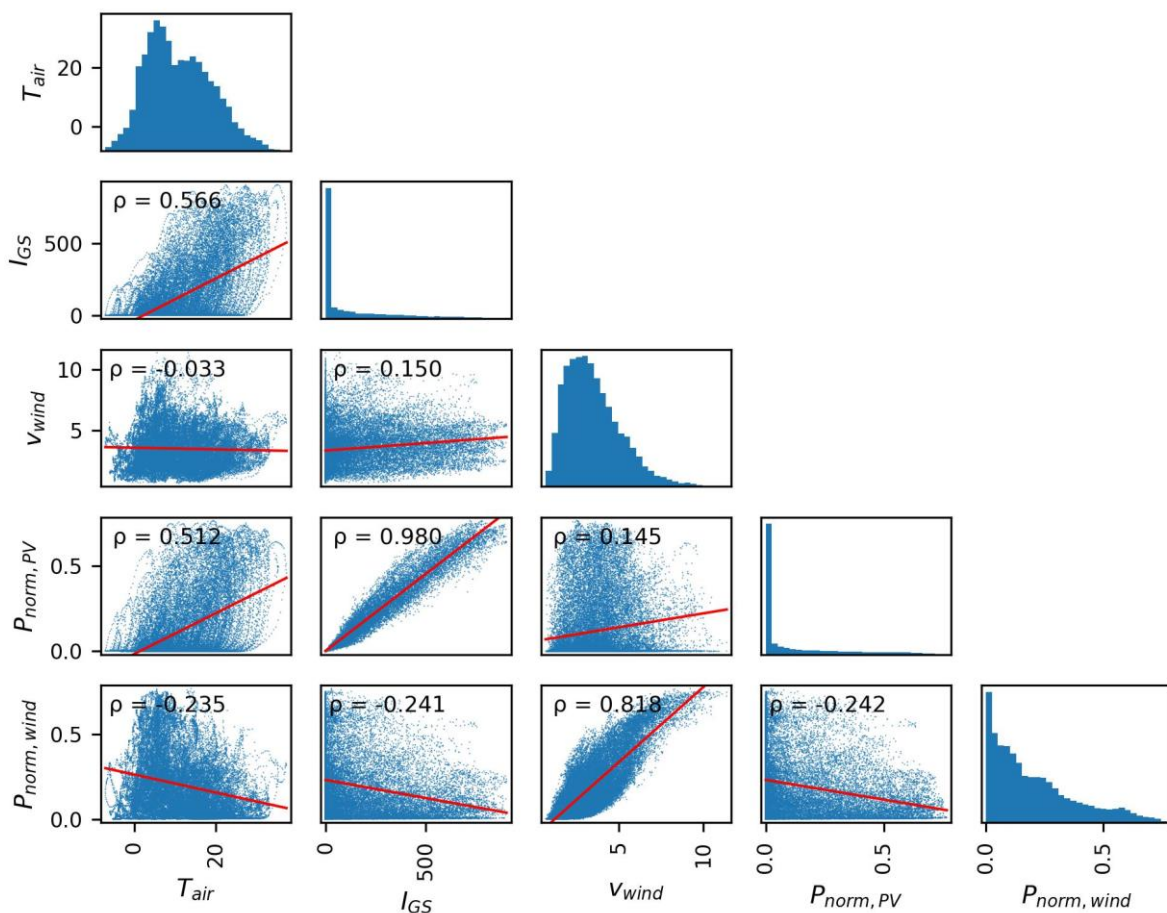


Figure 4-3: Scatter matrix and correlation matrix of data

It can be seen from the correlation coefficients in the Figure 4-3 that the weather parameters have different impacts on the RES power outputs. The selection of weather data for the regional prediction will be set as case studies. According to the input features, the input layer neurons can recognize different information from the environment; here are the measured time series of weather data.

The input features are divided into three different categories (CT). Different weather data are set as input features for AI models and are shown in the Figure 4-4. CT-1 covers the input features with the category of highly correlated weather data. PV power prediction model uses the solar irradiation of the regional weather stations, and wind prediction model uses the wind speeds as input data respectively. CT-2 and CT-3 extend the input features with two other weather parameters. The corresponding input layer units are listed in the Table 4-2 and Table 4-3, where, for example, $I_{GS,i}$ donates the measured solar irradiance from the DWD station i . R donates the observed regional area. Therefore, the sizes of corresponding input features are 13 and 14 respectively. In order to reduce the impact of the measurement of a single weather station on the power estimation of the entire region and take into account the spatial smoothing effect, three additional AI models are designed to receive the mean value of the input features of the CT-1, CT-2 and CT-3 respectively. Their input categories are named as CT-1-mean, CT-2-mean and CT-3-mean. Meanwhile, the models with mean categories have less input features and consequent computations.

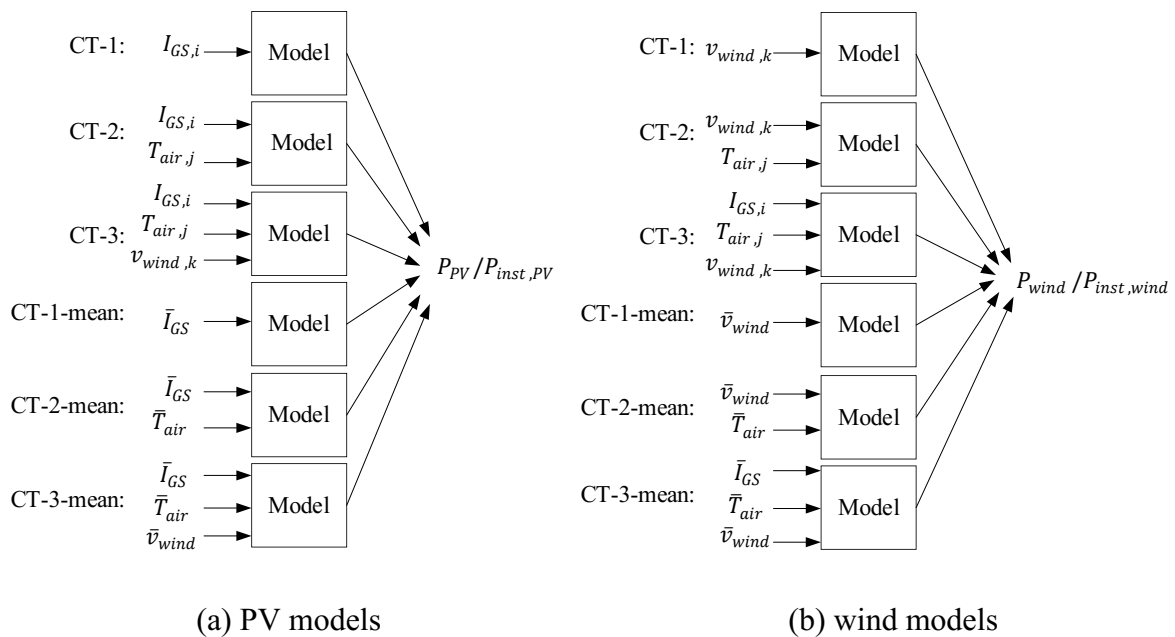


Figure 4-4: PV and wind prediction models based on different input features

Table 4-2: PV neural network based on different input category parameters

Category	Input features for PV models	Corresponding input layer units
CT-1	$[I_{GS,i}], \forall i \in R$	13
CT-2	$[I_{GS,i}, T_{air,j}], \forall i, j \in R$	28
CT-3	$[I_{GS,i}, T_{air,j}, v_{wind,k}], \forall i, j, k \in R$	42
CT-1-mean	$\left[\frac{\sum I_{GS,i}}{N^{GS}} \right], \forall i \in R$	1
CT-2-mean	$\left[\frac{\sum I_{GS,i}}{N^{GS}}, \frac{\sum T_{air,j}}{N^{air}} \right], \forall i, j \in R$	2
CT-3-mean	$\left[\frac{\sum I_{GS,i}}{N^{GS}}, \frac{\sum T_{air,j}}{N^{air}}, \frac{\sum v_{wind,k}}{N^{wind}} \right], \forall i, j, k \in R$	3

(CT: category)

Table 4-3. Wind neural network based on different input category parameters

Category	Input features for Wind models	Corresponding input layer units
CT-1	$[v_{wind,k}], \forall k \in R$	14
CT-2	$[v_{wind,k}, T_{air,j}], \forall j, k \in R$	29
CT-3	$[I_{GS,i}, T_{air,j}, v_{wind,k}], \forall i, j, k \in R$	42
CT-1-mean	$\left[\frac{\sum v_{wind,k}}{N^{wind}} \right], \forall k \in R$	1
CT-2-mean	$\left[\frac{\sum v_{wind,k}}{N^{wind}}, \frac{\sum T_{air,j}}{N^{air}} \right], \forall j, k \in R$	2
CT-3-mean	$\left[\frac{\sum I_{GS,i}}{N^{GS}}, \frac{\sum T_{air,j}}{N^{air}}, \frac{\sum v_{wind,k}}{N^{wind}} \right], \forall i, j, k \in R$	3

(CT: category)

In this work, one simple ANN architecture with three hidden layers, each equipped with 10 neuron units, is used. The size of neurons in input layer is dependent on the corresponding number of input features, which can be found in the Table 4-2 and Table 4-3. In addition, the hidden layer is activated by using *ReLU* function to filter the negative value and to obtain the process layer output in positive range $[0, +\infty]$. The output layer is activated by *sigmoid* function to ensure the output in the range of $[0, 1]$, which corresponds to the normalized power output. The parameter setting for neural network is summarized in the Table 4-4. The neural network components and structures built in this work are based on TensorFlow, which is a free and open-source machine learning and artificial intelligence platform [81]. The training epoch is set as 20 and an EarlyStopping function with patience of three is used to avoid the over-fitting.

Table 4-4: Parameter setup for ANN model

Items	Parameters
Input layer size units	According to number of input features
Number of hidden layer	3
Hidden layer size neurons	10
Output layer size neurons	1
Activation function for hidden layer	<i>ReLU</i>
Activation function for output layer	<i>sigmoid</i>
Pre-set training epochs	20

4.1.2 Comparison of Different Input Categories

Based on the experimental setup of the ANN models, the dataset of year 2019 are divided into two datasets, 70 % as training dataset and 30 % as validation dataset. Different input categories for power prediction are investigated. The profiles of the predicted and real measured PV power are compared in the Figure 4-5 for two selected periods in June and December 2019. The ANN models with raw or mean categories data can track the changes of PV power generation both in summer and in winter. Figure 4-6 shows the prediction results of wind power in the period of March 2019. The ANN models with different input features work also very well both in windy and calm periods.

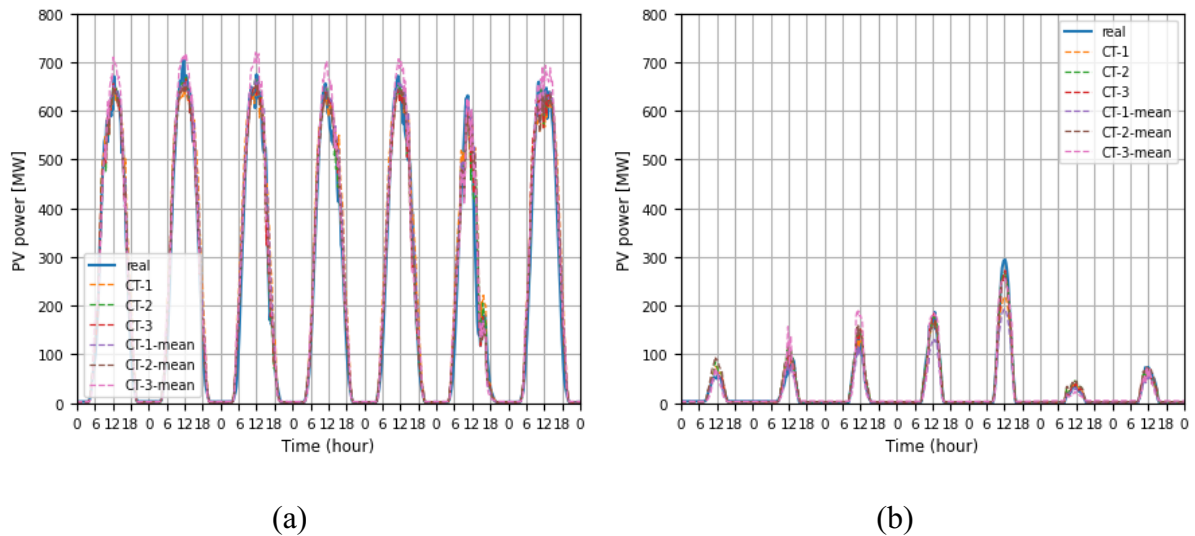


Figure 4-5: PV power prediction for 1st week of June (a) and 1st week of December in 2019 (b)

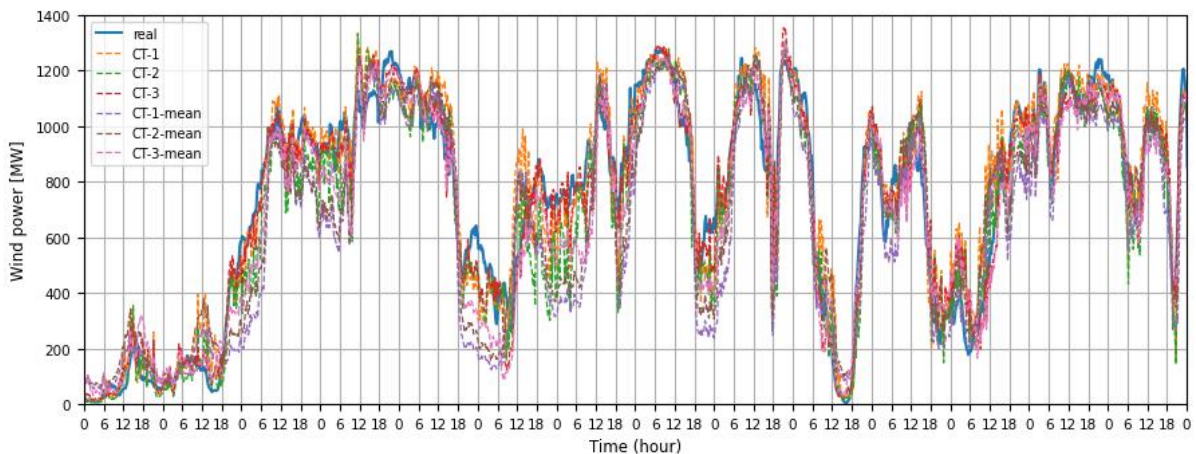


Figure 4-6: Wind power prediction for the first two weeks of March in 2019

For each experiment, the established trained model is tested with validation dataset. The detailed view on the regional power prediction error based on different input features for training and validation datasets is given in the Table 4-5 and Table 4-6. The values of MAE and nMAE for training and validation dataset for PV and wind power prediction differ not so much. This means that each model is trained very well and no over-fitting occurs. The minimal nMAE of PV trained model can reach 1.456 % relative to the regional installed capacity. ANN wind model with all raw measured weather data (CT-3) as input features performs best with 3.465 % nMAE.

By using all raw measured weather data (CT-3), the ANN model with CT-3 input features works better than CT-1 and CT-2. Increased input features lead to a better prediction result. More information from the different type of weather data can be recognized by the neurons in ANN models. The extended weather parameters can improve the prediction performance. Meanwhile,

the average data of the three categories (CT-1-mean, CT-2-mean and CT-3-mean) are used as input features. The trained model with input of CT-2-mean with only two input features results in a smaller nMAE in comparison with CT-1 with 13 input features. ANN PV model using mean solar irradiance and air temperature as input features performs better than the models using only solar irradiance as input data, whether it is the raw measured data or an average value.

Table 4-5: MAE and nMAE for PV ANN Model

PV model	Training dataset		Validation dataset	
	MAE [MW]	nMAE [%]	MAE [MW]	nMAE [%]
ANN_CT-1	17.29	1.842	17.30	1.843
ANN_CT-2	14.42	1.537	14.32	1.526
ANN_CT-3	13.66	1.456	13.81	1.472
ANN_CT-1-mean	17.07	1.820	17.54	1.871
ANN_CT-2-mean	16.10	1.716	16.25	1.732
ANN_CT-3-mean	18.79	2.003	18.98	2.023

Table 4-6: MAE and nMAE for Wind ANN Model

Wind model	Training dataset		Validation dataset	
	MAE [MW]	nMAE [%]	MAE [MW]	nMAE [%]
ANN_CT-1	95.05	5.572	95.71	5.610
ANN_CT-2	81.52	4.776	82.97	4.861
ANN_CT-3	59.14	3.465	59.91	3.509
ANN_CT-1-mean	137.54	8.056	138.37	8.105
ANN_CT-2-mean	118.02	6.913	120.60	7.063
ANN_CT-3-mean	85.20	4.990	83.94	4.916

Following this preliminary analysis, the distribution of the prediction deviation in the period of the whole year is further investigated. The cumulative distribution of normalized error is plotted in the Figure 4-7. It shows how the increasing input features with different categories affect the global performance of the prediction. Table 4-7 reports the number of hours of the over-estimation ($\varepsilon(t) > 10\%$), under-estimation ($\varepsilon(t) < -10\%$) and percentage within the tolerance ($\pm 10\%$). The distributions of the PV prediction within the tolerance of the normalized

error between $\pm 10\%$ are all over 96 %. Among them, the model with CT-3 input features gives the best prediction result. About 97 % can be achieved by ANN model with CT-2-mean input. However, wind power prediction differs from the input features. The model with CT-3 input performs well in wind prediction in comparison with other model inputs. More than 94 % predicted samples are in the tolerance ($\pm 10\%$). Meanwhile, the model with mean values of weather parameters (CT-3-mean) gives a better performance than the model with only one high-related parameter as input feature (CT-1).

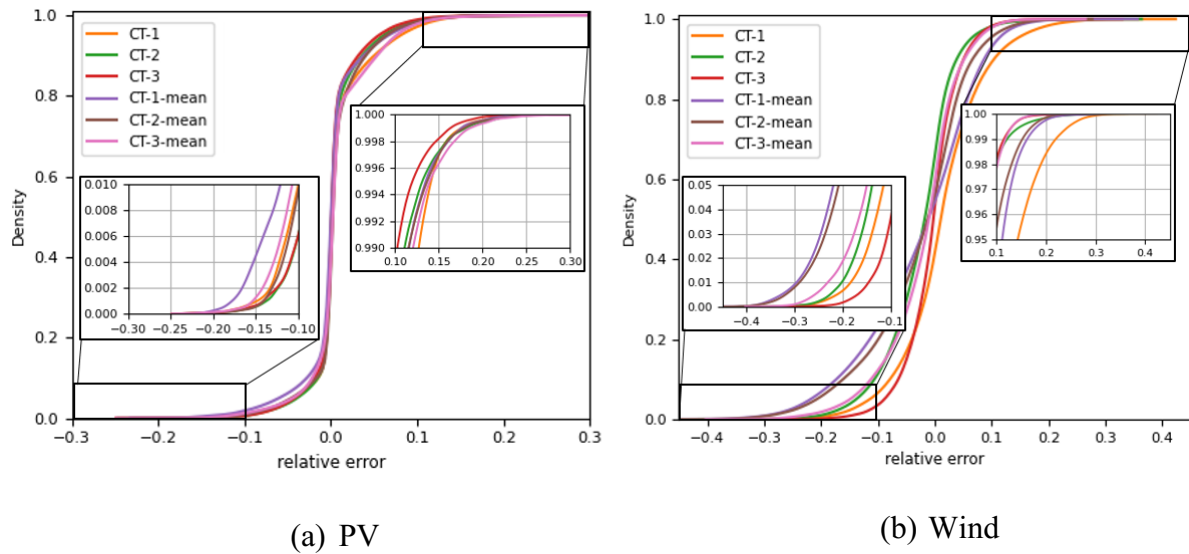


Figure 4-7: Cumulative distribution of normalized error

Table 4-7: Normalized error distribution

$\varepsilon(t)$	PV prediction			Wind prediction		
	$> +10\%$ [h]	$< -10\%$ [h]	$-10\% \sim +10\%$ [%]	$> +10\%$ [h]	$< -10\%$ [h]	$-10\% \sim +10\%$ [%]
ANN_CT-1	240.25	84.75	96.29	933.50	562.00	82.93
ANN_CT-2	125.25	52.25	97.97	160.50	980.00	86.98
ANN_CT-3	97.00	51.00	98.31	152.00	311.75	94.71
ANN_CT-1-mean	159.50	165.25	96.29	582.75	2011.00	70.39
ANN_CT-2-mean	140.00	80.25	97.49	376.25	1727.00	75.99
ANN_CT-3-mean	185.25	105.00	96.69	178.00	1069.75	85.76

Furthermore, Figure 4-8 shows the sensitivity of prediction error in different seasons. Summer has a higher value of nMAE than winter for PV power estimation. In summer, the PV power generation is higher than in winter, so the large prediction deviation leads to bigger nMAE.

Another reason may be the complicated local weather changes in summer. The overall performance of wind power prediction is more stable throughout the year for each input. In general, the combination of the three types of weather parameters gives better prediction in PV and wind prediction in different seasons. The work presented in [21] compared an up-scaling model based on reference measurements and a persistence model using last day data to estimate the current PV power for different months. Their estimation errors maintain 4 % and up to 8 % respectively.

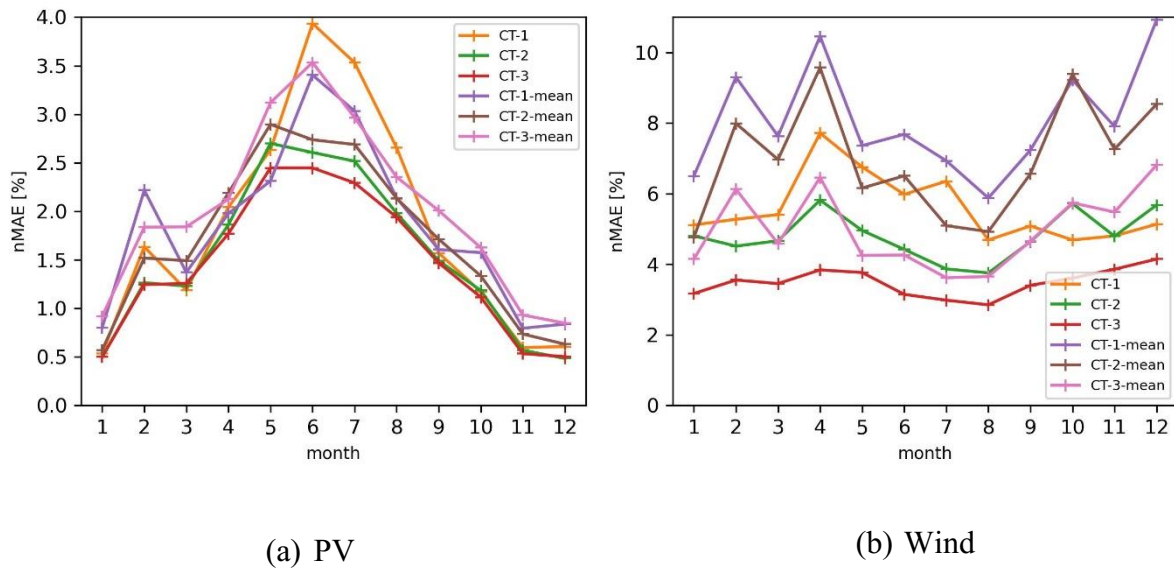


Figure 4-8. Monthly nMAE of ANN prediction based on different input features

In order to investigate the effectiveness of the proposed ANN method, two other widely used machine-learning methods are imported. One is machine-learning method of support vector machine (SVM) regression and another is multi-linear regression (MLR). The SVM and MLR used the same training datasets as the ANN model to fit the relationship between the input features and output target.

Figure 4-9 shows that the prediction approaches of ANN and MLR have very good performances by using the raw measured weather data (CT-1, CT-2, CT-3) and the reduced mean values of input categories features (CT-1-mean, CT-2-mean, CT-3-mean). In contrast, SVM cannot fit the PV prediction model very well. Because of the inherent characteristics of PV power generation, almost half of data values for the input feature of solar irradiance are zero during night time. This can lead to difficulties in fitting a suitable SVM hyper-plane for PV power regression, in particular at night time. However, SVM models with additional parameters (CT-2, CT-3) can mitigate the errors and have less nMAE than SVM model with CT-1 input features. The reason is that the performance of SVM depends highly on the features, samples

and kernel functions [68–70]. When focusing on the MAPE, which can give a view on the prediction in comparison with the real predicted samples. All PV prediction models have large values of MAPE, especially for SVM models. That is because the denominator of MAPE Equation (3-16) is very small at night time ($P_{true}^{PV}(t) \cong 0$) for PV power. A little deviation for small values can result in a bigger value in MAPE.

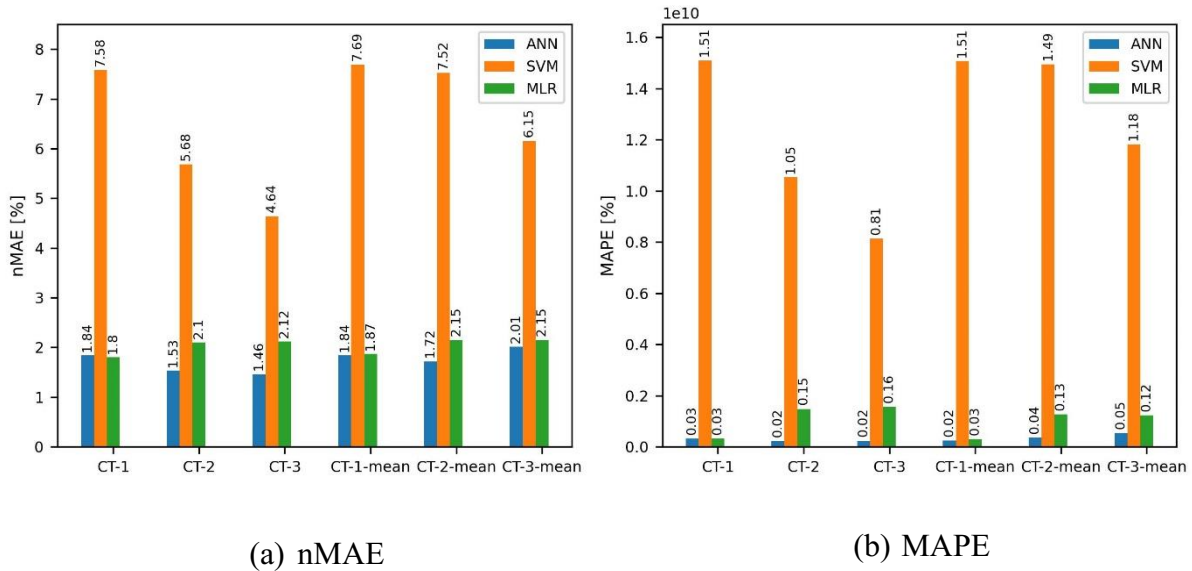


Figure 4-9. PV prediction metrics nMAE and MAPE

Figure 4-10 shows a comparison for the regional wind power estimation. All three prediction approaches perform very well. Especially, nMAE of SVM with CT-1 and CT-2 input are smaller than ANN model. ANN model with CT-3 input has the best prediction result. From the perspective of MAPE, it is obvious that ANN models have the best performance in wind power prediction.

Different machine learning approaches including SVM and Gauss Process Regression (GPR) are adopted for PV panel power estimation in [68]. SVM based on polynomial function with degrees from 1 (linear) to 3 (cubic) exhibit the nMAE result with the range of 10~15 %. GPR provides a higher performance with nMAE of 5~6 %. Different wind turbine power curves are fitted in [79]. They fitted the physical model based on wind speed. Due to the local on-site wind speed measurement for one single wind turbine, a lower normalized root mean square error (nRMSE) is obtained within about 3 %. Nevertheless, they have the limitation on the type of wind turbine.

For comparison, MLR works well in predicting PV power with one category of solar irradiance data due to its high linear relationship. SVM seems not suitable for the PV power prediction but it works well in the field of wind power prediction. In contrast, with raw measured weather

data or reduced mean weather data as input features, the ANN models provide a good prediction performance for PV and wind power.

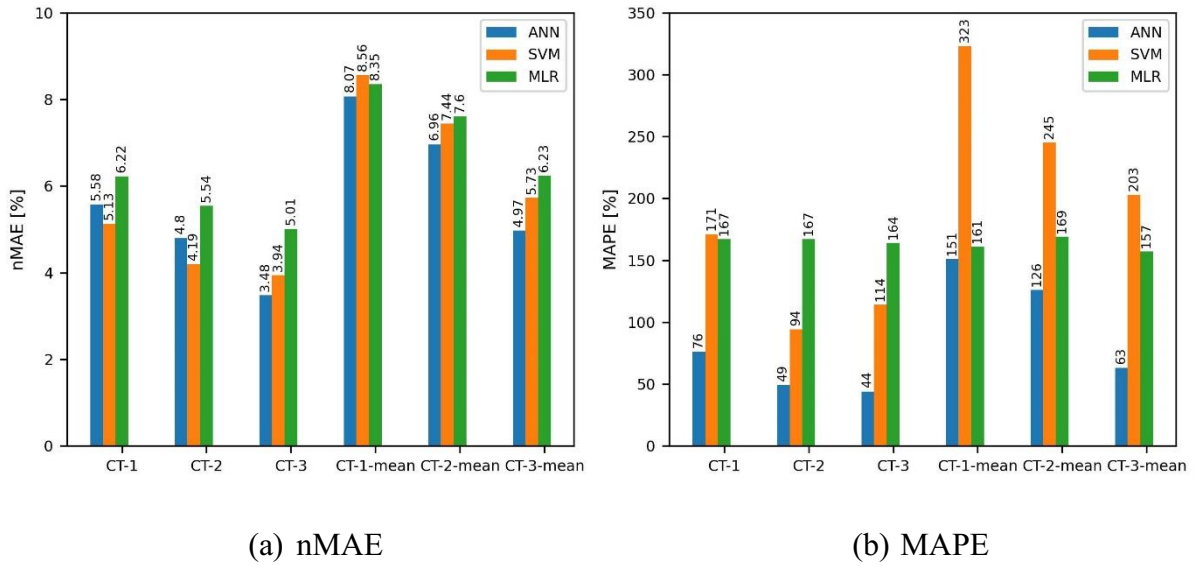


Figure 4-10. Wind prediction metrics nMAE and MAPE

4.2 Simulation Based on Aggregated Model

4.2.1 Experimental Setup and Data Pre-processing

The evaluation of regional PV and wind power generation is based on the regional aggregated characteristic curves. During the profile-based power conversion process, wind speed is used as input data. Moreover, regional PV power can be estimated based on solar irradiation and air temperature. The methodologies and processes of aggregated characteristic curves are already described in the Section 3.2 and 3.3. In order to reduce the individual measurement error, the regional average values of measurements from the different weather stations are used. The clustered mean weather data are calculated based on the nearest five measurement stations shown in the Figure 4-11. The weather data can be obtained from DWD database from the Section 2.3.3.

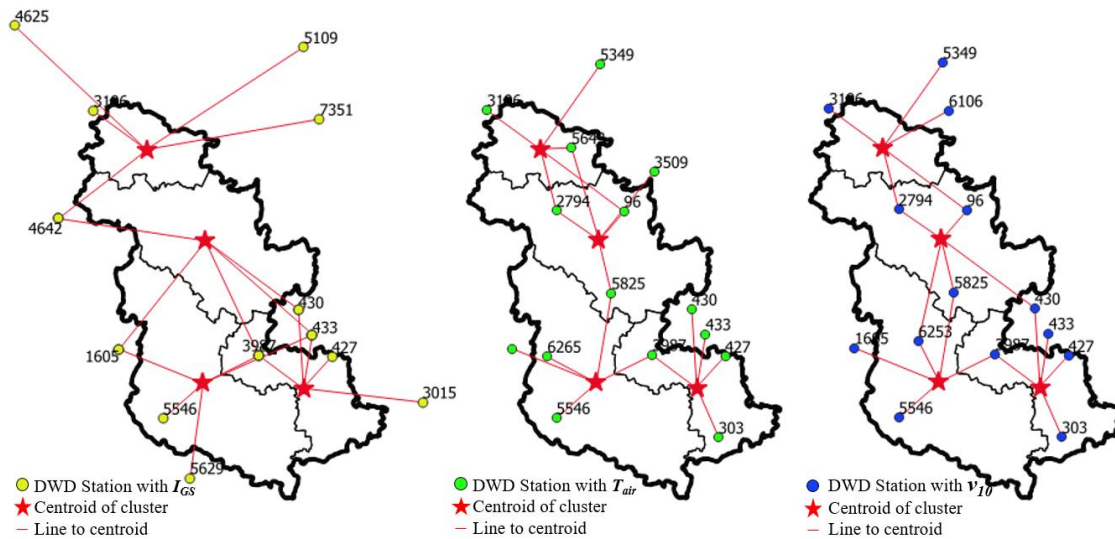


Figure 4-11: Mean weather data for each cluster

Setup for wind model

According to the installation date, the regional about 1067 WTs are further divided into 3 groups. There are about 115 new installed WTs between year 2015 and 2020. Compared with the installed rated power of WTs before 2005, the mean new installed capacity of a single WT has been almost tripled. The mean hub height is among 75 m until 2005. The new wind turbines installed after 2015 have a hub height of approx. 136 m. By the end of year 2019, the entire region has around 1721 MW installed wind capacity.

Figure 4-13 and Figure 4-14 illustrate the clustered installed wind capacity and mean hub height for different age groups respectively. By the end of 2019, cluster 1 has an installed wind capacity of 367 MW. Both cluster 2 and 3 have about 629 MW installed capacity, whereas only 95.2 MW installed WTs are connected to the cluster 4. All four clustered areas have an average hub height of above 130m for the G3 age group.

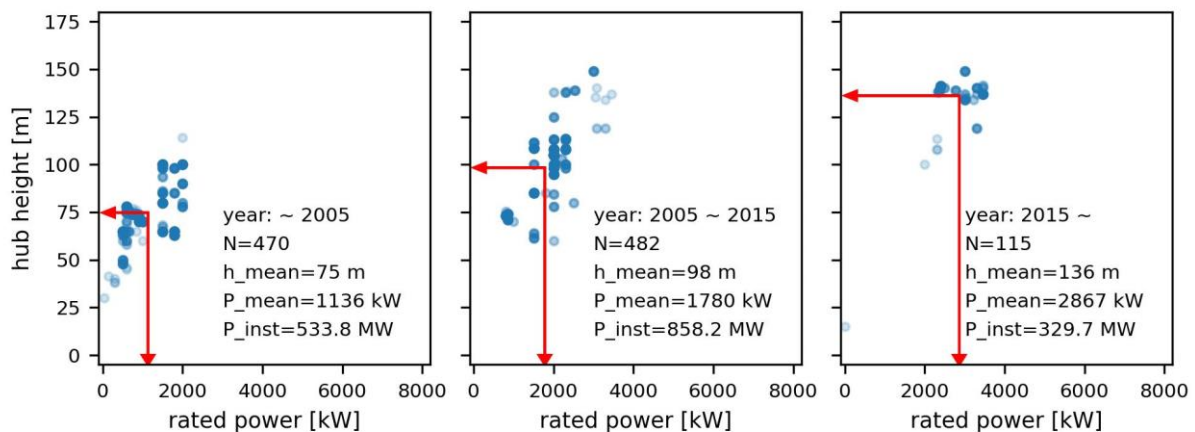


Figure 4-12: The mean installed capacity and hub height for the region

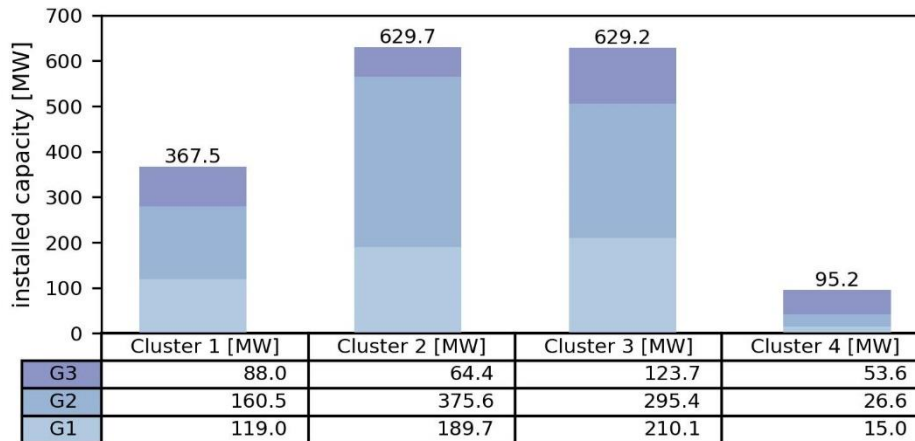


Figure 4-13: Clustered installed wind capacity for different age groups (status end year 2019)



Figure 4-14: Clustered mean hub height for different age groups (status end year 2019)

Setup for PV model

To simplify the calculation of the physical model and also because of no complete information about the installation type of each PV system, in this work the average temperature coefficient α is proposed as $0.43 \text{ }^\circ\text{C}^{-1}$ as treated in [65]. In addition, the inverter efficiency indicates how much DC power is converted into AC power. The efficiency of inverter is assumed approx. with 95 %. The parameter γ can be estimated by the mounting type of PV systems. Small PV power generation units connected to the low voltage grid system, such as households with rooftop PV systems, are supposed to use the parameter γ of $0.056 \text{ }^\circ\text{C}\cdot\text{m}^2/\text{W}$. In addition, large-scale PV systems connected to the medium- to high-voltage grids are mostly installed in open space. For this reason, a typical γ value of $0.02 \text{ }^\circ\text{C}\cdot\text{m}^2/\text{W}$ for ground-mounted PV system is used for large-scale PV systems.

The regional clustered installation capacity of PV systems is listed in the Figure 4-15. The installed capacity of PV connected to the medium- to high-voltage grids is significantly higher than that connected to the low-voltage grids. The installed PV systems in the DSO grid amount to approx. 946 MW for the region with four clusters for the end of 2019.

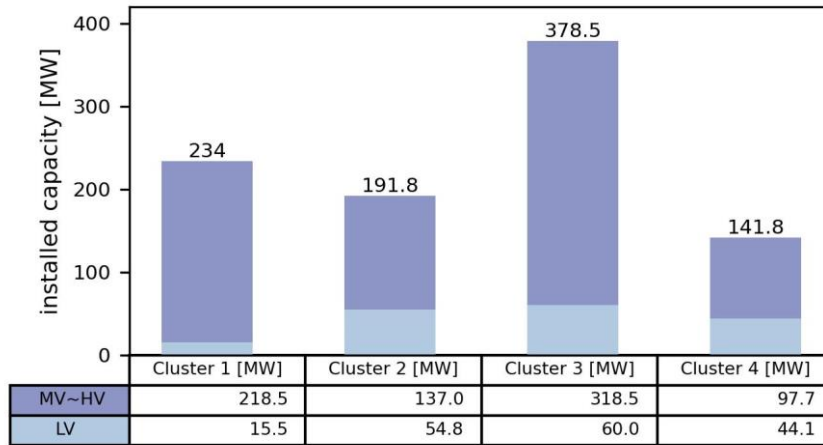


Figure 4-15: Clustered installed PV capacity for different voltage levels (status end year 2019)

4.2.2 Regional Wind Power Estimation

The clustered wind power is calculated from the classification due to the converting procedure in the Figure 3-3 in Section 3.2.2. Figure 4-16 illustrates the modelled wind power results for the regional four clusters. The modelled wind power can follow the changes and reach the maximum wind power for some periods. Although the installed wind power capacity of the two clusters 2 and 3 is almost the same, their power characteristics and weather conditions are different, so that the power generation curves vary. In addition, since cluster 4 has the highest mean hub height, it reaches its rated power fast and maintains it for a longer period. However, in terms of the regional aggregated and real measured wind power, there is a big difference in peak power. As can be seen in the Figure 4-17, the real average relative peak power with reference to the installed capacity differs monthly. The preliminary modelled wind power results can reach its rated power almost every month. The real mean wind peak power can only range between approx. 55 % in summer and 80 % in winter.

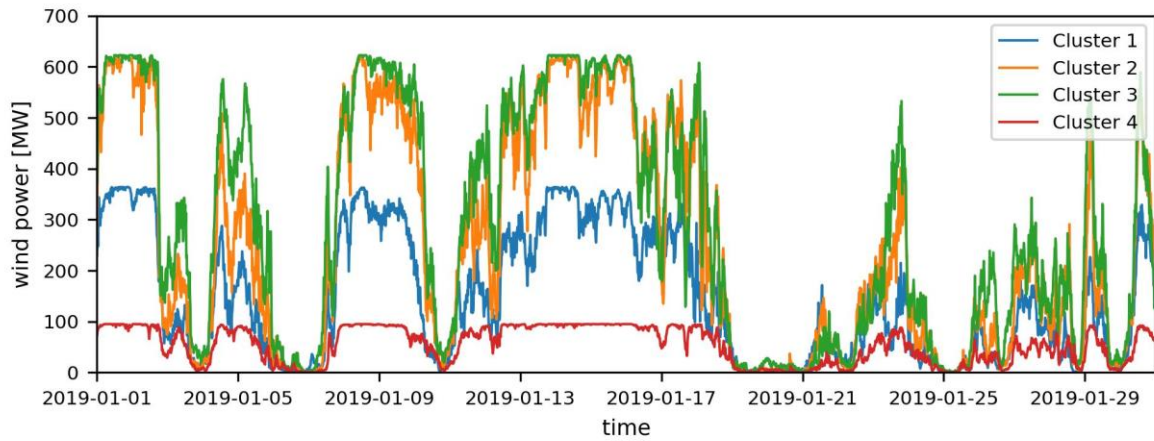


Figure 4-16: Modelled clustered wind power generation

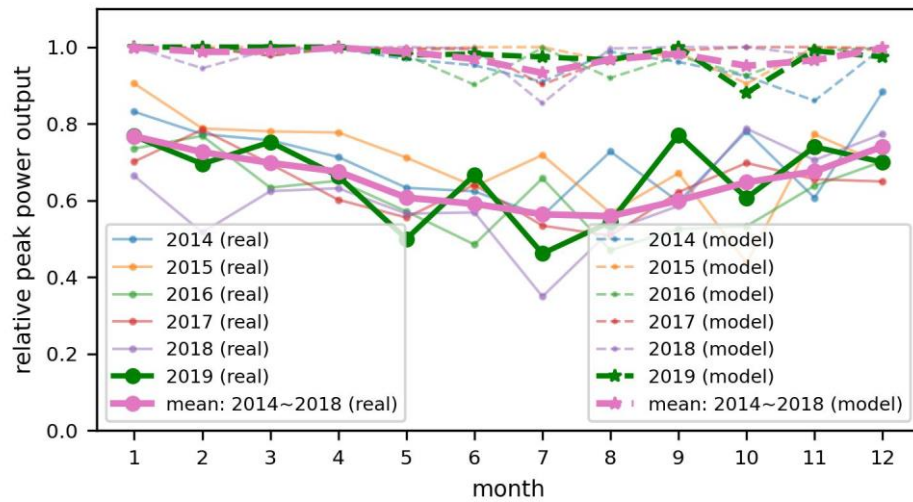


Figure 4-17: Monthly relative wind peak power output from historical measurements and model calculations

Such large power difference is because the preliminary aggregated characteristic model does not take into account the actual operating situation. The correction factor may include the following impact aspects. It is obvious that, some wind power generation systems are not available due to the planned and unplanned maintenance and repair work. In reality, it is difficult for a wind turbine to be operated along its characteristic curve due to the difference in the intensity and direction of the local wind in terms of the day and season. The direction of the wind turbine rotor should be adjusted accordingly, which results in electrical self-consumption and a delay in power generation. Furthermore, the start-up process in a small wind speed range and the influence of turbulences also brings great losses to the power generation process. In the case of the grid connection, WT power plants are not directly installed near the connection point of the electricity grid. The generated power is transmitted through cables to the connection point

and transformed to the corresponding voltage level. Generally, there are inevitable power losses during the transfer and transformation.

In addition to the mentioned technical factors, there are some requirements for switch-off or curtailment at certain time of day and season. Due to congestion management and limited grid capacity, the current grid cannot absorb that much electrical energy. In this case, the electricity generation cannot be fed into the grid system. Therefore, the new regional wind power can be expressed by adding a correction factor as a post-processing and it is described as:

$$P_{Wind,model}^* = P_{Wind,model} * f_c \quad (4-3)$$

where f_c is the dynamic correction index included all influencing factors, $P_{Wind,model}^*$ is the new corrected modelled wind power.

The dynamic correction factor is monthly updated based on the historical data investigation. The correction factor is calculated based on the multiple relationship between the relative mean measured peak power from last years and the corresponding mean modelled peak power. The resulted new regional aggregated wind power is plotted in the Figure 4-18. The preliminary calculated wind power (orange dash line) has higher peak power output than the real measurement (red line). By studying the historical data (i.e. last five years), a correction factor of 76.9 % for January is applied to reduce the modelled result. The correction factors are statistically examined using the historical mean values from 2014 to 2018 in the Figure 4-17. The new corrected power curve (blue dash line) demonstrates that the new aggregated model with correction factor can predict the regional real wind power well and accurately. A significant improvement is achieved for the aggregated model with the correction factor.

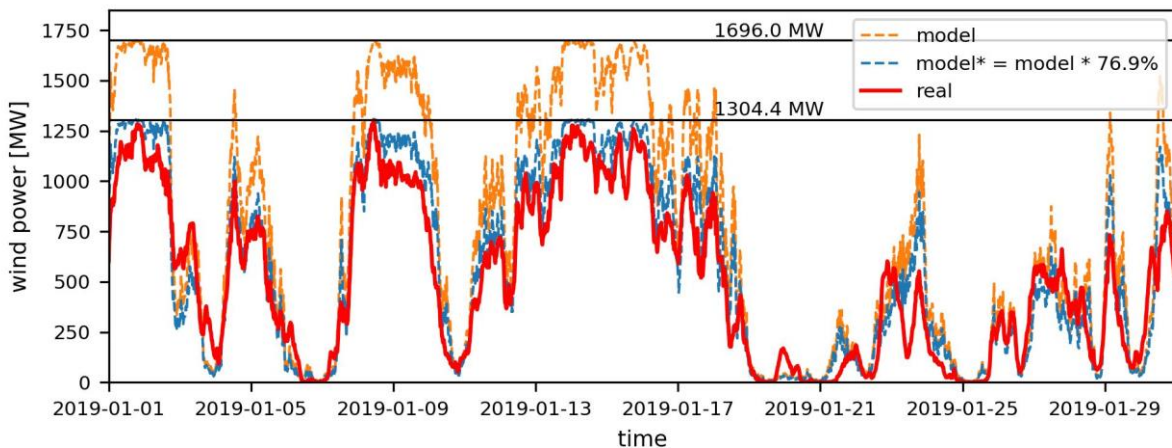


Figure 4-18: Regional aggregated wind power generation in January 2019

After applying the historical dynamic correction factor f_c on the modelled regional wind power, the monthly relative peak power and annual duration curve are shown in the Figure 4-19. The aggregated model has approx. 1639 full load hours (FLH) with a deviation of 155 hours to the real measurement. The following metrics of the proposed model, mean absolute error (MAE), root mean squared error (RMSE), normalized mean absolute error (nMAE) and normalized root mean squared error (nRMSE) with reference to the rated power are listed in the following Table 4-8. The metric values are calculated based on year 2019. It is noticeable that nRMSE maintains between 8 % and 16 % through the year. The annual prediction error of nMAE and nRMSE can reach around 8.4 % and 10.9 % respectively.

Table 4-8: Monthly and annual average metrics of the regional wind power estimation for year 2019

Month	MAE [MW]	nMAE [%]	RMSE [MW]	nRMSE [%]
1	110.8	6.52	144.5	8.50
2	156.4	9.20	207.9	12.23
3	132.7	7.81	172.9	10.18
4	209,8	12.34	259.4	15.27
5	157,0	9.24	191.3	11.25
6	139,7	8.19	185.1	10.85
7	123,0	7.21	157.8	9.25
8	105,9	6.20	150.6	8.81
9	139,2	8.14	178.8	10.46
10	159,4	9.26	203.1	11.80
11	130,3	7.57	180.3	10.47
12	152,4	8.85	196.5	11.41
annual average	143.1	8.40	185.7	10.90

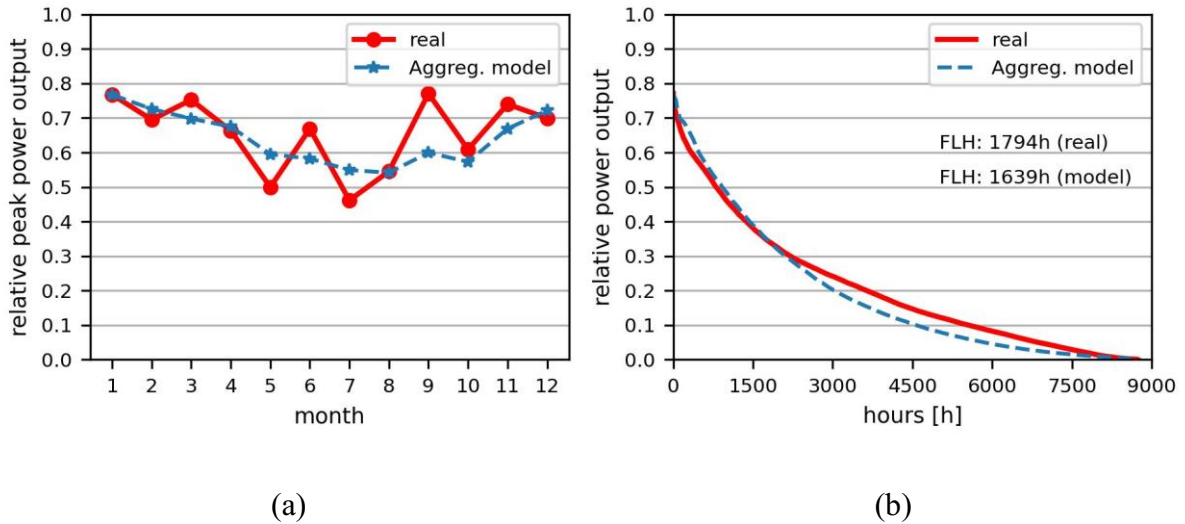
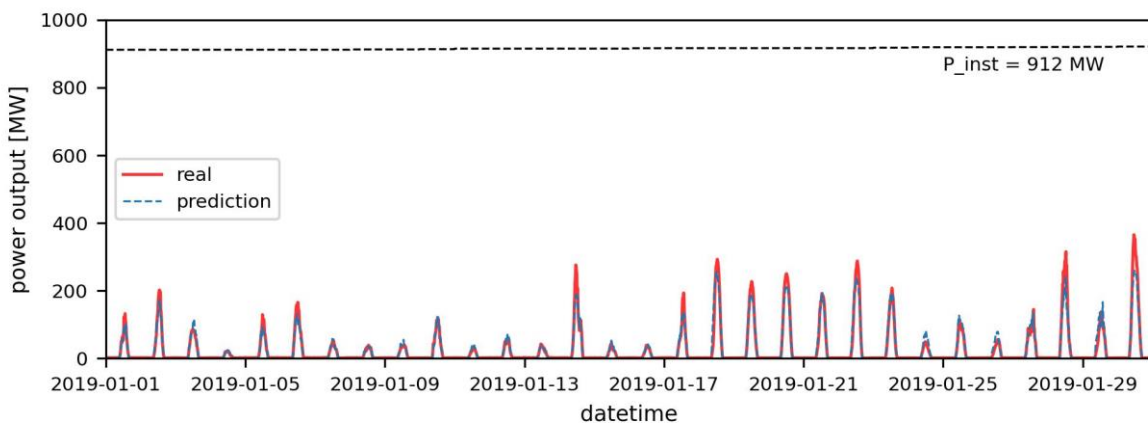


Figure 4-19: Monthly relative wind peak power output with correction factor (a) and the annual duration curve of regional wind power (b)

4.2.3 Regional PV Power Estimation

The aggregation of the regional cluster PV power output is determined by the procedure of regional PV power generation calculation in the Figure 3-6 in Section 3.3.2. The measured solar irradiation and ground air temperature are used as input parameters for the calculation. As can be seen from the Figure 4-20, the result of the PV power output prediction in January and June matches with real power measurement from DSO very well. It can provide a good prediction and track changes in regional PV power generation across different weather conditions and seasons.



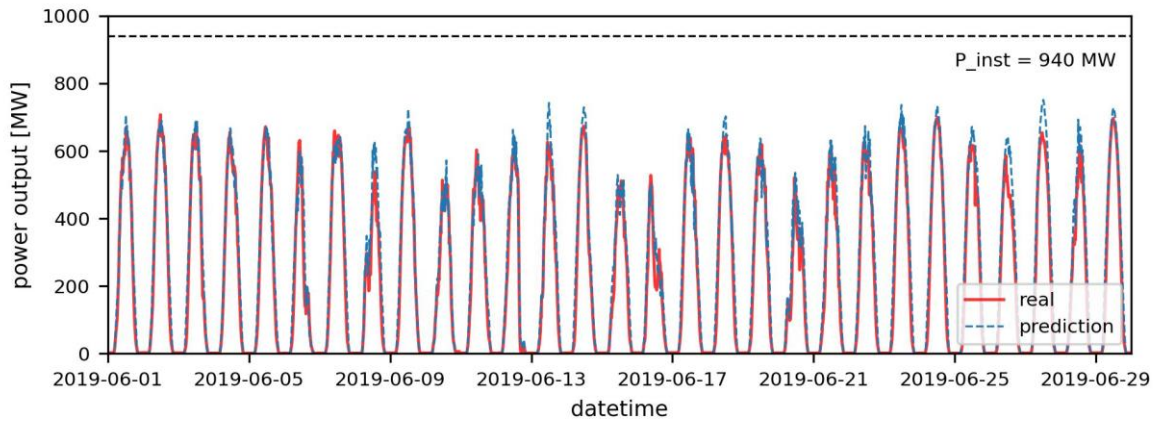


Figure 4-20: Regional aggregated PV power generation in January and June 2019

When focusing on the maximum power output, the relative peak power is formed in relation to the corresponding installed capacity. The monthly relative PV peak power output is shown in the Figure 4-21. It is noticeable that the mean value of the peak power from the year 2014 to 2018 can reach about 75 % of the rated power. In comparison with the historical real measurements, model calculations display a slightly higher peak power output, approx. 80 % of the rated power. The reasons of the peak power difference can be the feed-in power limitation for small-distributed power generators according to EEG-2017. In the case of large-scale PV systems, regulation is done by remote control by the local grid operator. Power grid systems have to absorb a large proportion of PV power on sunny days. The purpose of the feed-in power limitation is to stabilize the grid system and to ensure the electricity grid not to be overloaded.

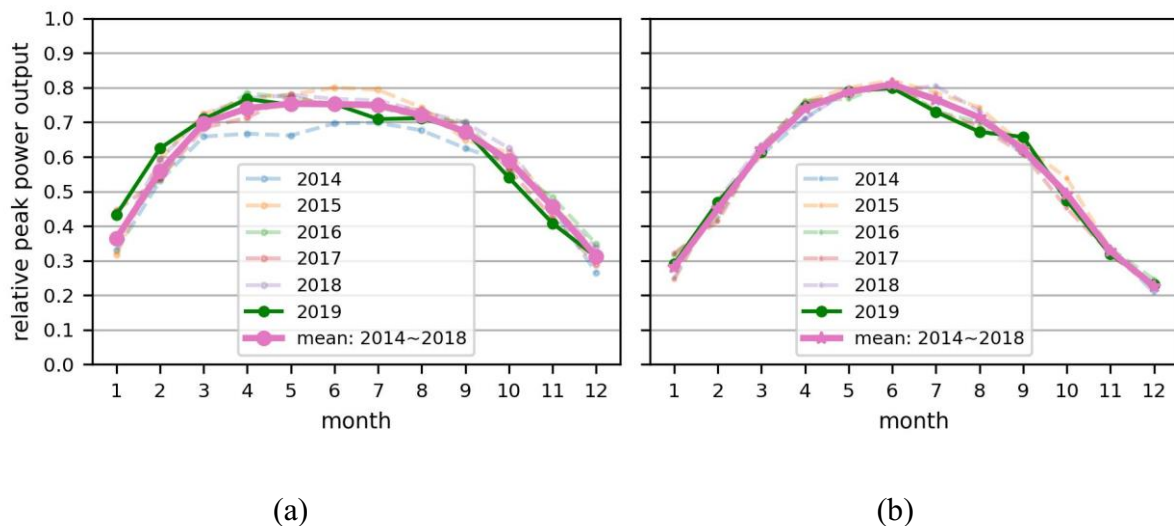


Figure 4-21: Monthly relative PV peak power output from historical measurements (a) and model calculations (b)

The power limitation and regulation are very difficult to simulate. Similar to the post-processing of wind power characteristic calculation model, historical measurements are investigated to mitigate the calculation errors from power limitation, regulation and other various reasons, such as power transport losses. As can be seen from the Figure 4-21, the mean value from the last five years (2014~2018) gives a good imitation for the monthly relative peak power of the year 2019. Therefore, a post-processing with historical power measurements can be implemented. To compare with the mean values from the historical real measurement and model calculation for the last years, the correction factors will be generated.

After adjusting the peak power, Figure 4-22 shows the results of the monthly peak power and annual duration curve of the real measurement and aggregation model. The aggregated model reproduces not only the peak power but also the annual duration curve in a very good range. The full load hour of regional PV systems is about 1042 hours for the model simulation, whose difference with the actual measurement is less than 25 hours.

Various metrics of the time series prediction are presented in the Table 4-9. The result shows that summer has higher PV power output with slightly larger prediction error in contrast to winter. The annual average MAE is 16.63 MW, which corresponds to 1.77 % of the regional installed capacity. Meanwhile, the annual RMSE amounts to about 32 MW.

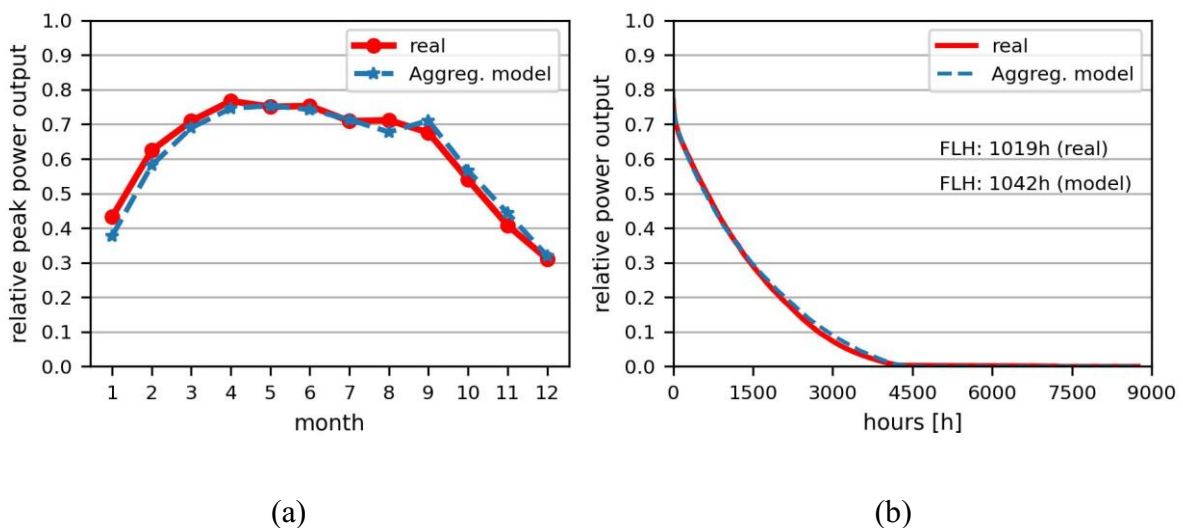


Figure 4-22: Monthly relative PV peak power output with correction (a) and its annual duration curve of regional PV power (b)

In reality, each solar module type has its own characteristic curve and parameter. Due to lack of the individual parameters for each RES system, the aggregated model applied typical average values to represent them, which brings errors for the calculation result. Furthermore, a slightly larger prediction error occurs due to the spatial compensation effects, which can be attributed

to different local weather conditions. Other influencing factors to be considered in the simulation based on PV characteristics are the orientation and inclination of the PV panels. The solar panels with the orientation to the southeast will start to produce power earlier than the panels with the orientation to the west. The tilt angle of ground-mounted solar panels can be adjusted to achieve maximum possible solar energy, while the roof-mounted is limited by the roof angle. The real-time solar irradiation on the inclined PV panel surface varies by the radiation incidence angle. Another important parameter is wind speed. It affects ventilation and cooling effects for the PV modules. This work does not consider these factors for each PV system in detail, but adopts a post-processing approach with correction factor based on the historical measurement to mitigate these errors.

Table 4-9: Monthly and annual average metrics of the regional PV power estimation for year 2019

Month	MAE [MW]	nMAE [%]	RMSE [MW]	nRMSE [%]
1	6.28	0.69	14.98	1.63
2	18.19	1.97	40.75	4.40
3	11.80	1.27	24.98	2.69
4	19.48	2.09	31.61	3.38
5	23.61	2.52	37.85	4.04
6	30.40	3.23	45.23	4.81
7	27.75	2.95	43.92	4.67
8	20.80	2.21	35.13	3.73
9	16.01	1.70	30.78	3.26
10	13.93	1.47	28.74	3.04
11	6.67	0.71	15.01	1.59
12	4.91	0.52	13.62	1.44
annual average	16.63	1.77	31.98	3.41

4.3 Simulation based on Deep Learning Method

In this work, deep learning models with LSTM, stacked LSTM (SLSTM) and CNN-LSTM are applied to extract the relationship between regional measured wind speeds from different

measurement stations and regional aggregated wind power. Moreover, different time steps of historical measurements are investigated to estimate the real-time regional aggregated wind power.

4.3.1 Experimental Setup and Data Pre-processing

A step of input features for deep learning model with a matrix shape of [time_steps, n_features] is set as input shape. It should be noted that the parameter “time_steps” actually corresponds to the size of look back window. In this work, four different time steps are defined to test different look back window size of historical data for the current wind power estimation. In particular, the values of look back window size $s = 12, 24, 48, 96$ correspond to the last 3, 6, 12, 24 hours respectively due to 15-min time resolution.

In addition, another four parameters are added to store the daily and seasonal features, namely $\left[\cos\left(2\pi\frac{t}{TD}\right), \sin\left(2\pi\frac{t}{TD}\right) \right]$ and $\left[\cos\left(2\pi\frac{t}{TA}\right), \sin\left(2\pi\frac{t}{TA}\right) \right]$. Since the temporal resolution is 15-min, it contains 96 time series data per day and 35040 time series data in one year. Therefore, the periods of TD and TA are 96 and 35040 respectively. Multivariable of wind speeds from different measurement stations (14 sites) and four daily and seasonally features form totally 18 input features. The parameter setups for the deep learning models itself are listed in the Table 4-10. The training epochs are set as 10 and an EarlyStopping function is used to avoid over-fitting.

Table 4-10: Parameters for neural networks

Deep Learning Models	Parameters setup
LSTM	LSTM(units = 10) → Dense(units = 1) activation = 'relu'; optimizer = 'Adam'
SLSTM	LSTM(units = 10) → LSTM(units = 10) → LSTM(units = 10) → dense(units = 1) activation = 'relu'; optimizer = 'Adam'
CNN-LSTM	Conv1D(filters = 10, kernel_size = 3) → Cov1D(filters = 10, kernel_size = 3) → MaxPooling1D(pool_size = 2) → Flatten() → LSTM(units = 10) → LSTM(units = 10) → Dense(units = 1) activation = 'relu'; optimizer = 'Adam'

4.3.2 Comparison of Different Models

After pre-processing of the measured data with different look back windows, the data are randomly split into training and validation datasets based on the data of the year 2018. After the deep learning model is trained, the year 2019 data is used as test dataset for the regional power prediction. Figure 4-23 shows the predicted regional wind power based on different deep learning models with look back window of $s = 12$. All the three models can follow the regional wind power based on the regional distributed wind speed measurements.

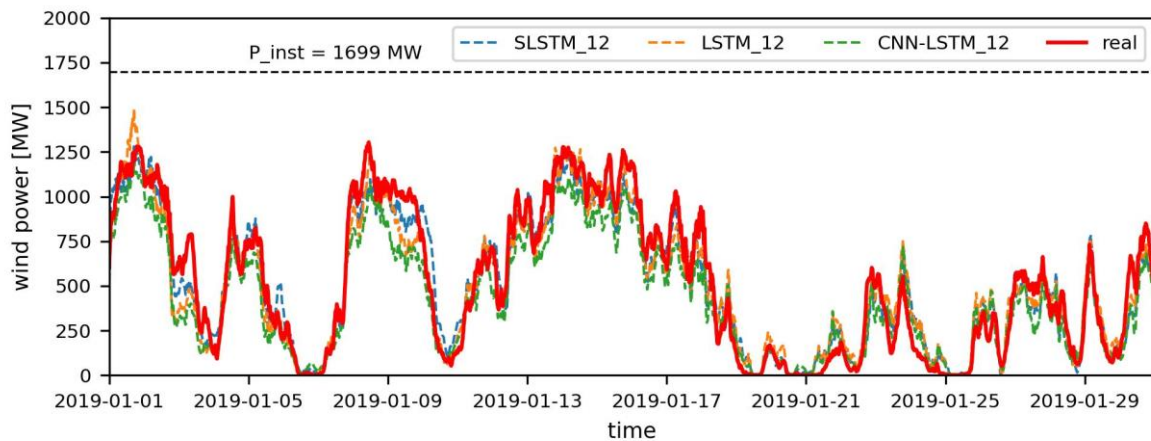


Figure 4-23: Wind power estimation based on different neural network models

In addition, Figure 4-24 shows the monthly comparisons from all tested models with different look back windows. It can be seen that neural network models have a better prediction performance than the aggregated model. Meanwhile, less prediction error can be obtained in summer than in winter. SLSTM can reach about 3 % nMAE from June to August and about 6 % in winter. Figure 4-25 presents as well the annual evaluation of nRMSE. Stacked LSTM (SLSTM) performs very well and stable with different look back window sizes. LSTM has also a better predictive computing ability with last 3 hours data ($s = 12$) or 24 hours data ($s = 96$). Relatively, CNN-LSTM model cannot be fitted as well as LSTM and SLSTM model. It may need more epochs for training process because of more weighting parameters and convolutional computation.

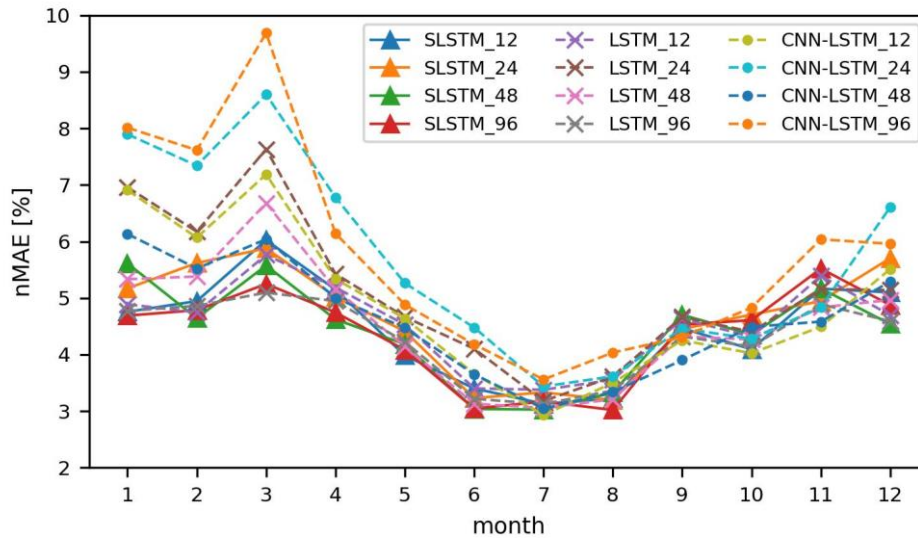


Figure 4-24: Monthly nMAE of deep learning models

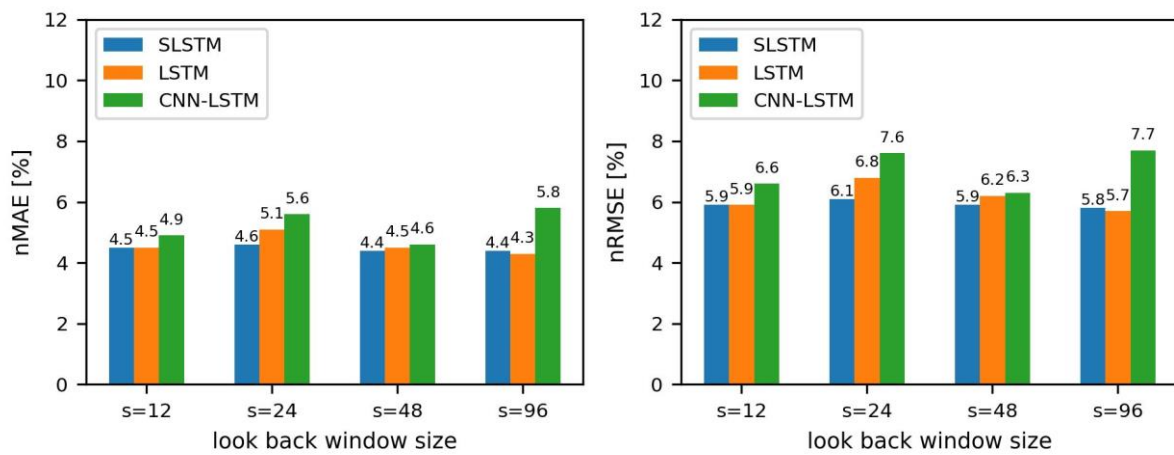


Figure 4-25: Annual nMAE and nRMSE of deep learning models

Furthermore, the annual duration curve for the modelled regional wind power are investigated and compared in the Figure 4-26. The annual duration curve of the estimated power from SLSTM model is very close to the actual duration curve. Concerning the full load hours (FLH), the real measured wind power generation in the investigated region has 1794 FLH for the year 2019. The peak power reaches about 80 % of the rated power. Additionally, Figure 4-26 shows the FLH difference between real measurement and the prediction results based on AI models with different look back window sizes. LSTM and SLSTM perform better in predicting the FLH than CNN-LSTM and the aggregated model. SLSTM with last 6 hours measure data ($s = 24$) has the best prediction effect with only 4 hours difference in FLH.

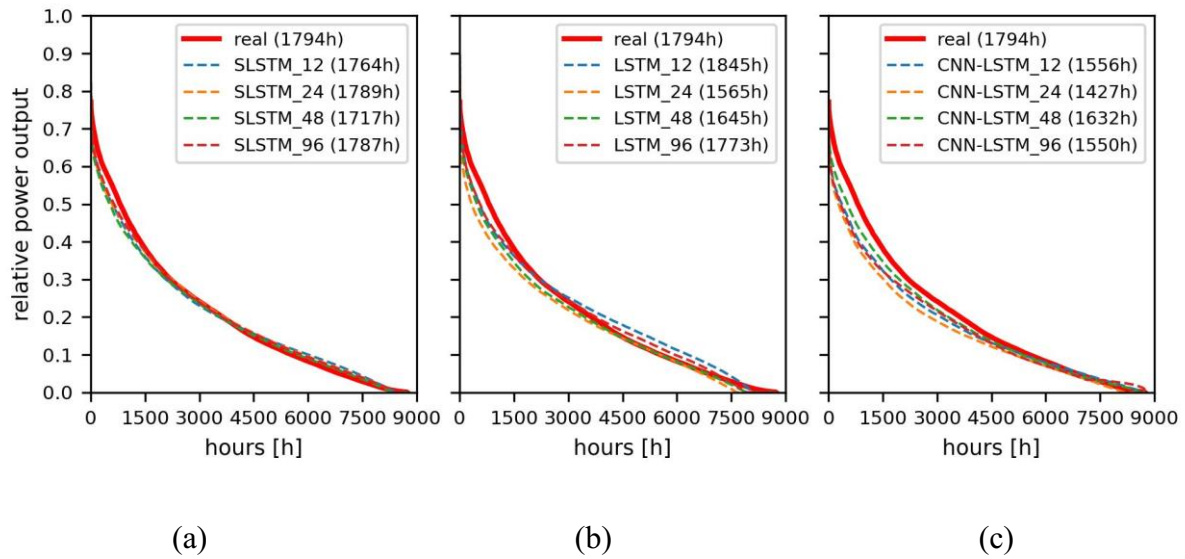


Figure 4-26: Annual duration curve of wind power estimation from deep learning models

4.4 Concluding Remarks

This chapter gives a view on the regional power estimation of large-scale PV and wind power systems based on real measurements. Different physical and data-driven AI methods are applied to the regional RES power estimation. This work allows for both the raw measured weather data (CT-1, CT-2, CT-3) and reduced mean weather data (CT-1-mean, CT-2-mean, CT-3-mean) as model input features for the regional PV and wind power estimation. A simple ANN architecture is adopted by using different input CT features. The ANN model can extract information from the related weather parameters. The results have shown that with increasing categories of input features, ANN works better than only one high-related category weather data. The extended weather data are not treated as redundant information for the regional power prediction in this work. The annual nMAE can reach about 1.5 % and 3.5 % for PV and wind power prediction with CT-3 as input features respectively, while with CT-1 input, the annual nMAE is about 1.8 % and 5.6 % for PV and wind power respectively. Especially, the sensitivity of monthly prediction error of nMAE shows that ANN model provides a stable prediction accuracy for wind power throughout the year for different seasons, and higher accuracy can be obtained for PV power estimation in winter due to lower power output.

In addition, the regional aggregated models based on the PV and wind power characteristics are proposed. The different aggregated characteristic curves represent synthetic curves of the local WTs installed from different periods and PV systems connected to different voltage levels. Theoretically, the investigated power characteristic curves can imitate the regional PV and wind

power output. Nevertheless, in reality, not all RES systems are in operation and cannot be always operated ideally under optimal weather conditions. Therefore, through the statistical analysis of the historical regional PV and wind power infeed, a dynamic correction factor can be derived. Under the assumption with the correction factor, better prediction results can be achieved. By using real measured data, the regional power prediction performance is investigated. The PV power prediction errors of nMAE and nRMSE can reach 1.77 % and 3.41 % respectively. The annual prediction errors nMAE and nRMSE for wind power maintain around 8.4 % and 10.9 % respectively.

Furthermore, different neural networks with deep learning models are applied in the wind power estimation. The regional measured wind speeds are treated as input features for the deep learning models. Essentially, these proposed neural computations achieve a good performance for the regional wind power estimation. Stacked LSTM model performs better than the other two models of LSTM and CNN-LSTM in time series prediction of wind power. It is noticeable that the annual prediction error of nRMSE decreases from 10.9 % for aggregated model to around 6 % for deep learning method. Particularly, the monthly nMAE demonstrates that, the neural networks achieve nMAE of about 3 % in summer and about 6 % in winter. The annual duration curve of wind power from the real measurement and the deep learning model of SLSTM match each other very well.

5 Implementation of Cluster Model in Virtual Regional Energy Storage System

In this chapter, two charging/discharging strategies for the grid-scale energy storage systems (ESS) are proposed to determine when and how much power to charge/discharge the ESS. In order to realize these strategies, a fuzzy logic control system is applied. The aim of the proposed strategies is to reduce peak power generation, consumption and grid fluctuations. In particular, the ratio between energy capacity and rated power of ESS is analysed. The performance of the proposed strategies is evaluated based on two aspects, the normalized power of ESS itself and its impact on the power grid. Simulation studies of rule-based control systems with different energy-to-power ratios are conducted.

The roadmap of this chapter is structured as follows. Section 5.1 describes the related work to the energy storage system. Section 5.2 describes the methodology for the charging/discharging strategies for the regional ESS. Section 5.3 focuses on the model development. To realize the strategies fuzzy control system is introduced and adapted to this unique purpose. The simulation procedures and results are presented in section 5.4. Section 5.5 gives the concluding remarks.

5.1 Overview of Regional Energy Storage System

In recent years, renewable power generation and energy storage systems (ESS) are widely researched. Wind-PV-Storage projects are established and evaluated to achieve smooth power output and can be operated in island- or grid-connected mode [82,83]. In [37] a grid-connected large-scale ESS is established. It is proven that multi-agent based energy-coordination control strategy increases the efficiency and robustness of global energy system. Fuzzy control method is used in the battery agent for the decision-making. In addition, a fuzzy inference system is proposed to control the operation of ESS in microgrids [84]. The effective control strategy results in a better power quality and life cycle of ESS. Fuzzy control technique is also implied to satisfy the economic goals and to reduce the imbalance between cost and benefit in a hybrid grid system [85]. Significant attention has been paid to energy management for a storage system. Data-based Energy Management System (EMS) has been researched to reach the optimum decision making [86]. Using real-time data for performing energy management optimisation of the EMS performance is achieved [87]. The analysis from a different perspective, such as capacity of RES, capacity of ESS and demand response should be pursued in the future. In terms of intelligent energy management, the integration of power generation, load and storage assets into the power grid entity is presented in [88]. The power management and control

architectures are defined to achieve the cooperative behaviour amongst the grid assets. Associated storage capacity can be determined according to the variation of power generation and load demands [89].

Based on the fluctuating electricity generation and load, the resulted power flows are converted into a synthetic power flow with generation and consumption. In addition, a model of electricity storage in a cluster is developed. The proposed grid-connected, large-scale energy storage system is investigated to balance local power generation and load. Furthermore, two charging/discharging strategies under observation of power imbalance between generation and consumption, and the current state-of-charge (SoC) are considered to set the operation process of ESS. Using real system measured data the algorithms are simulated and evaluated. Energy-to-power ratios instead of capacities are discussed in the further simulation procedure. To verify the efficiency of the control systems under the two charging/discharging strategies, the normalized power output and the new generated power saldo with ESS will be then evaluated.

5.2 Methodology for Storage System

The electricity generation and loads are connected to different voltage levels. The regional grid network cluster related to a distribution system ranging from low-voltage to high-voltage levels is proposed and created in Chapter 2. Within the network cluster, there are various local generations and loads, and the power difference between the electricity generation and consumption is defined as cluster power saldo from the Equation (2-2) in the Section 2.2, which describes the needed power from the external grid.

Electrical energy storage system is defined and used in this work to modify the influence of regional renewable power generation. It can reduce the fluctuation of the generation and load, and the transmission congestion can be reduced through the investigated ESS as well [90]. The high power imbalance between generation and load requires ESS to ensure the optimal operation of the power grid. It is necessary to reduce the peak of the power saldo. With the integration of ESS, the new generated power saldo can be formed in the Equation (5-1). In addition, Figure 5-1 illustrates a simplified power supply with ESS. The virtual energy storage system can store the regional surplus energy when the supply is greater than the demand. The ESS provides the energy to the local grid when the supply is less.

$$P_{Saldo_new}^{NC} = P_{Load}^{NC} - P_{Gen}^{NC} + P_{ESS}^{NC} \quad (5-1)$$

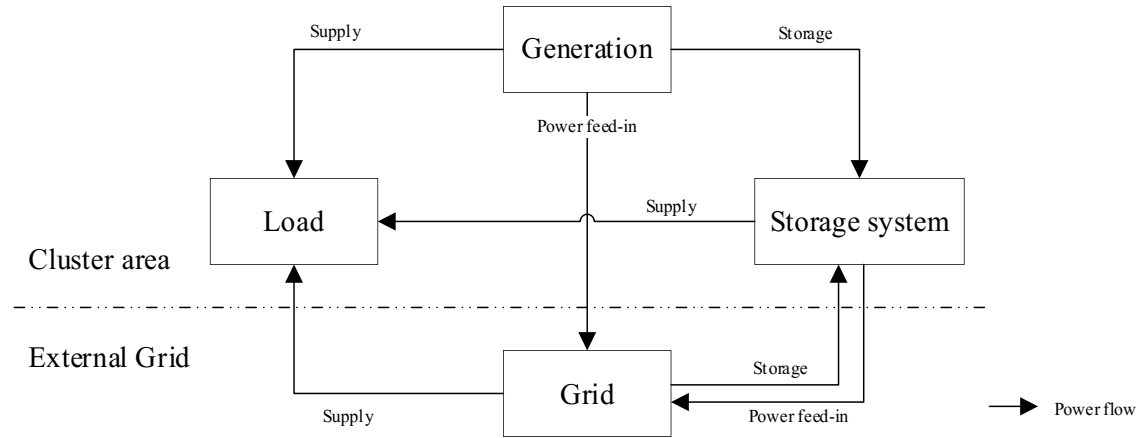


Figure 5-1: Energy storage system in power supply system

The virtual energy storage system is understood as a battery model with charging and discharging processes. Additionally, a storage system is characterized primarily by its storage capacity and its rated power. In this work, the energy-to-power ($e2p$) ratio is used and defined in the Equation (5-2). The rated energy capacity (E_{cap} in MWh) is divided by the rated power output (P_{rated} in MW), then the result forms the duration time ($e2p$ in h) for charging or discharging. The $e2p$ ratio describes how long the storage system can be operated at its rated power. State-of-Charge (SoC) represents the available capacity utilization of the ESS, which 0 % stands for empty and 100 % means full. After the charging/discharging process, the new SoC will be then formulated according to the Equation (5-3).

$$e2p = \frac{E_{cap}}{P_{rated}} \quad (5-2)$$

$$SoC(t + 1) = \frac{E(t + 1)}{E_{cap}} = \frac{E(t) + \Delta E(t)}{E_{cap}} \quad (5-3)$$

where $E(t)$ is the current capacity of the ESS and $\Delta E(t)$ represents the amount of charging/discharging energy. The charging/discharging strategy could be used to manage the operation mode of ESS based on the power saldo. Two charging/discharging strategies are illustrated in the Figure 5-2. Strategy 1 is used to discharge the ESS during the peak time of power saldo, and during the off-peak time the ESS is charged to reduce the peak power. When the power saldo is neutral, then no operation is undertaken in ESS, which means neither charging nor discharging happens. Strategy 2 considers the current situation of SoC . The storage system should be operated with higher power rate in extreme grid situations. From this

aspect, the energy storage system should be charged when the power saldo is low or negative with surplus energy, and discharged at a higher value of power saldo by lack of energy.

The operation power of ESS can be scaled in to the interval $[-1, +1]$ field with the implementation of normalized power of ESS $p_c(t)$, and it is formed by the Equation (5-4). The positive normalized power represents the charging process and the negative normalized power indicates the discharging process of the storage system. $P_{ess}(t)$ represents the operation power at timestamp t . The State of Charge (SoC) for the next time step is formulated in the Equation (5-5) with the implementation of $e2p$ ratio and normalized power $p_c(t)$. Δt donates the charging or discharging time.

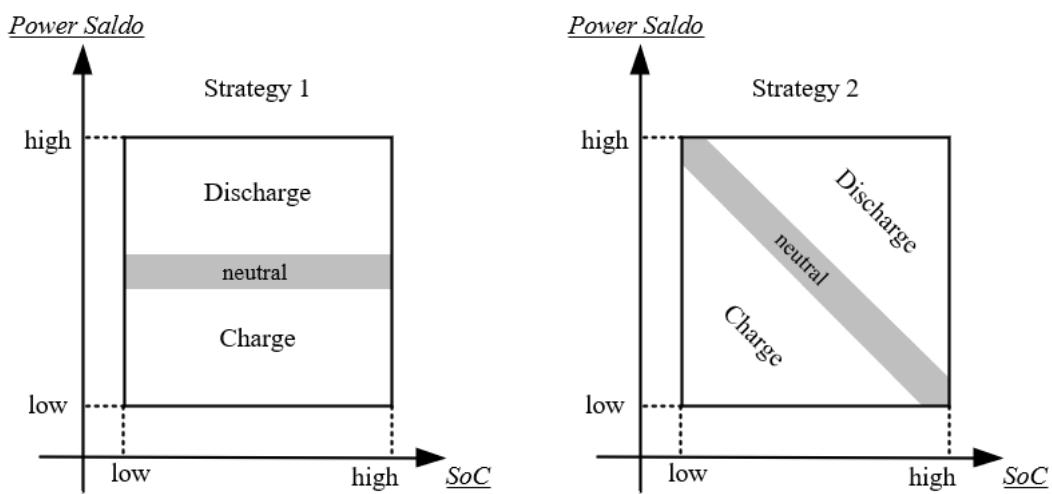


Figure 5-2: Charging/discharging strategies for a storage system

$$p_c(t) = \frac{P_{ess}(t)}{P_{rated}} \quad (5-4)$$

$$SoC(t+1) = \frac{E(t) + P_{ess}(t) * \Delta t}{E_{cap}} = SoC(t) + \frac{p_c(t) * \Delta t}{e2p} = f(SoC(t), p_c(t)) \quad (5-5)$$

5.3 Model Development and Experimental Setup

In order to realize the charging/discharging strategy for virtual storage system the fuzzy logic control is implemented. Fuzzy logic control can define the operation power rate quite well. Figure 5-3 shows the basic configuration of fuzzy inference system (FIS). Firstly, the input parameter will be fuzzified according to the fuzzy set. Secondly, IF-THEN rules are set and combined with the inference control system. Lastly in this work the Mamdani defuzzification

method is used to find the centroid of the membership distribution [91]. The output values are set as the normalized power of the storage system. The fuzzy system controller has two inputs and one output. The simulation of the proposed model architecture is realized in SciKit-Fuzzy [92], which is one fuzzy logic toolbox integrated in Python computing language.

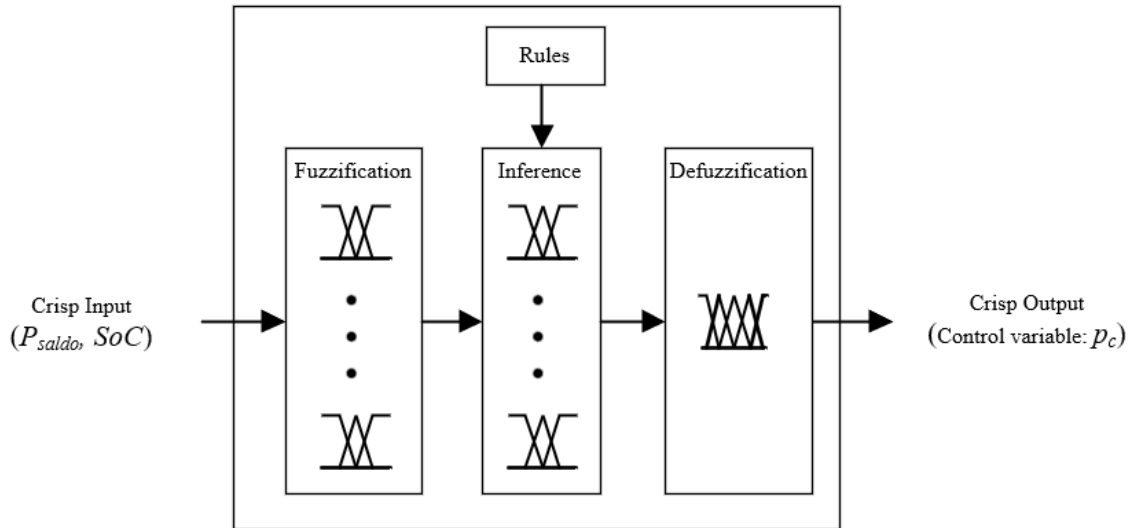


Figure 5-3: Basic configuration of a fuzzy system

Membership function

During the data fuzzification the membership functions are created. The membership function assigns the parameter into graded levels into fuzzy set. It determines the degree of the variable of each element. The parameter should be pre-processed before defining the membership function. In this work, the triangular functions are used to generate the input and output membership functions. With the aid of the min-max feature scaling, the input parameter is valued in the interval range [0, 1]. The scaling process is carried out for the power saldo. The transformation process for parameter is given by:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (5-6)$$

The membership function of input parameters and output rules function of normalized power are shown in the Figure 5-4. The centroid of the class area represents its typical value. The description of the fuzzy sets are explained in the Table 5-1 and Table 5-2, where the input parameters are classified in 5 groups, from very low to very high. The output power is graded into 5 levels with positive and negative characteristics. The negative power output means discharging and positive value means charging the storage system.

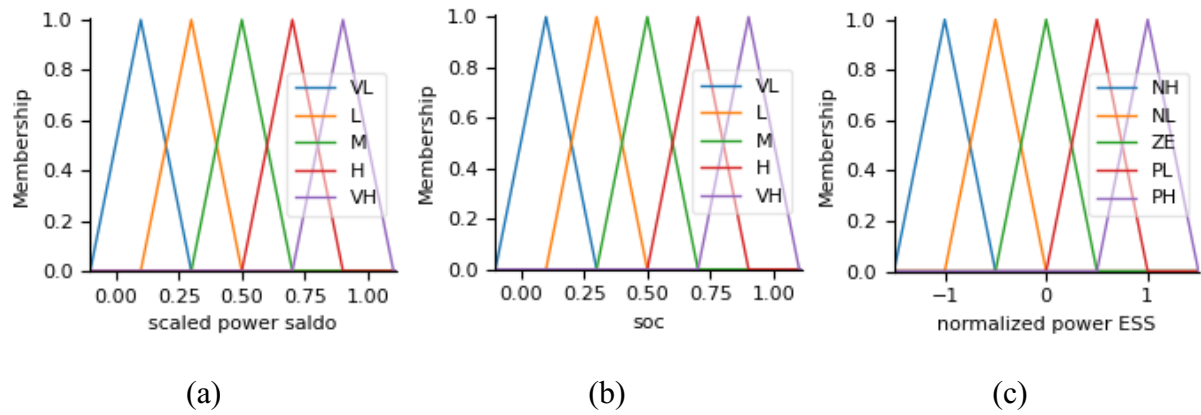


Figure 5-4 Input / Output membership functions

Table 5-1: Input fuzzy sets description

ID	Membership	Description	Scaled value
1	VL	very low	10 %
2	L	low	30 %
3	M	medium	50 %
4	H	high	70 %
5	VH	very high	90 %

Table 5-2: Output fuzzy sets description

ID	Rules	Rule description	Normalized power output
1	rule NH	rule negative high (discharging)	$p_c = -100 \%$
2	rule NL	rule negative low (discharging)	$p_c = -50 \%$
3	rule ZE	rule zero (no operation)	$p_c = 0 \%$
4	rule PL	rule positive low (charging)	$p_c = +50 \%$
5	rule PH	rule positive high (charging)	$p_c = +100 \%$

Fuzzy control system

The normalized power output of the storage system can be determined on the fuzzy control system. The fuzzy inference system is a matrix of the “IF-THEN” rules. The conditional statements are presented in the Table 5-3. For example, if the power saldo is VH and SoC is VH, then the ESS is in discharging mode and discharged with power level of NH according to the control strategies. In corresponding to the two charging/discharging strategies, the two control systems, which are assigned with different conditional statements, are defined. In addition, SoC is under the constraint of the Equation (5-7), where SoC_{min} corresponds to VL = 10 % and SoC_{max} is set as VH = 90 %.

$$SoC_{min} \leq SoC(t) \leq SoC_{max} \quad (5-7)$$

Table 5-3: Conditional statements for control systems

Control system 1						Control system 2					
Saldo \ SoC	VL	L	M	H	VH	Saldo \ SoC	VL	L	M	H	VH
VH	NH	NH	NH	NH	NH	VH	ZE	NL	NL	NH	NH
H	NL	NL	NL	NL	NL	H	PL	ZE	NL	NL	NH
M	ZE	ZE	ZE	ZE	ZE	M	PL	PL	ZE	NL	NL
L	PL	PL	PL	PL	PL	L	PH	PL	PL	ZE	NL
VL	PH	PH	PH	PH	PH	VL	PH	PH	PL	PL	ZE

Assessment

In order to evaluate the effectiveness of the investigated virtual, cluster-related ESS with fuzzy control system, there are four items under two aspects from ESS itself and its influence on the power grid to be compared in this work. The maximum and minimum power saldo with/without ESS are obtained from the simulation results, to see how much the peak power will be reduced. Power supply quality should take into consideration the power gradient amplitude. The 15-min deviation of the normalized ESS power describes how fast the storage system can be operated at a certain gradient amplitude. The 15-min deviation of power saldo represents the degree of change in required power from external grid. The assessment is set up and shown in the Table 5-4.

Table 5-4: Assessment items

Aspects	Items	Symbol
ESS	Normalized power ESS	$p_c(t)$
	15-min deviation of normalized ESS power	$\Delta p_c(t) = p_c(t) - p_c(t - 1)$
Power grid	Power saldo with/without ESS	$P_{saldo}(t)$ and $P_{saldo_new}(t)$
	15-min deviation of power saldo with/without ESS	$\Delta P_{saldo}(t) = P_{saldo}(t) - P_{saldo}(t - 1)$

5.4 Simulation and Results

With the measurement of the high voltage feed-lines for the regional supply area, the regional power saldo is then generated according to the definition in the Section 2.2.2. The profiles of synthetic renewable power generation are shown in the Figure 5-5 (a). The network cluster 3 is taken as an observed area. The regional cluster is supplied by two substations and the power flow of feedlines, which are connected to these substations, are plotted in the Figure 5-5 (b). The power saldo is then derived from these two horizontal power flows. The first three days sampled data of June 2019 for the regional network cluster is used for the case study. The high negative power saldo indicates the high potential utilization of the surplus energy for the regional network cluster. The surplus power of the cluster can exceed -250 MW at noon.

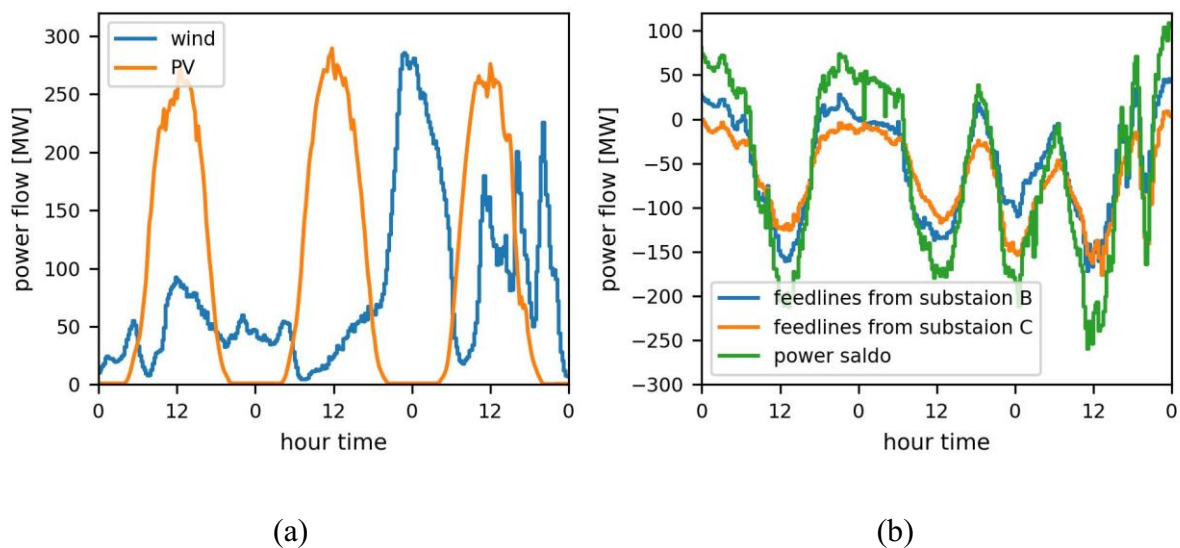


Figure 5-5: Profiles of RES generations (a), feedlines power and power saldo for Cluster 3 from 1st June to 3rd June (b) of year 2019

The procedure of the proposed simulation of the fuzzy logic control is shown in the Figure 5-6. It can be found that the fuzzy inference system (FIS) is under two control systems, which represents the two charging/discharging strategies. After getting the output from each fuzzy control system, the power output of ESS will be derived from the normalized power with ESS rated power, which is set as 100 MW. The capacity of the storage system is determined from the $e2p$ ratio with three different values as case studies. To obtain the new SoC in the next time step, the selection and setting of the $e2p$ ratio and control system play crucial roles. Thus, the SoC is generated from the normalized power and $e2p$ ratio based on the Equation (5-5). The operation time Δt in this study is set to be the same as the raw data temporal resolution of 15-min. The initial SoC for the simulation was set at medium level $M = 0.5$.

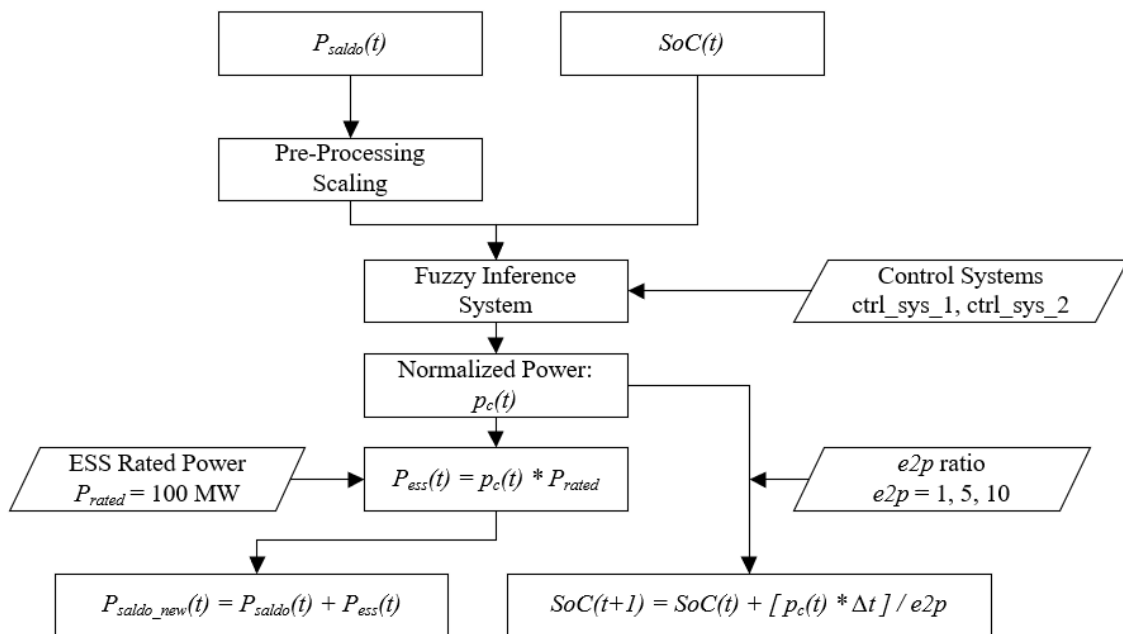


Figure 5-6: Procedure of the proposed simulation

In case studies the three different $e2p$ ratios (1, 5, 10) combined with two charging/discharging strategies are investigated and simulated. Figure 5-7 and Figure 5-8 show the normalized power of ESS and SoC by different $e2p$ ratios respectively. The two control systems are able to charge/discharge ESS. From the normalized power curve of ESS it can be seen that the ESS under control system 2 is more sensible. Control system 1 charges and discharges the ESS with a higher power rate and the ESS can reach its maximum or minimum SoC faster due to its control rules for PH and NH in the Table 5-3. However, the ESS is unable to be operated when it reaches its SoC limitation and it has to wait for the next available charging/discharging signal. It can be noted that the outcome from the control system 2 with $e2p = 5$ does not have much difference from the system with $e2p = 10$.

In regards to the dynamic range of SoC in the Figure 5-8, the capacity of ESS tends to be stable in a smaller range when a bigger capacity is implied. Higher depth of discharging has negative impact on the life cycle of ESS [84]. The extreme status of rest capacity can be avoided with control system 2. The SoC of ESS with $e2p = 10$ by strategy 2 remains in the range of 0.25 to 0.7.

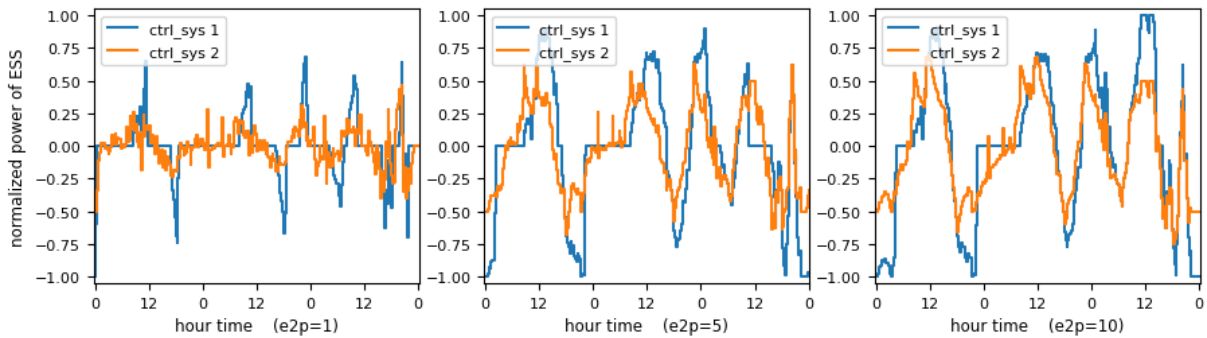


Figure 5-7: Normalized power output profiles

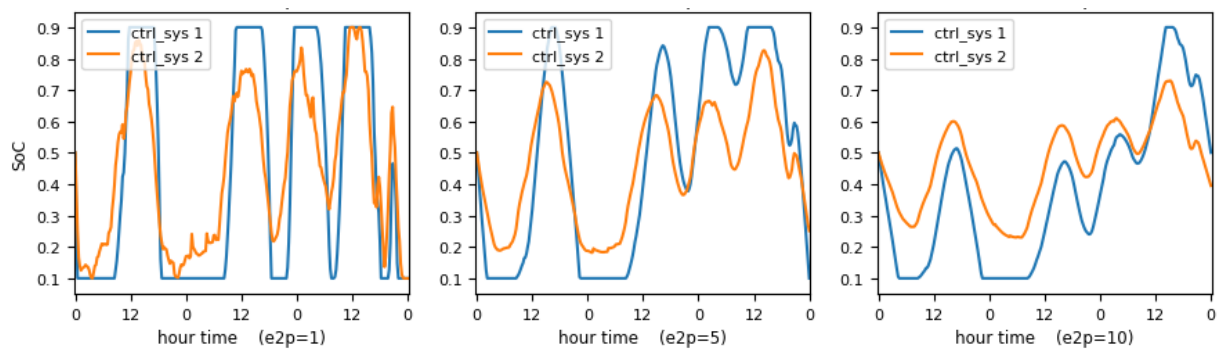


Figure 5-8: SoC of ESS

Descriptive statistics is used for the assessment of results. It can be found in the Figure 5-9 that ESS under control system 1 (blue line) can be operated at the rated power, which means the normalized power reaches an absolute value of 1. ESS under control system 2 (orange line) with $e2p = 1$ can be only charged or discharged at the normalized power of 0.5 (i.e. 50 %). With larger $e2p$ ratios both control systems can operate the ESS at a higher power rate. The reason is that a higher $e2p$ means a higher capacity ratio of the ESS, then SoC does not change very much at the same operating power. During the times of high power demand in grids and when the ESS is still near fully charged, the ESS can deliver high rate of power to the grid system.

As shown in the Figure 5-10, the 15-min deviation of the normalized power for control system 2 distributed in the range of -0.5 to 0.5, and the power change of ESS under the control system 1 is bigger than control system 2. No extreme charging/discharging power gradient is required for the ESS under control system 2 due to its intelligent control strategy.

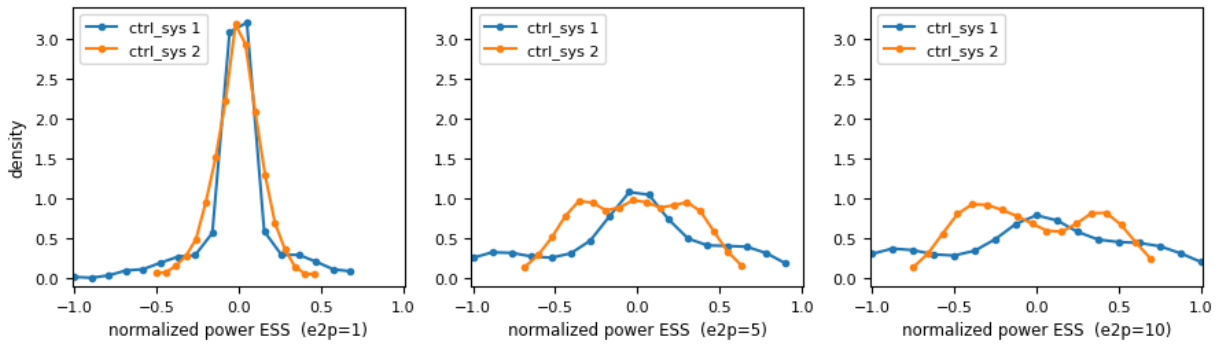


Figure 5-9: Distribution of normalized power ESS

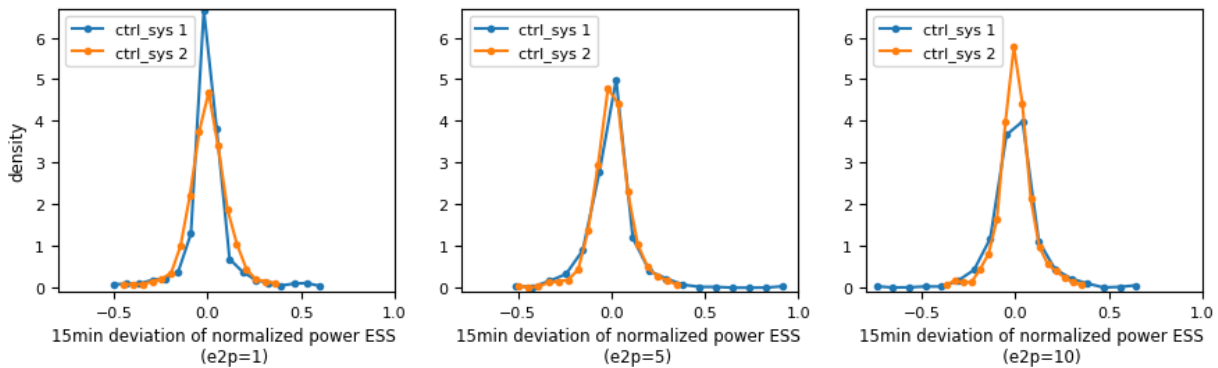


Figure 5-10: Distribution of 15-min deviation of normalized power ESS

In the Figure 5-11, the power profiles show the new power saldo results with ESS compared to the power saldo without ESS. It is noticeable that the two control systems are both able to match the high power saldo. Neither of the investigated control systems can reduce the peak power very well under $e2p = 1$. Nevertheless, the ESS with larger $e2p$ ratios gives a better performance.

Figure 5-12 shows the overall distribution profiles of the power saldo. A bimodal distribution of power saldo with positive and negative values implies that this region is an area of electricity consumption and production. By focusing on the maximum and minimum values, it shows how much peak power is reduced in terms of different control strategies and $e2p$ ratios. Both control systems deliver a better result in positive peak power shaving. By $e2p = 1$ and $e2p = 5$ the peak feed-back power (in negative) is not reduced by control system 1. That is because the ESS is already fully charged before the maximum surplus power occurs. Control system 1 charges the ESS directly without consideration of the mutual relation of current grid power imbalance and ESS available capacity.

It can be found by comparing the 15-min deviation of the power saldo with/without ESS (Figure 5-13) that the control system 1 aggravates the gradient of the power saldo. Aggregated power flows of electricity generation and consumption can be smoother under the control system 2. As a result, the frequent charging/discharging processes have a great influence on the stability

of power exchange with external grids. Therefore, the combination of observation of power saldo and status of ESS can give a better decision strategy for ESS. The ESS is controlled in such a way that it is not operated with a sudden high power in order to avoid a strong power change in the power grid. It contributes to the stability of the grid.

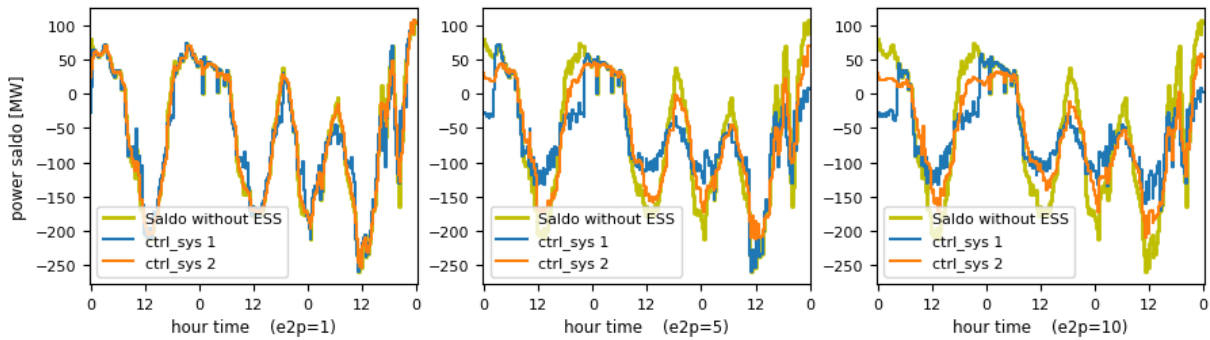


Figure 5-11: Power saldo profiles with / without ESS

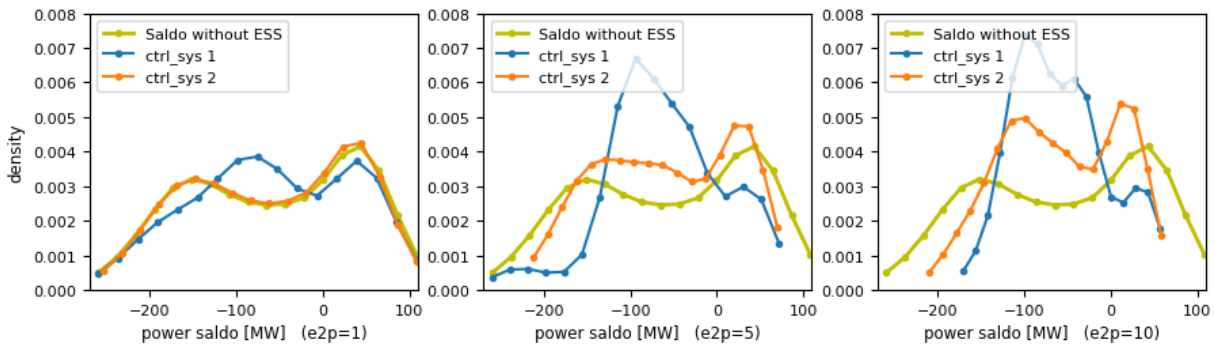


Figure 5-12: Distribution of power saldo with / without ESS

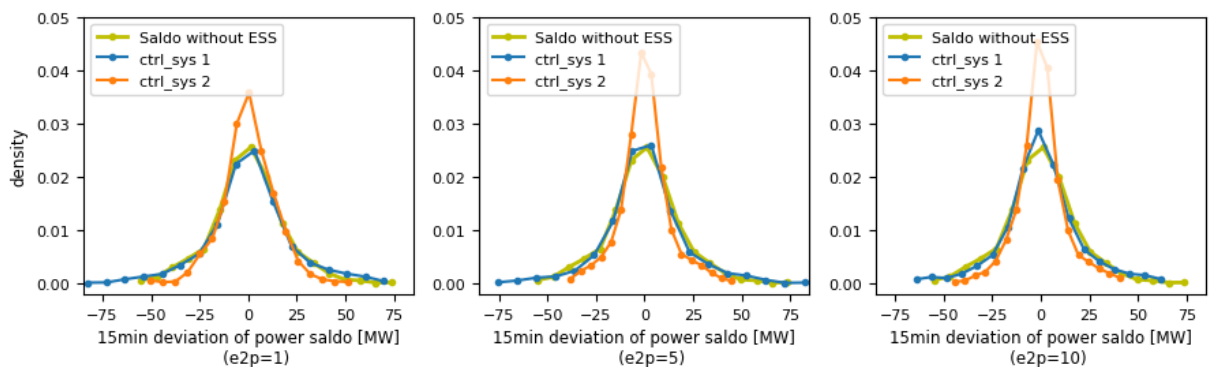


Figure 5-13: Distribution of 15-min deviation of power saldo with / without ESS

The calculated power saldo profile provides insights into the investigation of the ESS. The new investigated ESS can provide a smooth power curve based on the fuzzy control system. However, if the rest capacity of ESS is not considered (charging strategy 1), the investigation of ESS may bring additional challenges to the current power grid. The unexpected high power deviation may occur. However, the high power gradient of power saldo can also be avoided by

using ESS under control system 2, which combines the power grid situation of generation and consumption as well as the status of ESS. ESS only provides higher rate of power supply when power demand is high and the ESS is near fully charged.

5.5 Concluding Remarks

The idea of regional cluster is introduced along with a fuzzy approach to power saldo management. The supply-dependent electricity generation and demand are closely observed to investigate regional electricity surpluses. The investigation of the regional power load and generation is carried out based on the reduced network cluster. Two charging/discharging strategies for the virtual energy storage system are proposed and evaluated by using fuzzy logic. The regional power saldo and the *SoC* of ESS are used as input data to manage the local ESS. The normalized power output combined with rated power and different *e2p* ratios are then discussed in detail. After comparisons, it is found that the control strategy that observes the grade of power saldo and storage capacity performs better in implementing intelligent ESS control. Strategy 1 with a higher *e2p* ratio reduces peak power more effectively, but it faces challenges in maintaining the stability of the power exchange with external grids. Under strategy 2, extreme power requirements for fast fully charging or discharging is avoided, and the rest capacity of ESS provides power for local balance. This approach achieves an even profile of the resulting power saldo and effectively reduces the external power output. It is necessary to take a holistic view of the necessity of charging/discharging strategies for regional energy storage.

6 Implementation of Cluster Model in Cable Pooling Approach

This chapter implements the cable pooling approach to the regional network cluster model. Cable pooling allows adding additional RES installation to the existing RES system to increase the capacity factor of the local infrastructure. A combined system is proposed in this chapter to improve the efficiency of RES energy utilization. The specific installation capacity is evaluated based on single RES model. To extend the discussion, a regional model is established by using power measurements. The curtailment of energy and time duration, capacity factor and improvement of stability coefficient to the grid system are also discussed. The goal of this work is to evaluate the effect of cable pooling on the regional PV and wind power systems.

The roadmap of this chapter is structured as follows. Section 6.1 introduces the concept of the cable pooling. Section 6.2 describes the evaluation approaches for the combined system and the methodology for the system modelling. Section 6.3 discusses the results of the proposed models based on different case studies. Section 6.4 gives the conclusion remarks.

6.1 Overview of Cable Pooling

With the expansion of RES in electricity grid system, the current levelized cost of electricity (LCOE) of PV systems maintains 3.12 to 11.01 €Cent/kWh for different type of systems and solar irradiation, and wind power systems have a LCOE between 3.94 and 8.29 €Cent/kWh [93]. Delayed grid extension has hindered the integration of renewable energy and brings new challenges in grid congestion [94]. Generally, wind or PV power generation system has its own power cable and connection point to the electricity grid. In this way, the infrastructures of the cable and grid connection are not optimally used. Wind turbine power plant can only generate electricity when wind is blowing, and PV system feeds the electricity power in grid when the sun is shining. These two power plant systems rarely produce electricity at full capacity at the same time. In order to further reduce the costs and maximize the utilization of RES, the hybrid RES systems offer potential reliability benefits with stochastic wind and solar energies [95,96]. Cable pooling can solve the RES utilization of grid utility through 24-hour period and offers more opportunities for additional RES power system connections. Cable pooling is the concept that different RES generation systems share the cable and grid connection infrastructure [97]. Figure 6-1 presents two connection scenarios, independent grid connection and cable pooling with shared grid connection. Due to the shared grid connection, the capacity factor (CF) of the point of common coupling (PCC) is increased in comparison with the independent grid connection. For instance, PV and wind farms have CF of 11 % and 19 % respectively with

separated grid connection, but cable pooling allows a more efficient use of the existing grid connections and power cables. Obviously, the implementation of such a solution increases available connection capacity without additional expenditures on grid extension.

The purpose of this work is to provide a solution for the composite utilization of wind and solar power generation. The system adopting the solution of cable pooling enables large-scale development by utilizing the complementary advantages of the two power generation systems. This solution reduces the investment costs of power transmission and distribution. A techno-economic feasibility study of integration offshore PV and wind parks is investigated based on an existing Dutch wind farm [97]. Due to the inherent characteristics of renewable energies, the cable and grid connection are not fully used. The floating PV system could be added to the wind farm, but a forced curtailment is inevitable for the electricity overproduction due to the power limitation of the existing infrastructure. Wind power could compensate for the lack of PV power production at night time. They complement each other perfectly. As a result, it provides a more stable power supply for the shared grid connection at PCC.

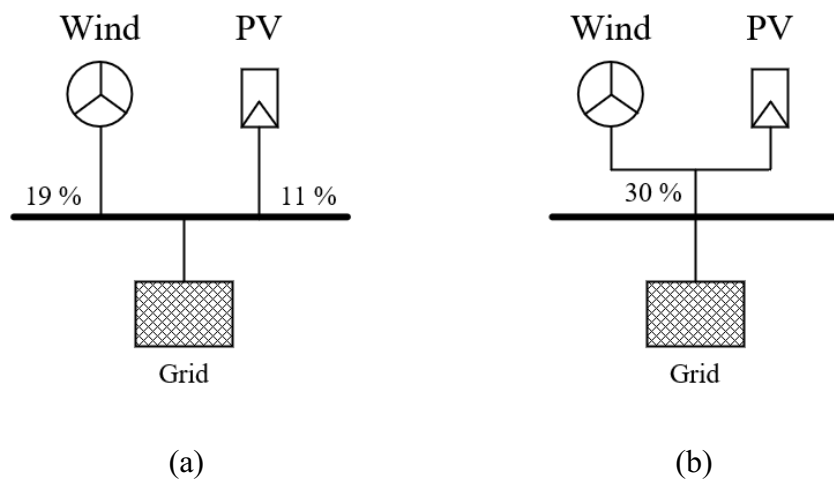


Figure 6-1: Independent grid connection (a) and cable pooling with shared grid connection (b)

6.2 Methodology for the System Modelling

6.2.1 Assessment Approaches

A combined system from PV and wind power systems could be evaluated by different metrics, such as capacity factor, stability coefficient and energy curtailment. These metrics are introduced in the following.

Capacity factor

Capacity factor (CF) is a measure of how much energy is produced by a power plant compared with its maximum power output. Equivalent to the full load hours, the capacity factor is the ratio of the actual electrical energy output over the maximum electrical energy output. It describes the continuous operation time at full capacity over one period. The definition of CF is described as:

$$CF = \frac{\text{actual energy output [MWh]}}{\text{power capacity [MW]} * \text{time period [h]}} * 100\% \quad (6-1)$$

Under a certain situation such as weather condition and installation, different technologies and regions have a various capacity factors. The capacity factor of a total combined system (CF_{total}) from wind and PV power generations can be calculated by the Equation (6-2). In order to broaden the utilization and comparison, the ratio of installed capacities of PV and wind systems is applied and converted in the Equation (6-3).

$$CF_{total} = \frac{E_{total}}{P_{inst,total} * \Delta T} = \frac{P_{inst,PV} * CF_{PV} + P_{inst,wind} * CF_{wind}}{P_{inst,PV} + P_{inst,wind}} \quad (6-2)$$

$$CF_{total} = \frac{\text{ratio} * CF_{PV} + CF_{wind}}{\text{ratio} + 1}, \text{ with ratio} = \frac{P_{inst,PV}}{P_{inst,wind}} \quad (6-3)$$

where E_{total} means the actual energy feed-in to the grid. $P_{inst,total}$ donates the sum installed capacity of the combined system. ΔT donates the observed time period.

Stability coefficient

An assessment of temporal complementarity for wind and PV systems is clarified in [95,98] for co-locating wind and PV combined systems. The stability coefficient (C_{stab}) is a robust indicator for the assessment for combined systems to quantify the synergies of PV and wind power [95,98]. The smoothness of power output from the combined systems is evaluated by the improvement of the stability coefficient in comparison with the existing reference system. Because the capacity factor of wind power system is higher than PV power system [97], wind power plant system is set as reference system in this work. The stability coefficient is then calculated based on the reference power system. Variance of the temporal capacity factor indicates the power output variability. It can be calculated by the variance of the combined system and the wind power system:

$$C_{stab} = 1 - \frac{Var(CF_{total})}{Var(CF_{Wind})} \quad (6-4)$$

where $Var(CF_{Wind})$ and $Var(CF_{total})$ donate the variance of capacity factor of wind power system and the total combined system respectively. As a result, high complementarity of PV and wind system leads to a high value of stability coefficient. By definition, the $C_{stab} = 0$ means that the combined system does not improve the stability compared to the existing wind system. When the combined system has the best synergy effect, which means the two systems compensate each other perfectly, then C_{stab} can reach the maximum value of 1.

Curtailement

The main constraint of the cable pooling is that the power transport and grid connection is limited by the existing RES power system. The cumulative power generation of the combined system can be obtained by adding the two technologies of PV and wind power generations by the Equation (6-5). The surplus power should be temporarily curtailed when the cumulative power generation exceeds the power cable limitation and grid connection capacity. The curtailed energy can be calculated by the Equation (6-6).

$$P_{gen}(t) = P_{wind}(t) + P_{PV}(t) \quad (6-5)$$

$$E_{curt} = \sum (P_{gen}(t) - P_{limit}) * \Delta t, \quad \text{where } P_{gen}(t) > P_{limit} \quad (6-6)$$

where $P_{gen}(t)$ means the cumulative power generation from PV and wind systems before curtailement. E_{curt} donates the curtailed energy due to the cable pooling limitation. P_{limit} donates the limited power of grid connection. Δt means the time step (i.e. 1/4h) of time series. The fraction of curtailed energy (E_{curt}) to total generation (E_{gen}) in the Equation (6-7) represents the percentage of energy curtailement. In addition, the duration time of curtailement measure can be expressed by the curtailement time percentage t_{curt} in the Equation (6-8).

$$f_{curt} = \frac{E_{curt}}{E_{gen}} = \frac{E_{curt}}{\sum P_{gen}(t) * \Delta t} * 100\% \quad (6-7)$$

$$t_{curt} = \frac{\sum t, \text{ where } P_{gen}(t) > P_{limit}}{\sum t} * 100\% \quad (6-8)$$

6.2.2 System Modelling

Wind power output performance

The four phases of power generation from wind turbines (WT) is introduced in the Section 3.2.1. The growth of wind power generation follows an “S-shape” curve with wind speed. Therefore, an asymmetric *sigmoid* function could be used to mimic the normalized WT power output between the cut-in and cut-out wind speed. It is described by the Equation (6-9) with the variable of wind speed at hub height h . Three parameters (M, B, A) control the shape of the curve including steepness of the slope and symmetry of the curve. They can be determined through hyper-parameter optimization of the WT power curve. In addition, the wind speed can be converted to a specific WT hub height h by the Equation (3-7), which can be found in the Section 3.2.2.

$$p_{norm, Wind}(v_h) = \begin{cases} 0 & (v_h \leq v_{cut-in}) \\ \frac{1}{\left(1 + e^{\frac{-(v_h - M)}{B}}\right)^A} & (v_{cut-in} < v_h \leq v_{cut-out}) \\ 0 & (v_h > v_{cut-out}) \end{cases} \quad (6-9)$$

where v_h donates the wind speed at hub height h . M, B, A donate the control parameters for asymmetric *sigmoid* function. v_{cut-in} and $v_{cut-out}$ represent the cut-in and cut-out wind speed for WT respectively.

One representative WT of Enercon E82/2300 (2.3 MW) with 138 meter hub height is used for single WT model. The normalized power simulation result is shown in the Figure 6-2. The parameters used are listed in the Table 6-1. The WT power curve can be obtained from the database of The Wind Power [99].

Table 6-1: Parameter setting for WT model

Parameters	M [m/s]	B [m/s]	A [-]	v_{cut-in} [m/s]	$v_{cut-out}$ [m/s]
Values	9.3	1.3	0.8	2	25

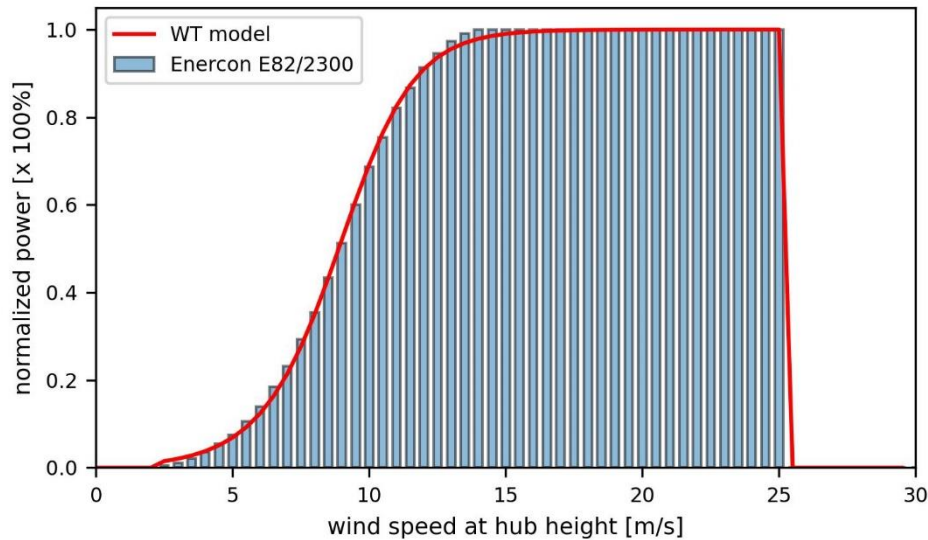


Figure 6-2: Wind power output performance model

PV power output performance

PV power generation can be as well calculated based on the PV power output characteristic with the solar irradiation and air temperature. The specific conversion process will not be repeated in this Chapter, and the details can be obtained from the Section 3.3.1. The following Table 6-2 gives the typical values for PV systems used in this work.

Table 6-2: Parameter setting for PV model

Parameters	Temperature coefficient	Parameter	Inverter efficiency
	α [%/°C]	γ [°C*m ² /W]	η_{inv} [%]
Values	0.43	0.02	95

Regional model based on power measurements

In order to study the effects of the concept of cable pooling on the regional grid system level, an assumption is made that all installed PV and wind power systems within the regional cluster are interconnected in the way of cable pooling. Regardless of voltage levels, the cumulative power generation is the object of this study. From the perspective of grid system operator at superordinate level, the sophisticated connections of regional RES systems are not taken into consideration, but the synergistic effects between the RES technologies are more important for the effective regional energy supply and grid stability. Therefore, the regional network cluster model established in the previous Chapter 2 could be used for this study. The normalized power

$p_{norm,RES}(t)$ can be converted by the Equation (6-10) based on the power generation $P_{RES}(t)$ and the installed capacity $P_{inst,RES}(t)$. Furthermore, it can be applied in the assessment indicator of capacity factor. The average regional capacity factor for the specific RES system in the investigated area is calculated by the Equation (6-11).

$$p_{norm,RES}(t) = \frac{P_{RES}(t)}{P_{inst,RES}(t)} \quad (6-10)$$

$$\begin{aligned} CF_{RES} &= \frac{\sum [P_{RES}(t) * \Delta t]}{P_{inst,RES} * \Delta T} = \frac{\sum [(p_{norm,RES}(t) * P_{inst,RES}) * \Delta t]}{P_{inst,RES} * (N * \Delta t)} \\ &= \frac{\sum p_{norm,RES}(t)}{N} \end{aligned} \quad (6-11)$$

where N donates the sample size of time series within the observed time period of ΔT , for example, there are 35040 ($35040=365*24*4$) sampling data for one year (365 days) based on 15-min time resolution. Δt donates the time step (i.e. 1/4h) of time series.

6.3 Results of Case Studies for Single and Regional RES Models

In the following section, the complementarity of the combined system is discussed based on one single wind turbine model and the extended regional cluster model. The averages of regional weather measurements from the year 2019 are used for the calculation of the wind and PV power output. The weather data and regional RES generation used in this work are extracted from the regional measurements. These data sources are already introduced in the Section 2.3.

Case study for WT and PV model

Figure 6-3 shows the cumulative power curve of the combined system from different PV capacities with one wind turbine generator. It can be seen that by adding PV capacity, more electricity power can be generated and transported to the grid. For example, *wind_2.3+PV_4.0* shows the combined system of 2.3 MW WT with 4.0 MW PV system, and their power generation can reach more than 5 MW. The power generation that exceeds the power limitation, which is set as the WT installed capacity of 2.3 MW, should be curtailed, in order to ensure the grid security. Table 6-3 shows the effects of additional PV system to the WT system based on the weather data of year 2019. WT with 2.3 MW delivers about 6633 MWh energy over one year. The average capacity factor of the single WT system is about 32.9 %, which is equivalent to approx. 2880 full load hours. The reason is that wind turbine can only work when the wind speed is higher than the minimum cut-in speed (i.e. 2 m/s), and if the wind speed is too high

(i.e. over 25 m/s), the wind turbine must be shut down due to the protection mechanism. The capacity factor of the combined system $wind_2.3 + PV_0.5$ is increased by 2.5 % due to the additional 0.5 MW PV in the way of cable pooling connection. The actual total feed-in energy E_{total} is about 7127 MWh, and only 0.4 % energy needs to be curtailed. These curtailment events occur for about 330 hours, which equates to 3.8 % time of one year. Nevertheless, the combined system brings about 20.8 % stability to the system in comparison with the existing WT system.

Table 6-3: Effects of adding different PV systems to one WT system

combined system	E_{gen}	E_{total}	E_{curt}	f_{curt}	t_{curt}	CF	C_{stab}
	[MWh]	[MWh]	[MWh]	[%]	[%]	[%]	[%]
wind_2.3 + PV_0.0	6633	6633	0	0	0	32.9	0
wind_2.3 + PV_0.5	7157	7127	30	0.4	3.8	35.4	20.8
wind_2.3 + PV_1.0	7682	7579	103	1.3	6.2	37.6	34.4
wind_2.3 + PV_1.5	8206	7997	209	2.5	8.1	39.7	43.7
wind_2.3 + PV_2.0	8730	8384	346	4.0	9.8	41.6	50.4
wind_2.3 + PV_2.5	9254	8737	517	5.6	11.7	43.4	55.3
wind_2.3 + PV_3.0	9778	9052	726	7.4	13.8	44.9	59.2
wind_2.3 + PV_3.5	10302	9322	980	9.5	16.1	46.3	62.5
wind_2.3 + PV_4.0	10826	9551	1275	11.8	18.1	47.4	65.3

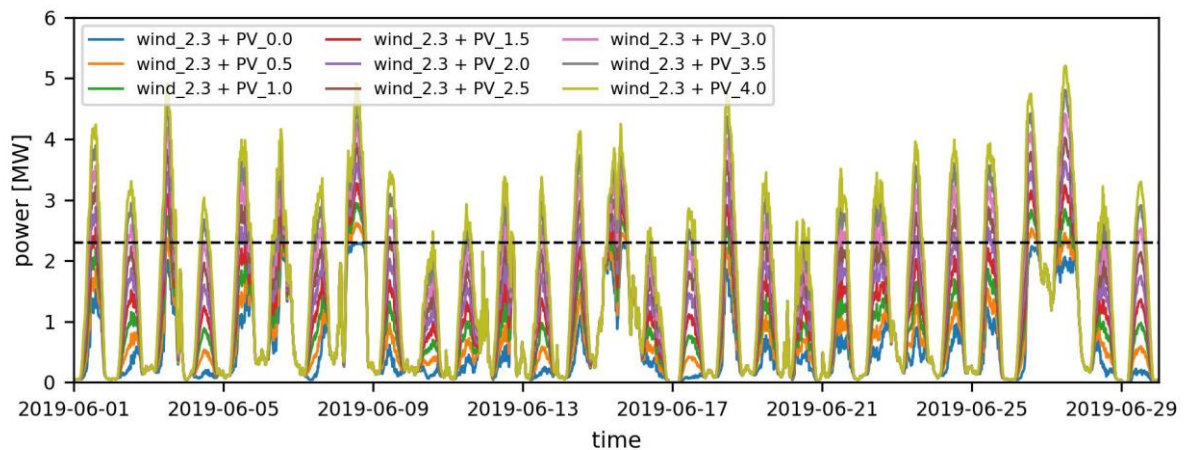


Figure 6-3: Power curves of combined system with different PV systems

Case study for the regional power measurements

Furthermore, one regional level case study is carried out based on power measurements in one DSO supply area. The developments of the regional RES installed capacity over last years are shown in the Figure 6-4. Monthly capacity factors of the regional wind or PV power systems are calculated and illustrated in the Figure 6-5. It highlights the monthly changes of the capacity factors. Monthly average of capacity factor for wind systems is about 12 % in summer and 30 % in winter. In addition, the capacity factor for PV systems in this region is approx. 2 % in winter and up to 20 % in summer. The annual average capacity factors are around 20 % and 11 % for wind and PV systems respectively. To make full and efficient use of the local infrastructure, electricity energy from wind and PV systems could be fed into grid via sharing cables and grid connection points. Taking advantage of seasonal differences, the combined system provides electricity power primarily from the wind energy in winter and solar energy in summer.

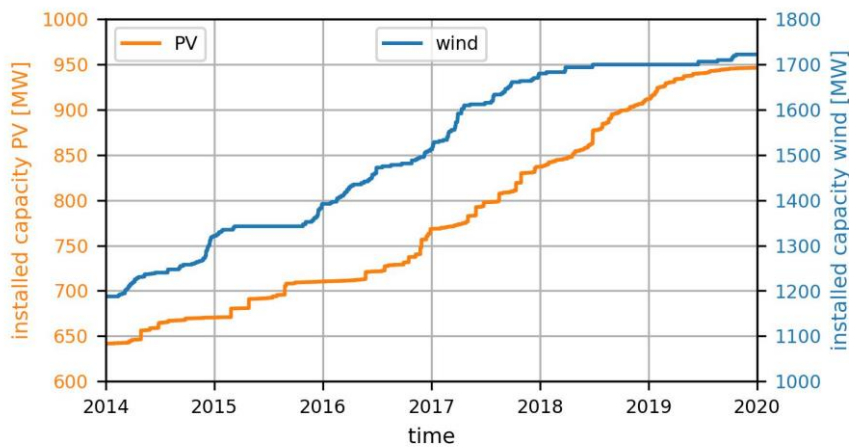


Figure 6-4: Regional installed capacity of PV and wind power systems from 2014 to 2019

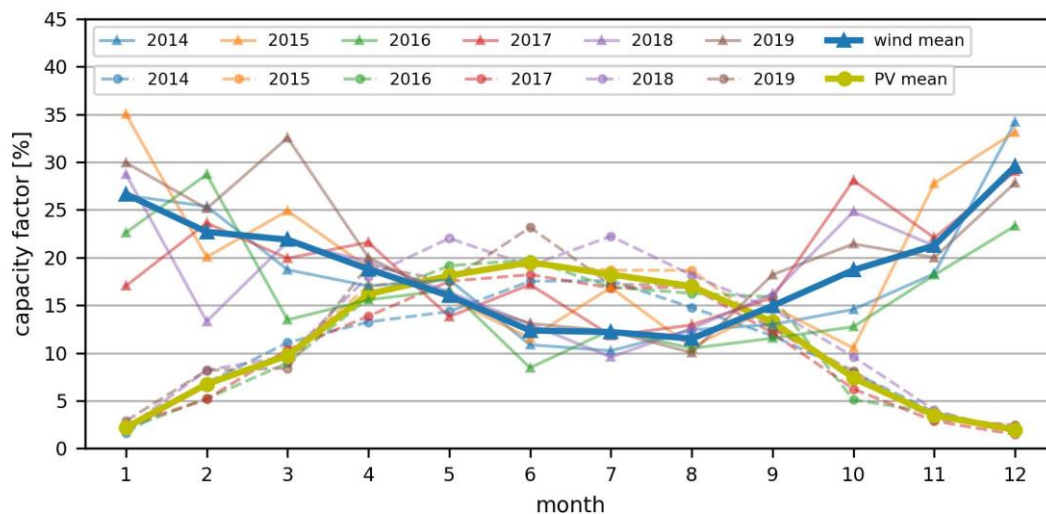


Figure 6-5: Capacity factor from regional measurements for PV and wind power systems

Comparison of different capacity ratios

Figure 6-6 reflects the variation in energy curtailment and capacity factor for different capacity ratios of PV to wind power system. Focusing on the curtailment percentage, the results provide more insights on the combination of the cable pooling system. There are big differences between the single WT and PV model based on calculation and regional cluster model based on regional power measurements. The power limitations of both models are set according to the installed wind power capacity. Optimal weather condition could be available for one single WT model. However, the total power output for large-scale RES systems would be smoothed due to the different spatial and temporal weather conditions, even different performance characteristics of each power system. It is difficult for large-scale RES systems to reach their nominal power at the same time. The maximum wind power output for the regional scale is about 80 % of nominal power, which is shown in the Figure 4-17 in Section 4.2.2. While for the single WT model, the theoretical maximum wind power can be achieved, the combined system with additional PV will inevitably exceed the capacity of cable or PCC. Therefore, the curtailment for one WT plus PV model occurs more frequently and earlier. For the regional model, the energy curtailment occurs slowly when capacity ratio is greater than one. The capacity factor still rises consistently from 20 % to over 30 % as the capacity ratio ranges from 0 to 1.

Significantly, more electricity is generated from the combined system with additional RES installations. Despite the improvement of capacity factor, the surplus power generation leads to a large amount of curtailment based on the cable-pooling concept. Figure 6-7 reveals the variation of stability coefficient due to new installed RES. It indicates the optimum capacity ratio for one investigated area. In the case of no additional energy curtailment in the way of cable pooling, the more integration of fluctuating PV system would degrade the stability coefficient of the grid. The new added PV system is inherently volatile, and the existing wind system can only compensate part of the PV power fluctuation. Therefore, when the capacity factor is greater than approx. 0.8 for regional model, the stability coefficient C_{stab} begins to decline. Only through energy curtailment, the power fluctuation can be then limited to improve the grid stability. In the case of power limitation, the maximum feed-in power remains constant by energy curtailment for the reason of cable capacity limitation. Due to the smoothing effect of the large-scale RES power generation, much more buffer space is for the new integration of RES power without curtailment events in the way of cable pooling. Starting from the capacity ratio of about 1.5, the grid stability coefficient at the regional level increases slowly due to the increase in energy curtailment. Two influencing factors for the increase in the stability

coefficient are the smoothing effect of the regional RES power and the power curtailment or limitation of feed-in power.

The current capacity ratios of the regional four clusters are as well marked in the Figure 6-7. Cluster 1 and cluster 3 are near the optimal range of stability coefficient. Cluster 4 has more installation of PV system in relative to wind power system. In order to increase the stability coefficient, it can be suggested that cluster 4 could install more wind power, and more PV system could be integrated into cluster 2. Cluster 1 and 3 can maintain the current synergic development trend of PV and wind power systems.

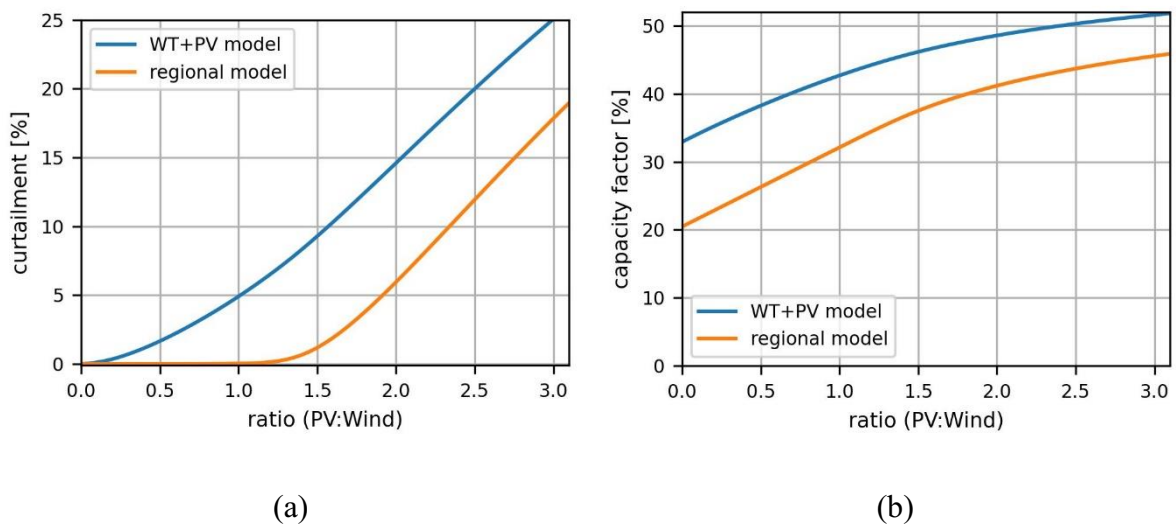


Figure 6-6: Energy curtailment (a) and capacity factor (b) for different capacity ratios

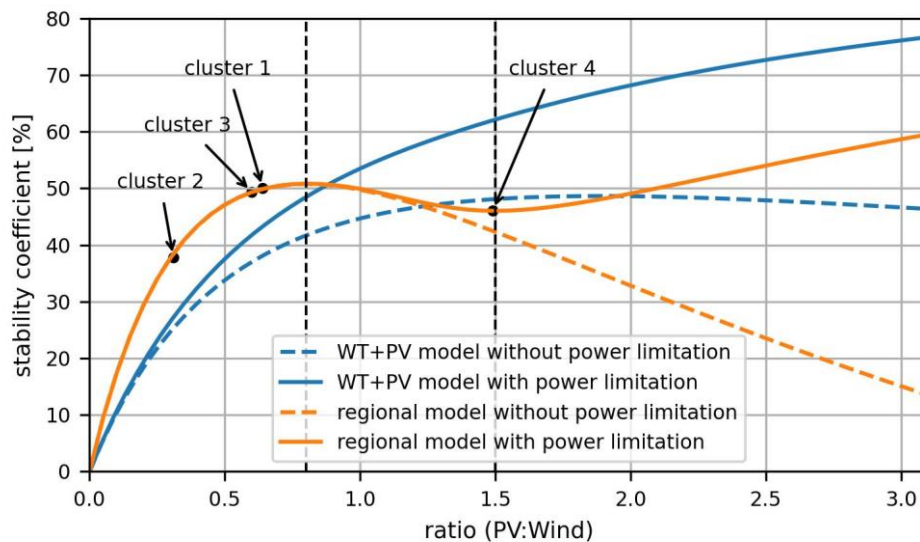


Figure 6-7: Stability coefficient for different capacity ratios

6.4 Concluding Remarks

It is noted that the local infrastructure is often underutilized during power production from wind or PV systems. This chapter presents a cable pooling approach that combines wind and PV systems to increase the capacity factor of the local infrastructure. Various metrics are used to evaluate the effectiveness of this approach. The regional model based on power measurements is further implemented and compared with a single WT and PV model. Theoretical curtailment and capacity factor improvements are discussed. The results of the proposed methods indicate that the power generation from PV and wind systems can balance each other due to the synergy effect. It would be appreciated at the regional level to take into account which ratio of RES installation can bring a greater stability coefficient to the grid. The concept of cable pooling applied to the regional cluster model provides more insights into the stability coefficient. As the relative capacity ratio increases, the instability caused by the fluctuation of PV power system becomes more evident. This means that even if there is no power limitation, the newly added PV system cannot bring stability to the existing grid, and the synergy effect between the two systems does not perform well. Therefore, the combination of different PV systems should take into account the currently existing wind power installations. The regional effect of cable pooling could decrease the stability coefficient of the grid and leads to a considerable amount of curtailment.

7 Conclusion and Outlook

7.1 Summary

In this dissertation, the focus is on the development and application of a network cluster model for interpreting the regional power generation and consumption. The regionalized network cluster model is specifically designed to fit the temporal and spatial features of the regional power systems. The proposed network cluster functions as an aggregator, encompassing power generation and consumption, and improving the observability of regional power performance. Furthermore, the effectiveness of the cluster model has been demonstrated in various applications, including large-scale RES power estimation, energy storage and cable pooling analysis throughout this thesis.

Chapter 2 presents the concept of a regional network cluster and creates different spatial and temporal databases for the regional grid-oriented network cluster. Regardless of the detailed connection of the grid components, the reduced network cluster model can be used as a tool to demonstrate the renewable energy allocation and regional power balance characteristics.

Chapter 3 and Chapter 4 present both the physical profile-based aggregated models and data-driven AI methods for the regional large-scale PV and wind power estimation. Due to the spatial inhomogeneity and temporal uncertainty of the distributed power generation, the proportion of fluctuating power generation from solar and wind energies is modelled and calculated according to the regional technologies and meteorological conditions. Different dimensions of the input features has been explored to improve the prediction accuracy. Physical aggregated model provides more energy conversion in detail, while AI models as black box have a better prediction performance.

Chapter 5 discusses the further application of network cluster in the virtual regional energy storage system, which adopts appropriate charging/discharging strategies according to the current power difference between generation and load. The remaining capacity of the ESS can provide power to stabilize the grid, minimize the fluctuation effect from renewable power, and obtain a smooth profile of power balance.

Chapter 6 presents the concept of cable pooling with shared grid connection to deal with the more effective utilization of the regional RES power generation based on the existing infrastructure. A combined system with cable pooling is studied in terms of curtailment percentage, capacity factor and stability coefficient. With the help of the regional network cluster, it reveals how much more stability coefficient can be obtained due to additional installations.

Overall, the proposed cluster model is well suited to capture the temporal and spatial features of regional RES power generation, and it provides a useful tool for assessing the potential of renewable energy sources and enhancing the stability and performance of regional power systems. The research questions and goals presented in the introduction have been successfully addressed through the development and application of the regionalized network cluster model.

7.2 Future Work

With the present work, the current state of the regional power generation and consumption is presented. The continuous update and refinement of the databases and models is essential. The spatial and temporal databases will be updated due to new installations and time series measurements.

The cluster model could be applied in further studies for the regional operation and stabilization of the local grids. The regional power plant system should be flexible with regard to how, where and when the power is consumed and produced, ensuring that it can always deliver electricity and utilize further storage and other technologies such as power-to-x to match the amount of the usage. More flexibility services related to the security of electricity supply could be regionally analysed based on the network cluster tool.

A good estimation performance of regional large-scale PV and wind power based on the cluster model can be obtained, but the estimation methods still have the following limitations. The mean WT hub height of each cluster is statistically calculated. However, regional decentralised weather measurements do not directly represent the actual wind speed at hub height. The transformation process of wind speed inevitably leads to errors for each cluster. Additionally, average values for two mounted types of PV systems are used. Furthermore, it is difficult to determine the actual operating conditions of individual PV and WT power systems throughout the year. Therefore, to make accurate predictions of regional RES power, it is recommended to use more measurement sites with corresponding spatial partitions. Nevertheless, using more data and small-scale partitions will require more computational cost, so a balance needs be found. An alternative option for comparison is different classifications of RES systems. Artificial intelligence with neural networks can be a solution to avoid the analysis of transformation processes. The internal mapping relations of input features can be trained and adjusted by self-learning. However, neural networks are understood as black boxes that lack explanatory power, and their complex structures require more training processes. Large

amounts of data are required, and the selecting appropriate input features and models is important for the prediction outcomes.

These useful and impressive results encourage more research activities in multiple inputs and charging/discharging strategies for the local energy storage system. It should be possible for the controller to generate effective control rules under different pre-processing, such as standard scaling, robust scaling, and new membership functions for different storage systems.

Specifically, the current regional PV and wind systems are not completely installed in the way of cable pooling concept. However, the model results give a future regional development direction. In order to get an effective application at the technical level, the regional wind and PV power systems need to take the advantage of synergy effect between them, which can effectively increase the use of cables and bring stability to the local grid. In addition, economic feasibility could be added for the further analysis based on these regional aspects.

Furthermore, there is an increasing need for open data platforms to exchange and manage the regional power analysis. The regional network cluster could be extended to different voltage levels, and be able to transfer the application on different sub-regions. Power system analysis is becoming increasingly interdisciplinary. Familiarity with geographic data, computer science, understanding of meteorological background, and power system characteristics, these abilities come together to better establish the foundation of the research. It is crucial for the regional development and effective utilization of renewable energies.

Bibliography

- [1] International Energy Agency (IEA), Renewables 2021 Analysis and forecasts to 2026, 2021. <https://www.iea.org/reports/renewables-2021>.
- [2] Federal Ministry for Economic Affairs and Climate Action (BMWK), Renewable energy sources in figures National and International Development 2020, Berlin, 2021. <https://www.bmwk.de/Redaktion/EN/Publikationen/Energie/renewable-energy-sources-in-figures-2020.html>.
- [3] Bundesministerium für Wirtschaft und Energie (BMWi), Kommission „Wachstum, Strukturwandel und Beschäftigung“, 2019. https://www.bmwk.de/Redaktion/DE/Downloads/A/abschlussbericht-kommission-wachstum-strukturwandel-und-beschaeftigung.pdf?__blob=publicationFile&v=4.
- [4] Bundesnetzagentur (BNetzA), Bedarfsermittlung 2021-2035 Bestätigung Netzentwicklungsplan Strom, 2022. <https://www.netzentwicklungsplan.de/de/netzentwicklungsplaene/netzentwicklungsplan-2035-2021>.
- [5] Netztransparenz.DE, EEG-Anlagenstammdaten, (n.d.). <https://www.netztransparenz.de/EEG/Anlagenstammdaten> (accessed May 16, 2022).
- [6] H. Schwarz, Will Germany move into a situation with unsecured power supply?, *Front. Energy*. 13 (2019). <https://doi.org/10.1007/s11708-019-0641-z>.
- [7] Agentur für Erneuerbare Energien, Bundesländer-Übersicht zu Erneuerbaren Energien Anteil Erneuerbarer Energien am Bruttostromverbrauch, (n.d.). <https://www.foederal-erneuerbar.de/uebersicht/bundeslaender/> (accessed November 1, 2021).
- [8] B. Yu, D. Fang, H. Yu, C. Zhao, Temporal-spatial determinants of renewable energy penetration in electricity production: Evidence from EU countries, *Renew. Energy*. 180 (2021). <https://doi.org/10.1016/j.renene.2021.08.079>.
- [9] D. Baltensperger, A. Buechi, F.R.S. Sevilla, P. Korba, Optimal Integration of Battery Energy Storage Systems and Control of Active Power Curtailment for Distribution Generation, in: *IFAC-PapersOnLine*, 2017. <https://doi.org/10.1016/j.ifacol.2017.08.1542>.
- [10] Next Kraftwerke GmbH, Was ist Einspeisemanagement?, (2021). <https://www.next-kraftwerke.de/wissen/einspeisemanagement> (accessed December 2, 2021).

- [11] Bundesnetzagentur (BNetzA), Monitoring Report 2021, Bonn, 2022.
https://www.bundesnetzagentur.de/SharedDocs/Downloads/EN/Areas/ElectricityGas/CollectionCompanySpecificData/Monitoring/MonitoringReport2021.pdf?__blob=publicationFile&v=2.
- [12] Bundesnetzagentur (BNetzA), Monitoring Report 2020, Bonn, 2021.
https://www.bundesnetzagentur.de/SharedDocs/Downloads/EN/Areas/ElectricityGas/CollectionCompanySpecificData/Monitoring/MonitoringReport2020.pdf?__blob=publicationFile&v=2.
- [13] Bundesnetzagentur (BNetzA), Bericht zum Zustand und Ausbau der Verteilernetze 2021, Bonn, 2022.
https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Energie/Unternehmen_Institutionen/NetzentwicklungUndSmartGrid/ZustandAusbauVerteilernetze2021.html.
- [14] BDEW, DSO 2.0 - Verteilnetzbetreiber der Zukunft, (2018).
<https://www.bdew.de/energie/dso20/> (accessed December 1, 2021).
- [15] L. von Bremen, Large-Scale Variability of Weather Dependent Renewable Energy Sources, in: A. Troccoli (Ed.), *Manag. Weather Clim. Risk Energy Ind.*, Springer Netherlands, Dordrecht, 2010: pp. 189–206. https://doi.org/10.1007/978-90-481-3692-6_13.
- [16] A. Steiner, C. Köhler, I. Metzinger, A. Braun, M. Zirkelbach, D. Ernst, P. Tran, B. Ritter, Critical weather situations for renewable energies – Part A: Cyclone detection for wind power, *Renew. Energy*. 101 (2017).
<https://doi.org/10.1016/j.renene.2016.08.013>.
- [17] T. Schmid, *Dynamische und kleinräumige Modellierung der aktuellen und zukünftigen Energienachfrage und Stromerzeugung aus Erneuerbaren Energien*, Technische Universität München, 2019.
- [18] M. Beer, *Regionalisiertes Energiemodell zur Analyse der flexiblen Betriebsweise von Kraft-WärmeKopplungsanlagen*, Technische Universität München, 2012.
- [19] N. Aoun, Methodology for predicting the PV module temperature based on actual and estimated weather data, *Energy Convers. Manag.* X. 14 (2022).
<https://doi.org/10.1016/j.ecmx.2022.100182>.
- [20] P. Balakishan, I.A. Chidambaram, M. Manikandan, *Smart Fuzzy Control Based Hybrid*

- PV-Wind Energy Generation System, Mater. Today Proc. (2021).
<https://doi.org/10.1016/j.matpr.2021.07.074>.
- [21] E. Lorenz, T. Scheidsteger, J. Hurka, D. Heinemann, C. Kurz, Regional PV power prediction for improved grid integration, Prog. Photovoltaics Res. Appl. 19 (2011).
<https://doi.org/10.1002/pip.1033>.
- [22] S. Stock, D. Babazadeh, C. Becker, Applications of artificial intelligence in distribution power system operation, IEEE Access. 9 (2021).
<https://doi.org/10.1109/ACCESS.2021.3125102>.
- [23] J. Xie, I. Alvarez-Fernandez, W. Sun, A review of machine learning applications in power system resilience, in: IEEE Power Energy Soc. Gen. Meet., 2020.
<https://doi.org/10.1109/PESGM41954.2020.9282137>.
- [24] Deutsche Energie-Agentur (dena), Artificial Intelligence for the Integrated Energy Transition Assessing the technological status quo and categorising fields of application in the energy industry, Berlin, 2019.
<https://www.dena.de/newsroom/publikationsdetailansicht/pub/dena-analyse-kuenstliche-intelligenz-fuer-die-integrierte-energiewende/>.
- [25] J. Scheffler, Verteilnetze auf dem Weg zum Flächenkraftwerk - Rechtlicher Rahmen, Erzeuger, Netze, 1st ed., Springer Vieweg Berlin, Heidelberg, 2016.
<https://doi.org/10.1007/978-3-642-55297-7>.
- [26] K. Heuck, K.-D. Dettmann, D. Schulz, Elektrische Energieversorgung - Erzeugung, Übertragung und Verteilung elektrischer Energie für Studium und Praxis, 8th ed., Vieweg+Teubner Verlag Wiesbaden, 2010. <https://doi.org/10.1007/978-3-8348-9761-9>.
- [27] A.J. Schwab, Elektroenergiesysteme - Smarte Stromversorgung im Zeitalter der Energiewende, 6th ed., Springer Vieweg Berlin, Heidelberg, 2020.
<https://doi.org/10.1007/978-3-662-60374-1>.
- [28] CEN-CENELEC-ETSI Smart Grid Coordination Group, Smart Grid Reference Architecture, 2012. [https://www.cencenelec.eu/media/CEN-CENELEC/AreasOfWork/CEN-CENELEC_Topics/Smart Grids and Meters/Smart Grids/reference_architecture_smartgrids.pdf](https://www.cencenelec.eu/media/CEN-CENELEC/AreasOfWork/CEN-CENELEC_Topics/Smart%20Grids%20and%20Meters/Smart%20Grids/reference_architecture_smartgrids.pdf).
- [29] H. Schwarz, X. Cai, Integration of renewable energies, flexible loads and storages into the German power grid: Actual situation in German change of power system, Front.

- Energy. 11 (2017). <https://doi.org/10.1007/s11708-017-0470-x>.
- [30] A. Alhamwi, W. Medjroubi, T. Vogt, C. Agert, Modelling urban energy requirements using open source data and models, *Appl. Energy*. 231 (2018). <https://doi.org/10.1016/j.apenergy.2018.09.164>.
- [31] W. Medjroubi, U.P. Müller, M. Scharf, C. Matke, D. Kleinhans, Open Data in Power Grid Modelling: New Approaches Towards Transparent Grid Models, *Energy Reports*. 3 (2017) 14–21. <https://doi.org/10.1016/j.egy.2016.12.001>.
- [32] W. Biener, K.R. Garcia Rosas, Grid reduction for energy system analysis, *Electr. Power Syst. Res.* 185 (2020). <https://doi.org/10.1016/j.epsr.2020.106349>.
- [33] A. Alhamwi, W. Medjroubi, T. Vogt, C. Agert, Development of a GIS-based platform for the allocation and optimisation of distributed storage in urban energy systems, *Appl. Energy*. 251 (2019). <https://doi.org/10.1016/j.apenergy.2019.113360>.
- [34] Q. Wu, J. Zheng, Z. Jing, X. Zhou, Optimal Operation of Large-Scale Integrated Energy Systems, in: 2019. https://doi.org/10.1007/978-981-13-6943-8_6.
- [35] S. Amanpour, D. Huck, M. Kuprat, H. Schwarz, Integrated energy in Germany—A critical look at the development and state of integrated energies in Germany, *Front. Energy*. 12 (2018). <https://doi.org/10.1007/s11708-018-0570-2>.
- [36] J. Albouys-Perrois, N. Sabouret, Y. Haradji, M. Schumann, B. Charrier, Q. Reynaud, F. Sempé, C. Inard, Multi-agent simulation of collective self-consumption: Impacts of storage systems and large-scale energy exchanges, *Energy Build.* 254 (2022). <https://doi.org/10.1016/j.enbuild.2021.111543>.
- [37] K. Wu, H. Zhou, A multi-agent-based energy-coordination control system for grid-connected large-scale wind-photovoltaic energy storage power-generation units, *Sol. Energy*. 107 (2014). <https://doi.org/10.1016/j.solener.2014.05.012>.
- [38] M. Azeroual, T. Lamhamdi, H. El Moussaoui, H. El Markhi, Intelligent energy management system of a smart microgrid using multiagent systems, *Arch. Electr. Eng.* 69 (2020). <https://doi.org/10.24425/ae.2020.131756>.
- [39] T.H. Ruggles, K. Caldeira, Wind and solar generation may reduce the inter-annual variability of peak residual load in certain electricity systems, *Appl. Energy*. 305 (2022). <https://doi.org/10.1016/j.apenergy.2021.117773>.
- [40] OpenStreetMap Wiki, DE:Grenze, (n.d.).

- <https://wiki.openstreetmap.org/wiki/DE:Grenze> (accessed October 20, 2021).
- [41] Deutscher Wetterdienst, CDC - Climate Data Center, (n.d.). <https://cdc.dwd.de/portal/> (accessed October 23, 2022).
- [42] Deutscher Wetterdienst, Stationsliste der 78 Messstationen (nach Stationsname sortiert), (n.d.).
<https://www.dwd.de/DE/leistungen/klimadatendeutschland/stationsuebersicht.html> (accessed October 23, 2022).
- [43] Energiewirtschaftsgesetz (EnWG), § 23c Veröffentlichungspflichten der Netzbetreiber, (n.d.). https://www.gesetze-im-internet.de/enwg_2005/_23c.html (accessed October 24, 2022).
- [44] E.DIS Netz GmbH, Veröffentlichungspflichten/ Mitteilungspflichten Strom, (n.d.).
<https://www.e-dis-netz.de/de/edis-netz/veroeffentlichungspflichten/strom.html> (accessed October 24, 2022).
- [45] Bundesnetzagentur (BNetzA), Market Master Data Register, (n.d.).
<https://www.marktstammdatenregister.de/MaStR/> (accessed May 16, 2022).
- [46] H. Su, L. Chi, E. Zio, Z. Li, L. Fan, Z. Yang, Z. Liu, J. Zhang, An integrated, systematic data-driven supply-demand side management method for smart integrated energy systems, *Energy*. 235 (2021). <https://doi.org/10.1016/j.energy.2021.121416>.
- [47] J. Ren, Z. Yu, G. Gao, G. Yu, J. Yu, A CNN-LSTM-LightGBM based short-term wind power prediction method based on attention mechanism, *Energy Reports*. 8 (2022).
<https://doi.org/10.1016/j.egy.2022.02.206>.
- [48] M. Pierro, D. Gentili, F.R. Liolli, C. Cornaro, D. Moser, A. Betti, M. Moschella, E. Collino, D. Ronzio, D. van der Meer, Progress in regional PV power forecasting: A sensitivity analysis on the Italian case study, *Renew. Energy*. 189 (2022) 983–996.
<https://doi.org/10.1016/j.renene.2022.03.041>.
- [49] D. Song, Y. Yang, S. Zheng, X. Deng, J. Yang, M. Su, W. Tang, X. Yang, L. Huang, Y.H. Joo, New perspectives on maximum wind energy extraction of variable-speed wind turbines using previewed wind speeds, *Energy Convers. Manag.* 206 (2020).
<https://doi.org/10.1016/j.enconman.2020.112496>.
- [50] Y. Zhang, P. Hao, H. Lu, J. Ma, M. Yang, Modelling and estimating performance for PV module under varying operating conditions independent of reference condition,

- Appl. Energy. 310 (2022). <https://doi.org/10.1016/j.apenergy.2022.118527>.
- [51] E. Taslimi-Renani, M. Modiri-Delshad, M.F.M. Elias, N.A. Rahim, Development of an enhanced parametric model for wind turbine power curve, *Appl. Energy*. 177 (2016). <https://doi.org/10.1016/j.apenergy.2016.05.124>.
- [52] M. Pierro, M. De Felice, E. Maggioni, D. Moser, A. Perotto, F. Spada, C. Cornaro, Data-driven upscaling methods for regional photovoltaic power estimation and forecast using satellite and numerical weather prediction data, *Sol. Energy*. 158 (2017). <https://doi.org/10.1016/j.solener.2017.09.068>.
- [53] Y.M. Saint-Drenan, G.H. Good, M. Braun, T. Freisinger, Analysis of the uncertainty in the estimates of regional PV power generation evaluated with the upscaling method, *Sol. Energy*. 135 (2016). <https://doi.org/10.1016/j.solener.2016.05.052>.
- [54] U. Focken, M. Lange, K. Mönnich, H.P. Waldl, H.G. Beyer, A. Luig, Short-term prediction of the aggregated power output of wind farms - A statistical analysis of the reduction of the prediction error by spatial smoothing effects, *J. Wind Eng. Ind. Aerodyn.* 90 (2002). [https://doi.org/10.1016/S0167-6105\(01\)00222-7](https://doi.org/10.1016/S0167-6105(01)00222-7).
- [55] W. Hu, Z. Liu, J. Tan, Thermodynamic Analysis of Wind Energy Systems, in: *Wind Sol. Hybrid Renew. Energy Syst.*, 2020. <https://doi.org/10.5772/intechopen.85067>.
- [56] M. Mohanpurkar, R.G. Ramakumar, Probability density functions for power output of Wind Electric Conversion Systems, in: *IEEE PES Gen. Meet. PES 2010*, 2010. <https://doi.org/10.1109/PES.2010.5590119>.
- [57] Z. Xiao, Q. Zhao, X. Yang, A.F. Zhu, A power performance online assessment method of a wind turbine based on the probabilistic area metric, *Appl. Sci.* 10 (2020). <https://doi.org/10.3390/app10093268>.
- [58] R.H. Barnes, E. V. Morozov, K. Shankar, Improved methodology for design of low wind speed specific wind turbine blades, *Compos. Struct.* 119 (2014). <https://doi.org/10.1016/j.compstruct.2014.09.034>.
- [59] D. Kang, K. Ko, J. Huh, Determination of extreme wind values using the Gumbel distribution, *Energy*. 86 (2015). <https://doi.org/10.1016/j.energy.2015.03.126>.
- [60] H. Shi, Z. Dong, N. Xiao, Q. Huang, Wind Speed Distributions Used in Wind Energy Assessment: A Review, *Front. Energy Res.* 9 (2021). <https://doi.org/10.3389/fenrg.2021.769920>.

- [61] N. Ayuketang Arreyndip, E. Joseph, Generalized Extreme Value Distribution Models for the Assessment of Seasonal Wind Energy Potential of Debuncha, Cameroon, *J. Renew. Energy*. 2016 (2016). <https://doi.org/10.1155/2016/9357812>.
- [62] K. Knorr, Modellierung von raum-zeitlichen Eigenschaften der Windenergieeinspeisung für wetterdatenbasierte Windleistungssimulationen, Universität Kassel, 2017.
- [63] G. Reich, M. Reppich, Nutzung der Windenergie, in: G. Reich, M. Reppich (Eds.), *Regen. Energietechnik Überblick Über Ausgewählte Technol. Zur Nachhalt. Energieversorgung*, Springer Fachmedien Wiesbaden, Wiesbaden, 2018: pp. 151–185. https://doi.org/10.1007/978-3-658-20608-6_4.
- [64] R. Bründlinger, D. Christ, H. Fechner, M. Kaltschmitt, J. Müller, G. Peharz, D. Schulz, L. Sens, Photovoltaische Stromerzeugung, in: *Erneuerbare Energien*, 2020. https://doi.org/10.1007/978-3-662-61190-6_5.
- [65] Georg Wirth, Modellierung der Netzeinflüsse von Photovoltaikanlagen unter Verwendung meteorologischer Parameter, Carl von Ossietzky Universität Oldenburg, 2014.
- [66] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, 2016. <https://www.deeplearningbook.org/> (accessed May 16, 2022).
- [67] J. Feng, X. He, Q. Teng, C. Ren, H. Chen, Y. Li, Reconstruction of porous media from extremely limited information using conditional generative adversarial networks, *Phys. Rev. E*. 100 (2019). <https://doi.org/10.1103/PhysRevE.100.033308>.
- [68] B. Zazoum, Solar photovoltaic power prediction using different machine learning methods, *Energy Reports*. 8 (2022). <https://doi.org/10.1016/j.egy.2021.11.183>.
- [69] D.A. Otchere, T.O. Arbi Ganat, R. Gholami, S. Ridha, Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models, *J. Pet. Sci. Eng.* 200 (2021). <https://doi.org/10.1016/j.petrol.2020.108182>.
- [70] N. Tutkun, Minimization of operational cost for an off-grid renewable hybrid system to generate electricity in residential buildings through the SVM and the BCGA methods, *Energy Build.* 76 (2014). <https://doi.org/10.1016/j.enbuild.2014.03.003>.
- [71] V. Suresh, P. Janik, J. Rezmer, Z. Leonowicz, Forecasting solar PV output using

- convolutional neural networks with a sliding window algorithm, *Energies*. 13 (2020). <https://doi.org/10.3390/en13030723>.
- [72] M. Abdel-Nasser, K. Mahmoud, Accurate photovoltaic power forecasting models using deep LSTM-RNN, *Neural Comput. Appl.* 31 (2019). <https://doi.org/10.1007/s00521-017-3225-z>.
- [73] A. Agga, A. Abbou, M. Labbadi, Y. El Houm, I.H. Ou Ali, CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production, *Electr. Power Syst. Res.* 208 (2022). <https://doi.org/10.1016/j.epsr.2022.107908>.
- [74] X. Chen, X. Zhang, M. Dong, L. Huang, Y. Guo, S. He, Deep Learning-Based Prediction of Wind Power for Multi-turbines in a Wind Farm, *Front. Energy Res.* 9 (2021). <https://doi.org/10.3389/fenrg.2021.723775>.
- [75] S. Hochreiter, J. Schmidhuber, Long Short-Term Memory, *Neural Comput.* 9 (1997). <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [76] F. Shahid, A. Zameer, M. Muneeb, A novel genetic LSTM model for wind power forecast, *Energy*. 223 (2021). <https://doi.org/10.1016/j.energy.2021.120069>.
- [77] Y. Chen, Y. Wang, Z. Dong, J. Su, Z. Han, D. Zhou, Y. Zhao, Y. Bao, 2-D regional short-term wind speed forecast based on CNN-LSTM deep learning model, *Energy Convers. Manag.* 244 (2021). <https://doi.org/10.1016/j.enconman.2021.114451>.
- [78] QGIS Tutorials, Nearest Neighbor Analysis, (n.d.). http://www.qgistutorials.com/en/docs/nearest_neighbor_analysis.html (accessed May 16, 2022).
- [79] B. Jing, Z. Qian, H. Zareipour, Y. Pei, A. Wang, Wind turbine power curve modelling with logistic functions based on quantile regression, *Appl. Sci.* 11 (2021). <https://doi.org/10.3390/app11073048>.
- [80] Y. Wang, R. Zou, F. Liu, L. Zhang, Q. Liu, A review of wind speed and wind power forecasting with deep neural networks, *Appl. Energy*. 304 (2021). <https://doi.org/10.1016/j.apenergy.2021.117766>.
- [81] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, others, Tensorflow: A system for large-scale machine learning, in: *12th USENIX Conf. Oper. Syst. Des. Implement.*, 2016: pp. 265–283.

- [82] Y. Ren, X. Yao, D. Liu, R. Qiao, L. Zhang, K. Zhang, K. Jin, H. Li, Y. Ran, F. Li, Optimal design of hydro-wind-PV multi-energy complementary systems considering smooth power output, *Sustain. Energy Technol. Assessments*. 50 (2022). <https://doi.org/10.1016/j.seta.2021.101832>.
- [83] R. Benadli, M. Bjaoui, B. Khiari, A. Sellami, Sliding Mode Control of Hybrid Renewable Energy System Operating in Grid Connected and Stand-Alone Mode, *Power Electron. Drives*. 6 (2021). <https://doi.org/10.2478/pead-2021-0009>.
- [84] T.T. Teo, T. Logenthiran, W.L. Woo, K. Abidi, Fuzzy logic control of energy storage system in microgrid operation, in: *IEEE PES Innov. Smart Grid Technol. Conf. Eur.*, 2016. <https://doi.org/10.1109/ISGT-Asia.2016.7796362>.
- [85] D. El Bourakadi, A. Yahyaouy, J. Boumhidi, Multi-Agent System Based on the Extreme Learning Machine and Fuzzy Control for Intelligent Energy Management in Microgrid, *J. Intell. Syst.* 29 (2020). <https://doi.org/10.1515/jisys-2018-0125>.
- [86] A. Mohamed, S.S. Refaat, H. Abu-Rub, A Review on Big Data Management and Decision-Making in Smart Grid, *Power Electron. Drives*. 4 (2019). <https://doi.org/10.2478/pead-2019-0011>.
- [87] M.S. Alam, S.A. Arefifar, Energy Management in Power Distribution Systems: Review, Classification, Limitations and Challenges, *IEEE Access*. 7 (2019). <https://doi.org/10.1109/ACCESS.2019.2927303>.
- [88] C.M. Colson, M.H. Nehrir, Algorithms for distributed decision-making for multi-agent microgrid power management, in: *IEEE Power Energy Soc. Gen. Meet.*, 2011. <https://doi.org/10.1109/PES.2011.6039764>.
- [89] J. Li, W. Wei, J. Xiang, A simple sizing algorithm for stand-alone PV/Wind/Battery hybrid microgrids, *Energies*. 5 (2012). <https://doi.org/10.3390/en5125307>.
- [90] L.S. Vargas, G. Bustos-Turu, F. Larrain, Wind Power Curtailment and Energy Storage in Transmission Congestion Management Considering Power Plants Ramp Rates, *IEEE Trans. Power Syst.* 30 (2015). <https://doi.org/10.1109/TPWRS.2014.2362922>.
- [91] M.R. Hossain, A. Maung Than Oo, A.B.M.S. Ali, A hybrid machine learning using Mamdani type fuzzy inference system (FIS) for solar power prediction, *Ann. Fuzzy Sets, Fuzzy Log. Fuzzy Syst.* 2 (2013).
- [92] skfuzzy 0.2 docs — skfuzzy v0.2 docs, (n.d.). <https://pythonhosted.org/scikit->

- fuzzy/index.html (accessed January 9, 2022).
- [93] C. Kost, S. Shammugam, V. Fluri, D. Peper, A.D. Memar, T. Schlegl, Fraunhofer ISE: Levelized Cost of Electricity - Renewable Energy Technologies, June 2021, 2021. <https://www.ise.fraunhofer.de/en/publications/studies/cost-of-electricity.html>.
- [94] M. Kuprat, M. Bendig, K. Pfeiffer, Possible role of power-to-heat and power-to-gas as flexible loads in German medium voltage networks, *Front. Energy*. 11 (2017) 135–145. <https://doi.org/10.1007/s11708-017-0472-8>.
- [95] D. Harrison-Atlas, C. Murphy, A. Schleifer, N. Grue, Temporal complementarity and value of wind-PV hybrid systems across the United States, *Renew. Energy*. 201 (2022) 111–123. <https://doi.org/10.1016/J.RENENE.2022.10.060>.
- [96] S. Venkataraman, C. Ziesler, P. Johnson, S. Van Kempen, Integrated Wind, Solar, and Energy Storage: Designing Plants with a Better Generation Profile and Lower Overall Cost, *IEEE Power Energy Mag.* 16 (2018). <https://doi.org/10.1109/MPE.2018.2793478>.
- [97] S.Z.M. Golroodbari, D.F. Vaartjes, J.B.L. Meit, A.P. van Hoeken, M. Eberveld, H. Jonker, W.G.J.H.M. van Sark, Pooling the cable: A techno-economic feasibility study of integrating offshore floating photovoltaic solar technology within an offshore wind park, *Sol. Energy*. 219 (2021). <https://doi.org/10.1016/j.solener.2020.12.062>.
- [98] S. Sterl, S. Liersch, H. Koch, N.P.M.V. Lipzig, W. Thiery, A new approach for assessing synergies of solar and wind power: Implications for West Africa, *Environ. Res. Lett.* 13 (2018). <https://doi.org/10.1088/1748-9326/aad8f6>.
- [99] The Wind Power Wind Energy Market Intelligence, Manufacturers and turbines, (n.d.). https://www.thewindpower.net/turbines_manufacturers_en.php (accessed November 26, 2022).