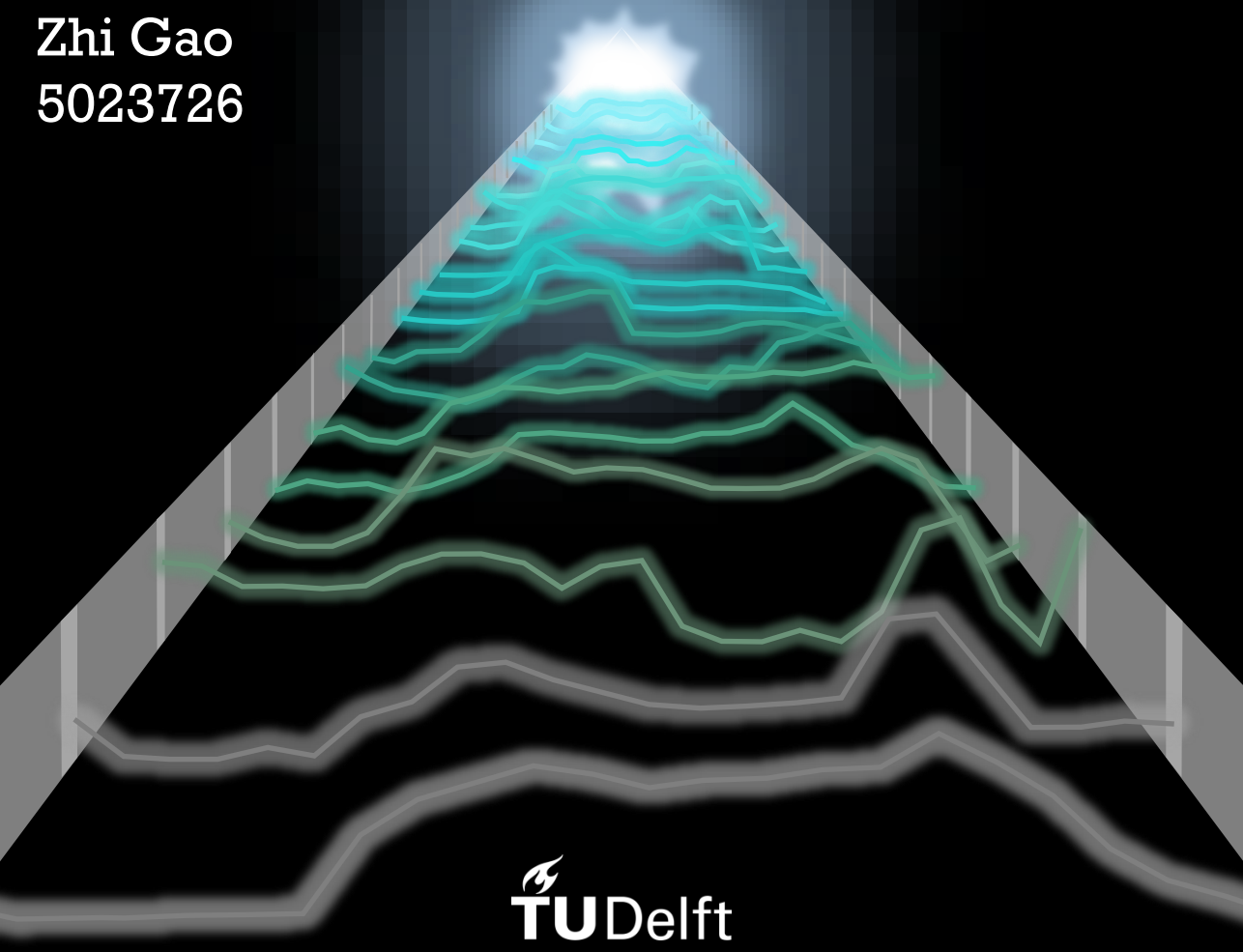
Back to the Future

Projecting Dutch Electricity Day-ahead Market Price Series under 2050 Energy Scenarios

Zhi Gao
5023726



TU Delft

BACK TO THE FUTURE

PROJECTING DUTCH DAILY DAY-AHEAD MARKET PRICE SERIES UNDER 2050 ENERGY SCENARIOS

Master Thesis

to obtain the degree of Master of Science in Sustainable Energy Technology

by

Zhi GAO

Faculty of Electrical Engineering, Mathematics and Computer Science,
Delft University of Technology, Delft, Netherlands.

Student number:	5023726	
Email:	gaozcme@gmail.com	
Project duration:	Q2 2020/21 - Q3 2021/2022	
Thesis committee:	Ir. Ties van der Heijden	Daily supervisor
	Dr.ir. Edo Abraham	Supervisor
	Prof.dr. Peter Palensky	Supervisor

An electronic version of this thesis is available at http://repository.tudelft.nl/.

ACKNOWLEDGEMENT

Looking back from the end of my thesis project and master study at TU Delft, I see myself shaping up as a leaner, a thinker, and most importantly as a growing contributor in tackling challenges among the human society. Excitements and happiness accompanied when I learned about electricity prices and neural networks. This project gave me the perfect opportunity to look for a grand landscape behind technical profundities.

I could not have done it without my daily supervisor and my friend, Ir. Ties van der Heijden, who generously welcomed me to the project, introduced me the field of adversarial learning, tutored me in coding and writing with the finest patience, and treated me with the kindest hospitality when meeting in person. Many many credits should also go to my supervisor Dr.ir. Edo Abraham for enlightening me with his critical insights and persistence of scientific robustness. Last but not least, my gratitude shall be expressed towards Prof.dr. Peter Palensky for encouraging me to always aim high and push forward. Thanks again to my supervisory team for working with me. You will be my exemplars in the upcoming research career.

Furthermore, I want to thank TU Delft and the SET program for offering the unforgettable educational experience. I really enjoyed knowing the local culture, hanging out with the community, getting help from the career center and exploring subjects freely in the campus. Going to Delft has been one of the best decisions I made in my life.

Finally, I would like to acknowledge my appreciation to my beloved parents who offer me the privilege to live a good life and reach for my dreams in Delft. I am especially grateful for the support from my grandfather and my dearest friend Taimuzi Fu during difficult times.

Zhi Gao
Guangzhou, January 2022

ABSTRACT

Under the increasing electrification of end uses in the energy transition towards more renewable integration, the electricity price keeps gaining importance on every scale from individual well-being to the competitiveness of an economy. Though scarce in the scientific literature, Long-Term Electricity Price Projection (LEPP) has great potentials in decision-making and planning, as well as complementing the long-term energy scenarios. This study takes features from the Dutch, Spanish and Danish data in five years (2015-2019) to train deep neural networks in the conditional Wasserstein Generative Adversarial Nets with Gradient Penalty (cWGAN-GP) framework, in order to project Day-Ahead Market (DAM) price series under Dutch 2050 energy scenarios.

The LEPP to 2050 is made possible by normalising the selected markets. As a result, the conditions unprecedented in Dutch data are covered in the normalised and combined data set. Generally, under scenarios with high proportions of hydrogen power in the energy portfolio, the cWGAN-GP model projects that DAM price series would have slightly lower mean and daily standard deviation than the 2019 level. Whereas much lower mean and daily standard deviation are projected when natural gas is still the fuel of the most frequent final generating technology. To explore the possible application of the projector model, the German DAM prices series in 2019 have been projected and evaluated, and the projections under Dutch 2050 energy scenarios have been used in calculating the generic profit potential of energy storage.

Five findings can be summarised from the main results. Firstly, from a literature survey and importance analyses, seven features are shown relevant to the DAM price in the combined data set, namely month of the year, day of the week, total hourly load forecast, national daily mean temperature, fuel cost of the most frequent final generating technology, hourly renewable power generation forecast and total installed renewable power capacity. Secondly, it has been found that two of the four proposed market state normalisation solutions, the Renewable Scarcity Factor (RSF) and the Renewable-Load Ratio (RLR) help the cWGAN-GP model strike a balance between price value distribution and hourly inter-dependencies. Thirdly, in this LEPP study, the cWGAN-GP model performs better than the Conditional Variational Auto-Encoder (CVAE) and multivariate Gaussian distribution (mGaus) models. Compared with the two alternatives, the cWGAN-GP model produces samples in better quality while remaining sensitive to temporal conditions. Fourthly, projections by the cWGAN-GP model are more realistic than those made by the Energy Transition Model (ETM), with price values varying continuously in smooth boundaries. Finally, the fuel cost of the most frequent final generating technology is found critical to LEPP. The annual mean and daily standard deviation of the DAM price series are expected to rise significantly when natural gas is mostly replaced by hydrogen power in the national energy portfolio.

The study is limited in that no information on energy storage is included in the data set, as it is also missing in Dutch 2050 energy scenarios. And the pan-European market

integration considered in some Agent-Based Modelling (ABM) studies is neglected in this study due to a lack of valuable data. Moreover, supplementary efforts can be made on further differentiating the RM, NM and IM scenarios, because projections to these scenarios are similar in this study. Last but not least, it is possible that the cWGAN-GP model trained with DAM prices and conditions prior to the COVID-19 pandemic can be taken advantage of in finding impacts on the energy system from the pandemic.

CONTENTS

Acknowledgement iii

Abstract v

1 Introduction 1
- 1.1 Electricity Prices . . . 1
- 1.2 Long-term Electricity Price Projection . . . 3
- 1.3 Dutch 2050 Energy Scenarios . . . 5
- 1.4 This Study . . . 8

2 Feature Selection 11
- 2.1 Day-Ahead Market (DAM) Price Series . . . 11
- 2.2 Candidate Variables and Pre-processing . . . 12
- 2.3 Feature Selection Procedure . . . 14
- 2.4 Exploratory Data Analysis . . . 15
- 2.5 Importance Analyses . . . 18
- 2.6 Market State Normalisation Methods . . . 21
- 2.7 Causality Validation . . . 23
- 2.8 Selected Feature Variables Summary . . . 24
- 2.9 Effectiveness of the Market State Indicators . . . 25

3 Conditional Wasserstein Generative Adversarial Nets with Gradient Penalty (cWGAN-GP) 27
- 3.1 History of Neural Networks in a Nutshell . . . 27
- 3.2 The Neurons . . . 28
- 3.3 Gradient Descent and Back-Propagation . . . 30
- 3.4 Weight Optimization Approaches . . . 33
- 3.5 Conditional Generative Adversarial Nets (cGAN) . . . 34
- 3.6 Improvements of the GAN (cGAN) Framework . . . 35
- 3.7 Other Potential Methods . . . 37
- 3.8 Test Case on Reproducing Sinusoidal Curves . . . 38

4 Model Training, Evaluation and Projection 43
- 4.1 Evaluation Metrics . . . 43
- 4.2 The Training Procedure . . . 46
- 4.3 Conditions for Projection under Dutch 2050 Energy Scenarios . . . 49

5 Results and Discussion 51
- 5.1 Market State Normalisation . . . 51
- 5.2 Selected Price Projection Methods . . . 53
- 5.3 Projection under Dutch 2050 Energy Scenarios . . . 55

5.4 Projection under German 2019 Conditions 60
5.5 Generic Profit Potential of Energy Storage. 62
5.6 Discussion from Price Projection Results 65

6 Conclusion **69**

References **73**

1

INTRODUCTION

1.1. ELECTRICITY PRICES

Since the second industrial revolution, electricity has been playing an integral role in human societies, from ubiquitous applications in industries to enabling a modern lifestyle [1]. The share of electricity in global total final energy consumption has risen evidently from 9.4% in 1973 to 19.3% in 2018, mostly generated from coal and oil [2]. With the ongoing electrification of end uses, e.g. in the soaring market of electric vehicles in the transport sector, electricity is gradually becoming the most essential form of energy use [3, 4].

The electrification process is under reinforcement by the energy transition in which societies are shifting from fossil fuels to renewable energy sources, with the main motive to reduce emissions of greenhouse gases. As the emissions so far has resulted in a rise of global temperature unseen in many millennia [5], societies are urgently calling for actions. The European Union has announced an ambition to reduce its greenhouse gas emissions by more than 55% to the levels in 1990 while eventually reaching net zero by 2050 [6]. Furthermore, electricity is the direct product of most burgeoning renewable energy technologies. Specifically, solar, wind and hydro power, whose energy generation takes 90% of the total renewables, produce electricity, though some other renewable energy technologies contribute in heating [7].

With the growing dependency on electricity as an energy carrier, electricity prices are inevitably drawing attentions in the socio-economical discourse. On one hand, electricity prices have been frequently linked with industrial competitiveness. For instance, Serrano González and Álvarez Alonso [8] emphasised industries of all scales in Spain spend more than 50% of the cost in energy, and the low grid inter-connectivity to other EU countries threatens the industrial competitiveness of the country. Another study revealed the inter-relation between the shrinkage of the car manufacturing industry in Ontario, Canada with increasing electricity prices there [9]. A numerical research on the productivity of sectors in China suggested that an increase in electricity prices does not affect the industrial productivity directly, instead, it promotes the companies to improve

1

their energy efficiency, which can incur extra needs for capital investment [10]. On the other hand, electricity prices signal the adequacy of electricity supply. An example is the 2021 Texas electricity crisis, where a massive shortage occurred in the state due to a surge in demand, causing 4.5 million Texans suffer in the winter storm. During this period, the electricity price for most parts of Texas reached the price ceiling, 9000 $/MWh, about 100 times the usual level [11].

In the Netherlands, the wholesale electricity price is determined in the liberalised electricity markets. The final consumer-end prices depend on many factors, such as taxes, grid fees, the supplier, among others [12]. In the electricity supply chain, as shown in Figure 1.1, the generated electricity flows into the transmission system after it has been sold in the market. The trading is done before the generation of electricity, setting the wholesale market price. And the Program Responsible Parties (PRP) are the entities eligible to buy or sell electricity. To become a PRP, an entity must have at least one connection to the grid and report daily on both the electricity load applied on the transmission line and its net electricity demand to Tennet, the Dutch Transmission System Operator (TSO).

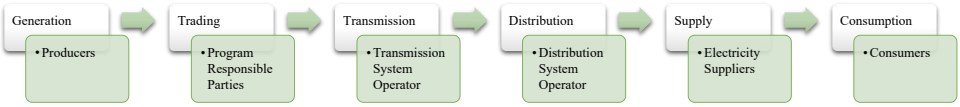


Figure 1.1: Supply chain of electricity and responsible parties in the Netherlands, adapted from [13].

The PRPs trade on the power futures and spot markets [14]. Main markets and their bidding time requirements are illustrated in Figure 1.2. The futures market closes two days before the delivery day. One day ahead of the delivery day, the Day-Ahead Market (DAM) opens, where bids are submitted for every hour in the delivery day at the same time. On the delivery day, PRPs can adjust their production or consumption on the Intra-Day Market (IDM), which is cleared down to 15 minutes before delivery. There is also an imbalance market where PRPs with a capacity over 60MW are obligated to bid for capacity services that balance the supply and demand. In DAM, the price is set by a double-blind auction where buy-orders are sorted in a descending manner, and sell-orders are ranked in an ascending way. The Market Clearing Price (MCP) is set at the cross of the two curves. Hence, all the buy-orders higher and sell-orders lower than the price are activated. While in IDM, a first come-first served principle is executed that the highest sell prices and lowest sell prices are only available in the beginning.

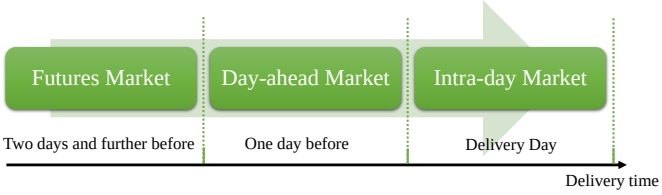


Figure 1.2: Dutch electricity markets and their opening times, adapted from [15].

It is expected that spot market prices will be influenced by the energy transition. A study on the Danish energy system and the Nord Pool spot market discovered that the addition of offshore & onshore wind and solar energy instantly lowers the spot market prices [16]. The main cause is that the cheaper renewable energy bids displace others in the merit order, meanwhile the market clearing units, thermal plants in the paper, rather sell cheaper to avoid being shut down under such displacement. Furthermore, depending on the generation portfolio, penetration of these renewable energy technologies can lead to contradictory changes in price volatility. Rintamäki et al. [17], using Auto-Regressive (AR) models, have found that in the German market, price volatility is amplified only in the off-peak hours with the renewable energy penetration. Whereas in Denmark, the electricity output from wind is enough to flatten out peak-hour prices, resulting in suppressed price volatility.

The lower bids from renewable energy are mainly due to the fact that, unlike conventional power plants, renewable energy generators converting energy from the nature do not consume any fuel. Consequently, their marginal cost of generation is zero if operational costs are neglected. By adopting the optimal strategy of bidding maximum capacity at a price of the marginal cost, electricity generation companies are therefore willing to sell for much lower prices [18].

At the same time, spot market prices play a role in the energy transition. A design purpose behind such market clearing principles is to create investment incentives for developing energy technologies with lower marginal costs, and to phase out less efficient generators [19]. If electricity spot market prices keep dropping, income of energy generators, including the upcoming renewable energy projects, would decrease for dampened cash flow [16]. The possible income decrease could burden the sustainable development, especially renewable energy projects as they have been relying on subsidies. Till 2020, renewable energy projects of 23.7 GW had been subsidised in Netherlands under the SDE+ scheme, whereas only 2.3 GW, mainly offshore wind projects, are ratified without subsidy [20].

1.2. LONG-TERM ELECTRICITY PRICE PROJECTION

Stakeholders in the energy transition have motivations to forecast electricity prices. To PRPs, having information on future prices is beneficial for both the bidding operations and investment decision-making [21]. And policymakers model electricity prices to examine the impacts of different policies. For instance, in [22], probabilistic generation capacity models are created to examine the effectiveness of national capacity markets. And such effectiveness is evaluated by the social cost in paying the import electricity price. Forecasting electricity prices might even be more fundamental in the future. To give an example, with prior information on the electricity price, energy consumers are becoming able to conduct demand response or even switching the type of energy consumption to minimise their total energy cost [23].

The majority of research in electricity price forecasting (EPF) are in the timescale less than a year, despite the interest in this field rising for more than two decades [24]. Methods applied in EPF can be categorised into five, namely multi-agent models, structural models, price-focused models, statistical models and machine learning (ML) models [25]. The multi-agent models often consist of multiple programmed agents, each rep-

resenting a part in the electricity market. And the price is derived from interactions of the agents. The structural models emphasise on the relationship between the price and driver factors like load, temperature and energy prices. Price-focused models, inspired by financial approaches, attempt to pinpoint the time derivatives of the price with limited model complexity. A stochastic error term can be added in the price-focused forecasts. In statistical models, there are mathematical treatments of previous prices and external factors, which aim at giving accurate price forecast. In the day-ahead forecast, the most popular area of EPF, it has been shown that the ML models outperform other methods. And among the ML models, Deep Learning (DL) models have produced further improved results [26]. The DL models are typically based on neural networks with more than one hidden layer. Therefore, training these models often comes with higher computational costs than simpler ML models. For day-ahead EPF, standard features used are hourly load estimate, renewable energy forecasts, fuel prices, carbon prices and available generation capacities [27].

Potentials lie in EPF for a longer term. Taking studies in electricity load forecast as a reference, researches have been proposing methodologies for forecasting electricity load 10-50 years ahead [28]. In these methodologies, macro-economic factors, including GDP, population growth, fuel prices, are often included. Moreover, structural changes in electricity end-use, such as the rising electricity consumption by electric vehicles, are also considered. And the end users' characteristics can be fundamentally altered with integration of energy storage, not only in flattening and shifting the load regarding the electricity price, but also in reducing the connection requirements. A study separately modelling local and national grid has found that with sufficient local energy storage, a community with enough solar power generation could be independent from the national grid, hence indifferent to the electricity price [29]. In these long-term load projections, most adopted methods are statistical or agent-based [30]. That said, applications of ML methods in long-term electricity load forecasting are emerging with the advancements in ML regression models [31].

In comparison, literature regarding EPF for commensurate timescales is scarce, with only several that can be noticed in grey literature. The model of the electricity market developed by an energy consultancy firm CE Delft [32] extrapolates the DAM price and its characteristics in 2023 and in 2030. The model is based on the historical performance of power generators and exogenous fuel prices. Another report, the EU Energy Outlook 2050 [33], predicted annually that the European electricity price would have a steady increase til 2035, with enlarged seasonal price volatility. In spite of the clear results in these reports, the methods and their effectiveness in long-term electricity price projection are not academically examined, to the best knowledge of the author.

Nevertheless, insights into the electricity price on a similar timescale can be derived from studies on the energy transition. For instance, Leuthold et al. [34] created a large network of nodes and connections across mid-west Europe, with each node representing a region. An optimisation procedure was conducted to study the effect of congestion management on price discrimination, where electricity prices in neighbouring countries were computed. Another widely implemented method in energy transition studies is the Agent-Based Modeling (ABM), which are actively taken for exploring the long-term relationships between the electricity price, energy policy and electricity system adequacy,

optionally with other economic factors [35]. ABM concept predicates that any system encompasses sub-level agents. To understand the functions of the system, each agent and the interactions between agents need to be examined [36]. Therefore, ABM-based studies usually take all major actors in the energy transition into account, thus providing more holistic behavioral analyses, including prospects on the electricity price [37].

Additionally, the Energy Transition Model (ETM), a structural model built for studying the energy transition, is noteworthy [38]. Developed by an energy consultancy firm Quintel, ETM takes specifications of the energy sector comprehensively, covering demand, supply, flexibility, emissions, costs and energy efficiencies. Using the detailed specifications and ETM's built-in functional principles, the results including hourly electricity prices are calculated for a given weather year. By adjusting the specifications, the user is able to observe the impacts on electricity prices under different scenarios.

Annual average electricity price projections have been studied for the Netherlands in ABM-based studies. In one study by the EMLab, under all of the policy scenarios, namely emission trading, carbon taxation and no intervention, the annual average electricity price has been projected to have a decline of roughly 20% in an S-shaped curve till 2060 [39]. In another study, Fraunholz et al. [35] from the PowerACE have found that the Dutch DAM electricity price would experience significant growth, more than doubled from 2020 to 2050. The primary difference between the two models is the organization of the agents. The EMLab model was designed with the focus on projecting consequences of various national policies. Therefore the agents were built for a single country. Whereas in the PowerACE study, the key interest was on analysing the constraints on market integration in the Europe. And the agents simulated the national markets equipped with different capabilities in price forecasting. Although the main tasks of these studies were not to project the electricity price, it can be seen that choosing different models, in other words, having different representations of the energy system, can lead to contradictory electricity price projections.

1.3. DUTCH 2050 ENERGY SCENARIOS

In the socio-technical aspect of the energy transition, various scenarios have been proposed to provide foundations for scientific research and policy assessments [40]. Generally, scenarios are formed in two ways. One is to set specific actions and incidents, based on observed behaviours of the socio-technical system. The other is the opposite, revealing what activities and incidents are required to facilitate pre-defined goals [41].

Energy scenarios put forward in the Netherlands are mostly based on emission reduction targets. As announced by the Dutch government, the goal is to have a reduction of 49% by 2030 and 95% by 2050, compared to 1990 levels [42]. Scenarios drawn so far are formulated with hypotheses on how much the goal is fulfilled or in which different ways the goal is realised.

It is mainly the consultancy firms and research institutes that propose energy scenarios in the Netherlands. Looking forward to 2030, CE Delft [43] sketched the future under different degrees of carbon emission cut and renewable generation. Similarly, TNO Energy [44] evaluated the technological innovations vital for improving sustainability in the next ten years. As for previewing the final 2050 targets, there are two scenario reports in Dutch, one compiled by the association of all the gas and electricity network operators

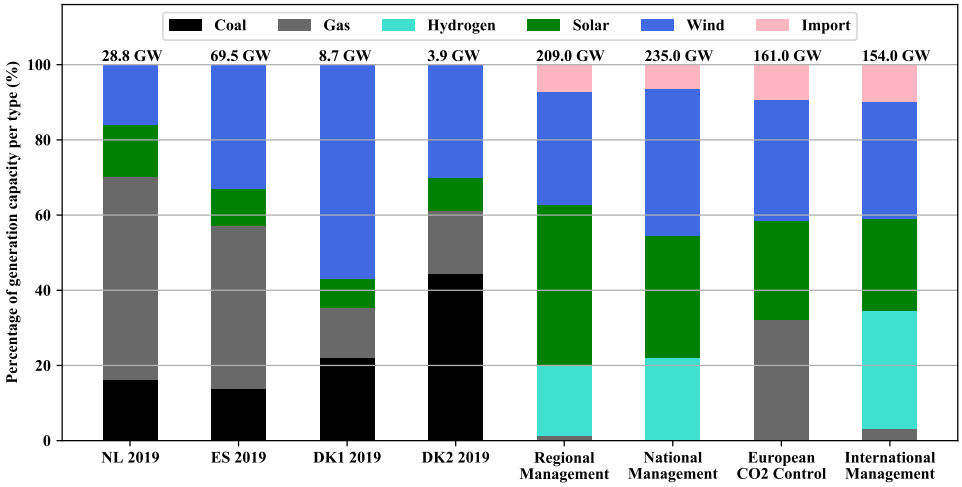
Netbeheer [45], which is based on the other jointly complied by two consultancy firms Berenschot and Kalavasta [46]. Both reports drew scenarios with regards to the level of management, namely regional, national, European and world-wide, but the former focuses on the energy infrastructure required in the scenarios. The one by Berenschot and Kalavasta was sent to the House of Representatives in the Dutch Parliament in 2020, as a material aiding in the decision making on energy infrastructure plans. Owing to its official acceptance, this report can be taken as a reference for conducting research on the energy transition in the Netherlands.

Table 1.1: Four Dutch climate neutral energy scenarios and assumptions [46].

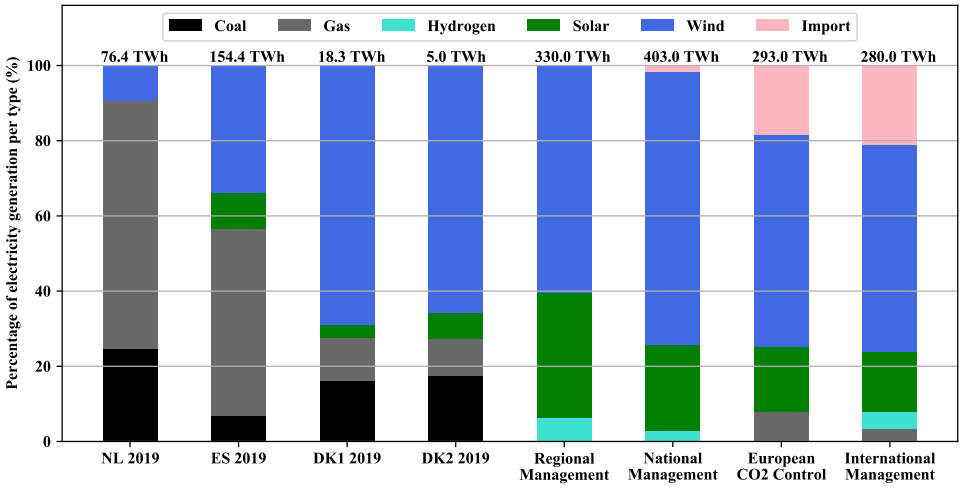
General assumptions	**Regional management**	**National management**	**European CO2 control**	**International management**
Who achieves CO2 target	NL	NL	Europe	The world
Measures level	Regional	National	Continental	Global
CO2 reduction	100%	100%	100%	100%
Policy installed	Unspecified	Unspecified	General CO2 tax	Unspecified
Dutch self-sufficiency	In principle	Very high	Unspecified	Unspecified
Import/export trading	Minimal import	Minimal import	EU duties and compensations	Stimulated globally
Energy intensive industry	Shrinkage	Remain current size	Growth	Growth
Society	Passionate citizens	Major national projects	Global hydrogen & biomass market	Global hydrogen & biomass market
Carbon capture & storage	Unspeficied	Unspeficied	Given a lot of space	Given space
Circularity for goods and food production	Spearheaded	Stressed	Unspecified	Unspecified

In the 2050 Dutch energy scenarios put forward by Berenschot and Kalavasta [46], there are four scenarios, namely Regional Management (RM), National Management (NM), European CO2 Control (EC) and International Management (IM), derived from different qualitative assumptions, as listed in Table 1.1. These assumptions are driven by the choice on the subject of responsibility, as well as the resulting scale of regulatory measures. For instance, Dutch energy self-sufficiency is required in RM and NM scenarios, as the national and regional governments are being accounted for. Whereas in EC and IM scenarios, such self-sufficiency is not specified, as it is allowed to import energy from other countries. These disparities in regulatory measures lead to different energy portfolios. As illustrated on the right four columns in Figure 1.3a and 1.3b, generation capacity and electricity generated under the four scenarios are discernible with

each other in both the total volumes and the proportions per type. As it is difficult to incorporate the qualitative assumption, the capacities and generations can be taken as features when conducting a numerical study. The left four columns in the figures shows the generation and capacity portfolio of the NL, ES, DK1 and DK2 markets in 2019. Data from the latter three markets will be included for this study, which will be explained in Section 1.4.



(a) Percentage of generation capacity per type.



(b) Percentage of electricity generation per type.

Figure 1.3: Generation capacity and electricity generation portfolios in the NL, ES, DK1 and DK2 bidding zones in 2019 and the Dutch 2050 scenarios [47, 46]. Combined heat & power, biomass, hydro power and nuclear energy are not included for their absence in the ENTSO-E data and the energy scenarios, respectively.

1

1.4. THIS STUDY

The target of this study is to project daily electricity prices with Dutch 2050 energy scenarios. Given the time gap of three decades, this study is distinguishable from the majority of EPF studies. To make it clearer, let us define a term, Long-term Electricity Price Projection (LEPP), to denote generating credible electricity price series more than ten years ahead, which is the case of this study.

The DAM price is chosen as the subject of this study, because most of the power transactions are done in DAM. Except for the bids submitted one day ahead, the contracts in the futures market are also realised in DAM. In 2020, the volume of energy exchanged in DAM is 28581.1 GWh, more than six times the volume of 4262.8 GWh of IDM [48]. Plus, the variation of DAM prices is confined in a range from -500 to 3000 €/MWh, whereas in IDM, the limits are -99,999.90 and 99,999.90 €/MWh [15]. Considering together with the difference in submission deadlines, it can be inferred that DAM price reflects the demand-supply relationship better, with fewer uncertainties (e.g. from abrupt changes in weather or unforeseen shortages) [49].

In this study, DAM prices from three other bidding zones, namely ES (in Spain), DK1 and DK2 (both in Denmark), are included to deal with the conditions unprecedented in the Dutch market. As can be noticed in Figure 1.3a and 1.3b, capacity and generation portfolios of the four 2050 scenarios are significantly different from the Dutch state in 2019. In fact, such high renewable penetration is unseen in historical Dutch data. Therefore, we select the bidding zones in Spain and Denmark to extend the data set. In the figures, the DK1 bidding zone seems much closer to the scenarios than NL, in both capacity and generation. And ES bidding zone is selected for possible references on a warmer climate and larger market volume, though changes in these aspects are not discussed in this study. The key issue is how we can incorporate data from multiple bidding zones. One potential solution is to normalise the markets by creating a state indicator to quantify states of the four markets together. Faced with this challenge, creation and evaluation of such indicator will be presented in this study.

Current popular methods capable of LEPP involve detailed modelling of the electrical system or parts of it, while data-based methods are scarcely applied. Including the emerging machine learning, these data-based methods would offer a new perspective and explores unattended potentials. In studies of long-term time series projection, machine learning methods have achieved fruitful results. For example, Herrera et al. [50] used Artificial Neural Networks (ANN) to predict prices of energy commodities five years ahead. Another study adopted multiple machine learning methods to produce long-term electricity load series [51]. Therefore, the machine learning methods are one of the focuses in this study. Among the machine learning methods, two generative methods, conditional Generative Adversarial Networks (cGAN) and Conditional Variational Auto-Encoders (CVAE) are particularly fit for the task. The generative models are able to produce stochastic samples with similar characteristics with the training data, which is desirable in LEPP. Principles and implementation of these generative models are going to be discussed in the upcoming chapters.

Subsequent to building the models, evaluation of DAM price projections is another study focus. As most published evaluation metrics on cGAN models are for image generation purposes [52]. In these tasks, there are categories of samples to learn and generate

from. Many of those metrics use probability estimates of presence of the category and generated samples. While in our case, there are plural dimensions of conditions depicting the market state, thus using such probability estimates is not viable. Alternatively, examining separately the distribution of price values and their hourly dependencies would be of greater value for LEPP. In answer to the shortage of suitable metrics, evaluation metrics novel for GAN model evaluation will be introduced in this study.

Through this study, the following Research Questions (RQ) will be answered.

Main RQ: To what extend can Dutch day-ahead market price series be projected under 2050 energy scenarios?

Sub-RQs:

- *What factors should be considered as features for price series projection?*
- *Which method is most suitable for this task, and how can it be effectively applied?*
- *Do data from other bidding zones help in projecting future Dutch DAM price series?*
- *What methods provide reliable validation of the projected price series?*
- *What would be valuable applications of the projector model?*

A structural overview of this thesis is as follows. Secondly, collection of DAM prices and candidate features, preliminary illustration of feature relevance with Exploratory Data Analysis (EDA) and feature importance rankings with Principal Component Analysis (PCA) and Random Forest (RF) will be shown in Chapter 2. In the same chapter, the market state indicators for normalising the selected DAMs are proposed and compared. They indicate the relative relationship between renewable power generation and electricity demand per hour. In Chapter 3, the history and mathematical theories of the Neural Networks (NN) are explained. Further from this, this chapter elaborates the formation of the cGAN framework and improvements to conditional Wasserstein Generative Adversarial Networks with Gradient Penalty (cWGAN-GP), attached with introduction of two rival methods, CVAE and multivariate Gaussian (mGaus). Prior to training the models, building evaluation metrics, choosing the model parameters and preparing the conditions for LEPP under Dutch 2050 energy scenarios are explained in Chapter 4. Next, Chapter 5 present results of the model training, including model performance by using only Dutch data or combined Dutch, Spanish and Danish data, by each proposed market state normalisation solution, and by each selected projection method. This chapter also compares projected price series with those from the Energy Transition Model (ETM) on price characteristics, hourly inter-dependencies and profit potential acquired by solving a optimization problem designed in this study. The last part of this chapter adds extensive discussions on the presented results. Finally, the thesis is ended with conclusions in Chapter 6.

2

FEATURE SELECTION

Since this study is to conduct Long-term Electricity Price Projection (LEPP) 30 years ahead, selecting features relevant to this task is particularly important. In this chapter, the choice of projecting 24-hour Day-Ahead Market (DAM) price series is reasoned. A candidate feature set is formed through a literature survey, with corresponding data acquired from open sources. Due to the novelty of price series projection, no explicit feature sets for similar tasks has been pointed out, to the author's knowledge. Hence features relevant to DAM price point prediction are assumed useful for the LEPP in the study. The features are examined via Exploratory Data Analysis (EDA), as well as importance analyses using Principal Component Analysis (PCA) and Random Forest (RF) regressors. And the results from these are verified with Recursive Feature Elimination (RFE). For the purpose of market state normalisation, market state indicators measuring the relative hourly difference between renewable power generation and electricity demand are put forward, namely Renewable Scarcity Factor (RSF), Renewable Load Ratio (RLR) and Renewable Load Satisfaction Ratio (RLSR). Finally, The effectiveness of market state normalisation with each of these indicators is compared with a simple scaling with total installed generation capacity.

2.1. DAY-AHEAD MARKET (DAM) PRICE SERIES

As shown in Figure 1.3a and 1.3b, it is needed to include data from the Spanish market (ES) and two Danish markets (DK1 and DK2) because their market states are obviously closer to the Dutch 2050 energy scenarios in renewable power capacity and generation (hydrogen not considered temporarily). Therefore, DAM prices of the Dutch market (NL) and the three mentioned markets are gathered from the ENTSO-E transparency platform [47].

It is proposed to project DAM price in daily 24-hour series. This is originated from the fact that these prices are cleared in the market at the same time prior to the delivery day. Hence the information (relevant features) to derive the 24 hourly prices shall be the same. To consolidate this, a frequency analysis on the DAM price is conducted. The

frequency analysis identifies the variational periods of the DAM price. This can be done with the discrete Fourier transform.

The discrete Fourier transform is suitable in this case as the DAM price can be seen as consecutive in time, thus one-dimensional. And, its values are discrete by hour. The one-dimensional discrete Fourier transform takes an input trace $\boldsymbol{x} = [x_1, x_2, ..., x_N]$ and produces the period amplitude trace $\boldsymbol{X} = [X_1, X_2, ..., X_K]$ [53], each element is given as follows,

$$X_k = \sum_{n=1}^{N} x_n \cdot e^{-\frac{i2\pi}{N} kn} = \sum_{n=1}^{N-1} x_n \cdot [cos(\frac{2\pi}{N} kn) - i \cdot sin(\frac{2\pi}{N} kn)],\ k = 1, 2, ..., K,\ K \leq N, \quad (2.1)$$

where e is the base of the natural logarithm, i the imaginary number. All possible frequencies are evaluated when the dimension of input trace N is equal to the dimension of output period amplitudes K. To calculate, the transformation can be realised with the *fft* function from the *Numpy* library in Python [54]. Instead of complex numbers, the function returns the magnitude of the values from the transformation in Eq. (2.1). The index k of the output period weights denotes the period of variation (number of steps in the input trace).

The discrete Fourier transform is conducted on the DAM price in the combined NL, ES, DK1 and DK2 DAM price set. Result of the discrete Fourier transform is shown in Figure 2.1. Considering the variational periods from 1 to 100 hours, all the significant period amplitudes considerably higher than 10^5 € occur in periods less than 24 hours. Hence, it can be concluded that projecting DAM prices in daily 24-hour series would capture the most fundamental patterns of variation, while fitting with the way of DAM clearing.

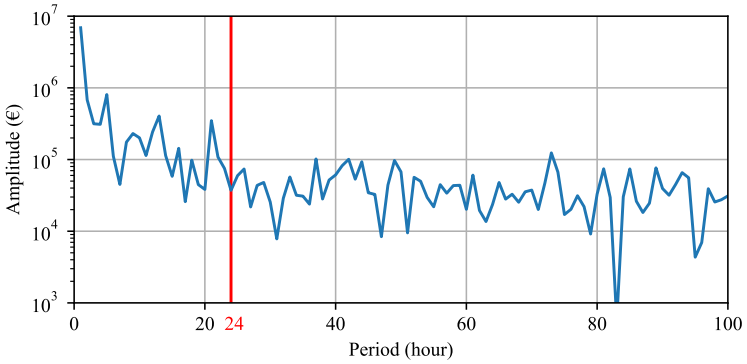


Figure 2.1: Frequency analysis of the combined NL, ES, DK1 and DK2 DAM price from 2015 to 2019.

2.2. CANDIDATE VARIABLES AND PRE-PROCESSING

As daily 24-hour DAM price series are to be generated, it can be hypothesised that the information on time, i.e. the month of the year and the day of the week should be taken as features. This hypothesis is bolstered by the monthly and weekly patterns of DAM prices in many markets around the world, which is well-noticed in scientific literature,

e.g. [55]. With that said, it is assumed that which week in the same month and which year it is are not pertinent to be chosen as candidate features.

As introduced in Section 1.1, electricity market prices are determined by bids of both the supply and the demand. From this knowledge, we decide to include the electricity load forecast from the ENTSO-E transparency platform [47] as one of the candidate features for projecting DAM prices. The day-ahead load forecast is used because, likewise, the DAM is cleared one day before delivery. Plus, in Electricity Price Forecast (EPF) studies, the load has been frequently taken as a feature variable as well [27].

It has been demonstrated, e.g. in [56, 57], that local temperature influences the electricity demand and capital investments of energy project. Moreover, the temperature is regarded relevant to power output of solar energy generators, owing to the correlation between the in-situ temperature and solar irradiance [58]. As both the supply and the demand are related, the temperature is anticipated to be a relevant candidate feature for projecting DAM prices. Thus, population-weighted national mean temperature series of the Netherlands, Spain and Denmark are acquired and derived from the Open Power System Data platform [59]. Population-weighted temperature is chosen to avoid distractions from the geographical distribution of weather stations, as well as to limit contributions from rarely inhabited areas. Nevertheless, for the purpose of LEPP, temperature in hourly resolution would be unnecessary as a weather indicator. The temperature series is then down-sampled to daily mean, in order to be presented as a daily condition.

Owing to the double-blind merit-order clearing principle in the DAM, the cost of primary energy (fuel price for fossil-based generators) paid by the final generating technology shall be considered as a candidate feature. Judging from the market states of electricity generation in Figure 1.3b and generation flexibility of the technology, gas has often been that technology in the Netherlands, Spain and Denmark. Besides, including gas price as one of the predictor variables has been done in many EPF studies in EPF [60]. For example, Mosbah and El-hawary [61] concluded that gas price has high impact on electricity prices. Therefore, the non-household natural gas price are taken as a candidate feature. Within the author's reach, the Dutch domestic non-household gas price is taken from the Eurostat [62]. On this platform, the gas price is updated twice per year. This would be favorable because it is not the fluctuations in energy markets but the trend of fuel cost of the final generating technology in the DAM that this feature should reflect on.

As explained in Section 1.1, electricity bidding prices from renewable energy generators are fundamentally lower, due to the zero marginal costs (operational costs in generation not counted). The renewable energy sources, primarily the solar radiation and wind, however, are not controllable. Therefore it can be hypothesized that the renewable power generation forecast would provide essential information for DAM price projections. Likewise to the load forecast, hourly renewable power generation forecast can be obtained from the ENTSO-E transparency platform [47].

Considerable increase in the renewable power generation capacity is still taking place. As can be seen in Figure 1.3a, different scenarios result in distinct generation capacity compositions. Hence, the annual total installed capacity of renewables is taken as another candidate feature variable from the ENTSO-E transparency platform [47].

The target variable and the chosen seven candidate feature variables are summarised

in Table 2.1. For each variable, a symbol is assigned for clearer reference later. Note that the fifth feature, F_{gp}, actually denotes the fuel price of the most frequent final generation technology in the DAM.

Table 2.1: The target variable and candidate feature variables in this study.

	Target variable	**Symbol**
	DAM price (€/MWh)	P
No.	**Candidate feature**	**Symbol**
1	Day of the week	F_D
2	Month of the year	F_M
3	Total hourly load forecast (MW)	F_L
4	National daily mean temperature (°C)	F_T
5	National non-household gas price (€/KWh)	F_{gp}
6	Hourly renewable power generation forecast (MW)	F_{gen}
7	Total installed capacity of renewables (MW)	F_{cap}

2.3. FEATURE SELECTION PROCEDURE

Unlike typical classification and regression models, generative models are difficult to be evaluated with certain scores. It is not desired for the generator to produce exactly the samples identical to the real ones. Usually, researchers prefer that the generated sample values have similar distribution with that of the real samples. In several studies like [63, 64], the distribution of generated sample values is examined with Quantile to Quantile (Q-Q) plots, where certain quantile values of the generated samples are plotted against values of the same quantiles in the real samples. Another approach is the Continuous Ranked Probability Score (CRPS), given as

$$CRPS(F, y_t) = \int_{-\infty}^{\infty} (F(\hat{y}_t) - H(\hat{y}_t - y_t))^2 d\hat{y}_t, \tag{2.2}$$

where F is the cumulative distribution function, $\hat{y}_t$ the projection for the time t, y_t the real sample value and H the Heaviside function which outputs 1 if $\hat{y}_t \geq y_t$, or 0 otherwise. CRPS, nonetheless, only examines the distribution of generated sample values against a single real sample or real samples under the same conditions. Therefore, it is not suitable for evaluating multi-conditional samples generation, as in our case, especially when the conditions can change continuously.

Faced with the challenge of lacking evaluating metrics, an assumption can be made to enable the feature selection in this study, that

Assumption: Features useful for DAM hourly price point prediction are considered relevant, and therefore taken for DAM price series projection.

One support behind this assumption is the relationship between the relevance and the usefulness of features for solving a problem. In general, the relevance of a feature is about how much information the feature contains that is intrinsically related to solving the problem. And the usefulness of a feature depends on how well a chosen model

can use the information in that feature to solve the problem [65]. In most of the time, useful features are relevant, whereas some relevant features can be useless [66, 67]. For example, a predictor variable that has a periodic functional relationship with the target variable is likely to be useless for a linear (one-dimensional) model. Thus, useful features are usually a subset of the relevant features.

Point prediction of the DAM price would be a sufficient surrogate task for the price projection. Although there is seemingly a gap that the interrelationship between hourly prices in a day is overlooked in point prediction, it is hypothesized that models trained by the cWGAN-GP framework is capable of extracting the hourly interrelationship in the real DAM price series. Hence only the information of price values, without the price patterns, is necessary for evaluating the relevance of the candidate feature variables.

For feature selection, the Dutch data of DAM prices and candidate feature variables listed in Table 2.1 are gathered. The data set of the target and the candidate feature variables are pre-processed in an hourly manner. The data on the time-saving hour in each year are removed. Data of days with any Not a Number (NaN) values are deleted. Finally, when necessary, the z-score standardisation is conducted on each variable, which is

$$z = \frac{x_z - \mu_z}{\sigma_z}, \tag{2.3}$$

where z and x_z are the standardised and the original values, while μ_z and σ_z are the mean and the standard deviation of the variable, respectively.

2.4. EXPLORATORY DATA ANALYSIS

Promoted by Tukey [68], an American mathematician and statistician, the Exploratory Data Analysis (EDA) is recommended before tackling any problem with data. Although there is no standard procedure in the knowledge of the author, the EDA usually includes plotting the target variable against the predictor variables, as well as showing the correlation between each two of the variables. In the following paragraphs, the correlation and selected plots of month of the year F_M, day of the week F_D, the load forecast F_L and the renewable power generation forecast F_{gen} are displayed and elaborated for their apparent relationships.

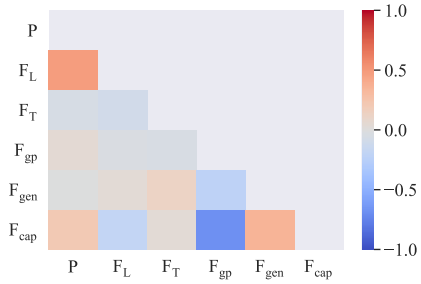
To unveil whether there is a linear relationship between each feature variable and the DAM price, correlations between these variables can be calculated. Invented by Francis Galton [69], the correlation is defined as

$$corr(X,Y) = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}, \tag{2.4}$$

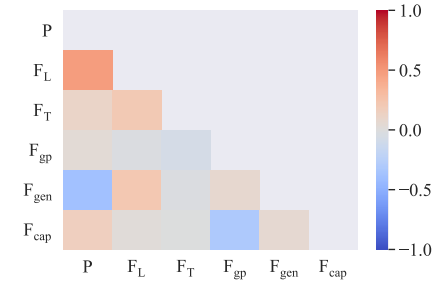
where X and Y are the two variables, E the mean function, μ_X, μ_Y, σ_X and σ_Y are the means and standard deviations of the two variables, respectively.

The calculated correlations are plotted in a heat-map, as shown in Figure 2.2. As expected, the DAM price correlates positively with the load forecast in all the four DAMs. It is interesting to see that in the NL, DK1 and DK2 markets, the electricity load forecast is slightly negatively correlated with the national daily mean temperature, while in the ES market, the correlation is positive. This could be explained with the demand disparity in air cooling and heating at different locations. Due to their low update frequency,

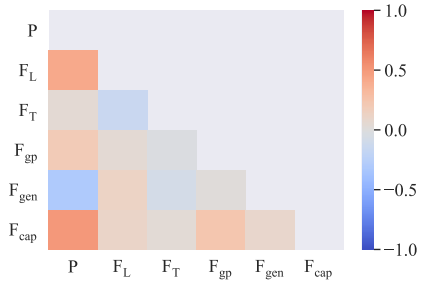
the relationship between the total installed renewable capacity and the national non-household gas price can be viewed in a contextual perspective. The negative correlation between these two variables in the Dutch market might be a result from the country's continuous additions in renewable capacity while the gas price was in decline. Last but not least, the negative correlation between the DAM price and the renewable generation forecast observable in the ES, DK1 and DK2 markets is absent in the Dutch market. It can be inferred that the renewable generation in the Dutch market had not been sufficient enough to have a significant influence on the DAM price. This will be investigated further in the plots below.



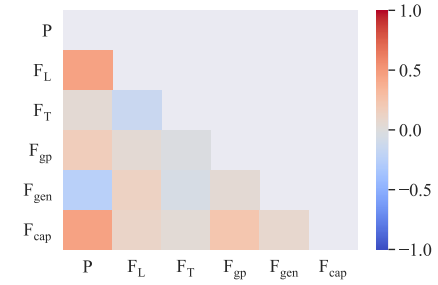
(a) Correlations of variables on the NL market.



(b) Correlations of variables on the ES market.



(c) Correlations of variables on the DK1 market.

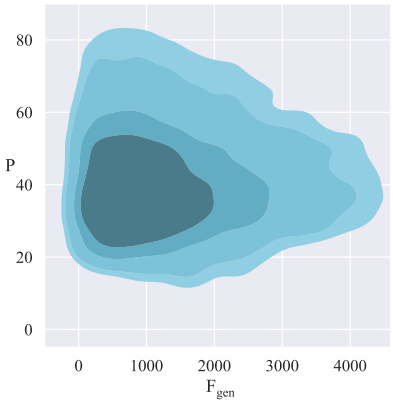


(d) Correlations of variables on the DK2 market.

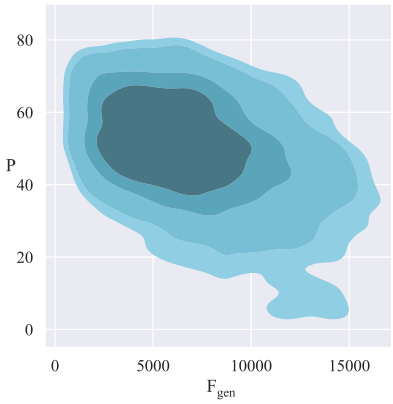
Figure 2.2: Correlation matrices of the target variable and the candidate feature variables on the NL, ES, DK1 and DK2 markets from 2015 to 2019, respectively.

When plotting the DAM price against the feature variables, two methods are adopted. The first is the filled Kernel Density Estimation (KDE) plot for feature variables that can change continuously. In the KDE-plots, regions in the variable space are color-filled based on the cumulative probability density function calculated with KDE from the data set. Principles and calculation steps of KDE can be found in [70]. The second is the enhanced box-plot, used when the feature variable is in discrete values. Put forward by Hofmann et al. [71], the enhanced box-plot shows more quantiles than the original box plot. Thus the enhanced box-plots are more suitable to display data distribution of large data sets.

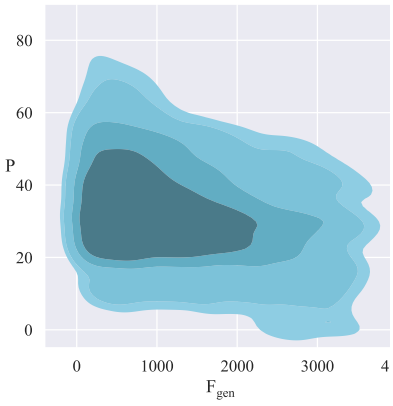
The absence of negative correlation between the hourly DAM price and the renewable generation forecast in the Dutch market is further investigated with KDE-plots. Fig-



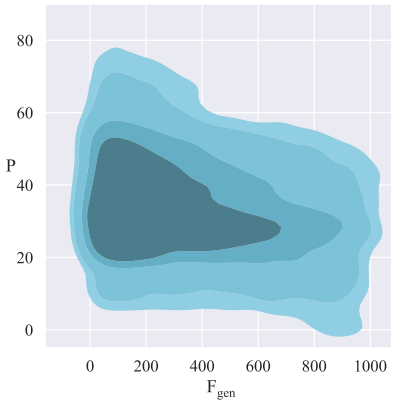
(a) P vs. F_{gen} on the NL market.



(b) P vs. F_{gen} on the ES market.



(c) P vs. F_{gen} on the DK1 market.



(d) P vs. F_{gen} on the DK2 market.

Figure 2.3: DAM prices against renewable power generation forecasts per hour in KDE-plots for the NL, ES, DK1 and DK2 markets from 2015 to 2019. Regions from dark to light includes approximately 50%, 70%, 90% and 95% of the data points, respectively.

ure 2.3 plots the two variables against each other for the NL, ES, DK1 and DK2 markets separately. Three points can be obtained from the plots. The first is that, as the renewable generation forecast increases, variance of corresponding DAM prices converges. This is clearest in the Dutch market in Figure 2.3a. Secondly, it can be seen that the trend of the DAM price declining with the renewable generation forecast does exist in the ES, DK1 and DK2 markets. Given that the three markets lead the Dutch market in shares of renewable generation, it is deduced that this phenomenon would appear when there is enough renewable power in the market. Lastly and interestingly, in the periphery (90% to 95% region) of the colored areas for the three markets, there are certain times when the DAM price drops considerably towards zero. These should be the cases when bids from renewable power generators alone are enough to meet the demand, since it only takes place on high renewable generation forecasts, judging from Figure 2.3b, 2.3c and 2.3d. As the negative correlations and the sudden drops of the DAM price are neither seen in the Dutch market, while they are regarded consequences of the renewable penetration, it is verified that the ES, DK1 and DK2 markets are beneficial supplements to the data set.

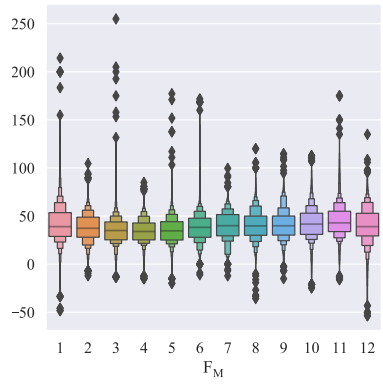
How the DAM price distributes in each month of the year and each day of the week in the NL, ES, DK1 and DK2 markets across 2015 to 2019 are illustrated together with enhanced box-plots in Figure 2.4. Regarding Figure 2.4a, DAM prices in warm seasons, from April to October, are relatively lower and less variant. Whereas during winter time, the prices are higher and more scattered towards large values over 100 €/MWh. Along with the higher prices, however, in December and January, the DAM price is more likely to go below zero. Possibly, this is owing to the lower demand during the vacation periods in Europe, when some generators rather submit negative bids to stay in operation. The weekly seasonality of the DAM price is displayed in Figure 2.4b. Notably, the mean of DAM prices during the weekend is lower than that in weekdays, also with lower variance. In the weekend, DAM prices on Sunday is typically lower than that on Saturday. Among weekdays, aside from a few higher outliers on Thursday, there are no conspicuous differences in distribution.

From the analysis above, intriguing relationships to the DAM price has been pointed out for the load forecast, renewable generation forecast, month of the year and day of the week. This does not rule out the other three candidate features, national daily mean temperature, national non-household gas price and total installed capacity of renewables. Possible exclusion of features depends on the feature usefulness, which will be presented with quantitative analyses in the next section.

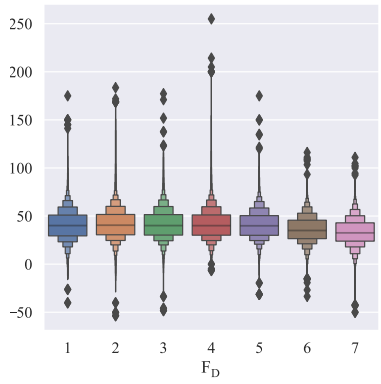
2.5. IMPORTANCE ANALYSES

To investigate the importance of the candidate feature variables, an unsupervised technique, Principal Component Analysis (PCA), and a supervised technique, Random Forest (RF) regression, are implemented. A technique is supervised if correct response is used in training, otherwise it is unsupervised. Both techniques can evaluate the candidate feature variables in a way that indicates the variable importance (usefulness). In this process, the data set is z-score standardised with Eq. (2.3).

When necessary, PCA is applied to help avoid over-fitting and focus on the main variable components extracted from the features, e.g. in [72]. Principles of PCA can



(a) Distribution of DAM prices per month of the year. F_M=1 denotes January.



(b) Distribution of DAM prices per day of the week. F_D=1 denotes Monday.

Figure 2.4: Distributions of DAM prices in the NL, ES, DK1 and DK2 markets by month of the year and day of the week in enhanced box-plots, respectively. The biggest boxes denote the middle 50% of data points in the category.

be explained as to find a linear coordinative vector conversion that merges all the feature dimensions to maximise variance of the data points to one axis [73]. Eigenvalue of the transformation matrix of each original dimension denotes the component of that dimension in the new axis, hence indicating its contribution in the maximised variance. Accordingly, the Contribution Rate (CR) of each variable is given by

$$CR_i = \frac{\lambda_i^{pca}}{\sum_{j=1}^{p} \lambda_j^{pca}} \times 100\%, \tag{2.5}$$

where λ_i^{pca} is the eigenvalue of the candidate features' co-variance matrix, corresponding to the ith principal component. And p is the number of candidate feature variables, seven in our case.

Calculated CR of the candidate feature variables in the combined data set of the NL, ES, DK1 and DK2 markets can be found in Table 2.2. Except for the Dutch market, month of the year and day of the week are ranked the highest, reflecting on the seasonalities. The higher contribution rates of these variables might result from the fact that they are nearly evenly distributed in the data set, resulting in higher eigenvalues of the co-variance matrix than others. Moreover, none of the candidate feature variables are rated unimportant, with the lowest CR of 7.4% by the national non-household gas price in the Spanish market.

Aside from the unsupervised importance analysis with PCA, RF regression is chosen as a more advanced technique that takes the target variable into account. RF is an ensemble model of multiple tree estimators, each taking features randomly to make predictions [74]. This way, the importance of a variable can be calculated by evaluating the improvement on a given metric, regarding all the trees and their nodes that use the variable [75]. In this study, each RF model contains 100 tree estimators, which is set by

Table 2.2: CR of the candidate feature variables in PCA, sorted in descending manner. Data set encompasses features in the NL, ES, DK1 and DK2 markets from 2015 to 2019.

No.	Feature	Contribution rate			
		NL	ES	DK1	DK2
1	F_M	27.2%	21.0%	21.7%	21.7%
2	F_D	3.7%	20.6%	18.1%	18.1%
3	F_L	17.8%	17.5%	17.2%	17.2%
4	F_T	18.9%	14.4%	14.1%	13.8%
5	F_{gp}	9.2%	7.4%	8.6%	8.5%
6	F_{gen}	12.2%	9.4%	9.8%	9.8%
7	F_{cap}	11.1%	9.9%	10.6%	10.8%
Sum		100.0%	100.0%	100.0%	100.0%

default like other model parameters from the *scikit-learn* library [76] in Python. Each RF model is evaluated with the R2 score, given as

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}, \tag{2.6}$$

where y is the actual value, $\hat{y}$ the prediction, $\overline{y}$ the mean of the actual values, and n the number of samples. Higher the R^2, the more predictions made by the model are better than simply using the mean.

RF model are trained with data of the target variable and the candidate feature variables for the NL, ES, DK1 and DK2 markets, giving variable importance as shown in Table 2.3. All the RF models give good performance, with the lowest R^2 score of 0.799. In contrast to the PCA results in Table 2.2, month of the year and day of the week are ranked much lower this time than other variables. An interpretation can be made that even though these time-related variables have relatively higher variance, they do not provide as much information to predict the DAM price as other feature variables. In fact, the two variables only provide information from one thing, the time. Day of week is the candidate feature variable ranked the least important. We still keep it in the feature set because the weekly pattern is obvious in 2.4b. Presumably, this feature can be substituted with a binary variable denoting whether it is in the weekend to get more importance. Yet, a sufficiently large and well trained neural network model should extract the differences from the day of the week. Furthermore, national non-household gas price, renewable generation forecast and total installed renewable capacity rise RF rating, meaning that despite their lesser contributions in data variance, they are useful for price prediction, on the opposite to the cases of month of the year and day of the week.

Viewing the two ratings, load forecast has been rated important consistently, with at least 17.2% in CR and at least 21.9% in RF importance. As for other feature variables, a reciprocal trend can be identified, i.e. the features low in CR are generally high in RF importance and vice versa. Also considering the demonstrated price seasonalities, it can be confirmed that all the selected variables are useful for DAM price point prediction, and therefore relevant for DAM price series projection.

Table 2.3: Variable importance from the RF model with R^2 scores. Data set encompasses the DAM price and features in the NL, ES, DK1 and DK2 markets from 2015 to 2019.

No.	Feature	Importance ×100%			
		NL	ES	DK1	DK2
R^2		0.799	0.932	0.837	0.804
1	F_M	7.0%	18.0%	6.1%	7.1%
2	F_D	3.4%	1.5%	3.5%	3.4%
3	F_L	34.1%	27.1%	21.9%	26.6%
4	F_T	11.1%	4.0%	9.3%	10.2%
5	F_{gp}	13.0%	12.7%	12.1%	12.3%
6	F_{gen}	12.3%	24.5%	21.4%	18.2%
7	F_{cap}	19.1%	12.2%	25.7%	22.2%
Sum		100.0%	100.0%	100.0%	100.0%

To bolster the decision to include all seven features, Recursive Feature Elimination (RFE) is conducted by comparing the best RF model performance on different number of features included, for each of the NL, ES, DK1 and DK2 markets separately. The best model performance is derived exhaustively, from results of all possible feature combinations. The model parameters are identical to those in the previous RF importance analysis. And R^2 is again the evaluation metric.

The result of the conducted RFE is shown in Figure 2.5. As can be seen for all four markets, there is always a growth in R2 score and the growth diminishes with the number of features, confirming that the RFE selects the most useful features each time. For the NL, DK1 and DK2 markets, considerable growths in model performance continue until the number of features reaches six. The final added feature variable is the day of week. As already stated, keeping this feature is necessary for the weekly seasonality. Although one can argue that such seasonality should already be embodied in other features like renewable generation forecast and load forecast, a dissection of the weekly pattern from the supply-demand relationship would still be preferred. This is because while the latter could change drastically across the 30 years' period, the day-to-day relationship in a week might not be as different.

2.6. MARKET STATE NORMALISATION METHODS

Although information in all the seven candidate variables are found indispensable from the importance analyses, the scales of the NL, ES, DK1 and DK2 markets are unaligned, e.g. the differences in total installed generation capacity listed in Figure 1.3a. Thus, it is meaningless to directly combine data of these four markets. Normalisation of the market states is necessary. More specifically, with the normalised market states, the four markets can be seen as homogeneous. Hence the higher renewable penetration in the ES, DK1 and DK2 markets can be taken advantage of to project DAM prices under the Dutch 2050 energy scenarios with conditions unprecedented in the Netherlands. To this end, three market state indicators, namely total installed capacity, Renewable Scarcity Factor (RSF), Renewable-load Ratio (RLR) and Renewable Load Satisfaction Ratio (RLSR), are

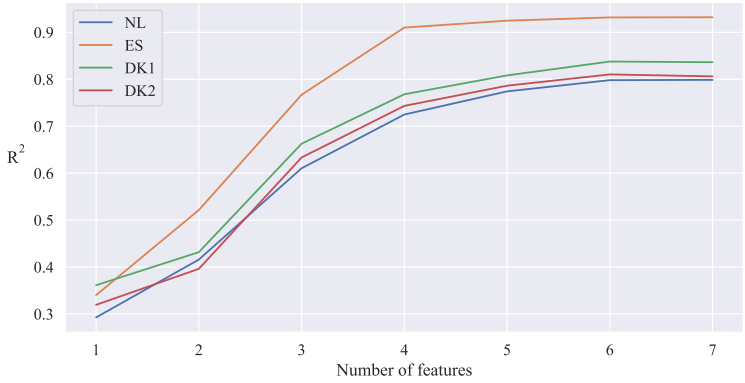


Figure 2.5: Best performance of the RF regressors vs. the number of features used to train. Data set used encompasses the DAM price and features in the NL, ES, DK1 and DK2 markets from 2015 to 2019.

proposed and compared with Normalising with total installed CAPacity (NCAP), respectively.

Features related to the characteristics of the energy system, the load forecast F_L, the renewable generation forecast F_{gen} and the total installed renewable capacity F_{cap}, should be normalised. To eliminate repetitive information, for each solution, features used in forming new features will be dropped in the data set for training. There are three solutions proposed. The first is the Renewable Scarcity Factor (RSF) F_{RSF}, which is calculated by

$$F_{RSF} = \frac{F_L - F_{gen}}{F_{cap}}. \tag{2.7}$$

Secondly, the Renewable-Load Ratio (RLR) F_{RLR} is the renewable generation forecast divided by the load forecast on the hour,

$$F_{RLR} = \frac{F_{gen}}{F_L}. \tag{2.8}$$

Finally, the Renewable Load Satisfaction Ratio (RLSR) F_{RLSR} derives from capping the RLR at 1,

$$F_{RLSR} = \min\left(\frac{F_{gen}}{F_L}, 1\right). \tag{2.9}$$

For comparison and evaluation purposes, a baseline normalisation solution, Normalising with total installed generation CAPacity (NCAP) is formed. In this solution, the scale-related features, F_L, F_{gen} and F_{cap} are simply normalised by total installed renewable capacity to F_{nL}, F_{ngen} and F_{ncap}, given by

$$F_{nx} = \frac{F_x}{\text{Total generation capacity}}. \tag{2.10}$$

where F_x is the feature to be normalised, and F_{nx} the normalised feature.

There is an issue with total installed renewable capacity F_{cap} that if it is being used as a feature and normalised with the total generation capacity as in Eq. (2.10), the corresponding condition in the Dutch 2050 scenarios is not covered by the data set, explained as follows. First of all, this feature is updated every year. Given the monotonic increase of the renewable capacity in these countries, the levels in the Dutch 2050 scenarios are actually unseen in the combined data set, and even in the whole Europe. In Sweden, where the share of energy consumption from renewables is the largest according to the European Commission [77], the total installed renewable capacity devided by the total installed generation capacity, F_{ncap}, is at around 0.63. Whereas in the RM and NM scenarios, the F_{ncap} would be 0.73 and 0.71. Therefore, in the cases of NCAP, RLR and RLSR, the total installed renewable capacity F_{cap} is removed from the data set.

From the information of power generation and capacities in Figure 1.3a and 1.3b, and with Eq. (2.7), (2.8), (2.9) and (2.10), annual mean of the market state indicators can be calculated. Calculation results are shown in Table 2.4.

Table 2.4: Annual mean of the market state indicators calculated from the four scenarios.

Scenario	**Annual mean**				
	RSF	RLR	RLSR	nL	ngen
NL 2019	0.92	0.094	0.094	0.30	0.028
RM	0.019	0.93	0.93	0.18	0.17
NM	0.014	0.95	0.95	0.20	0.19
EC	0.098	0.73	0.73	0.21	0.15
IM	0.12	0.69	0.69	0.21	0.14

Comparing with the 2019 reference, the scenarios in Table 2.4 have drastically different indicator values. When the renewable capacity is fixed, the RSF is proportional to the difference between load and renewable generation, the more demand that could be satisfied by renewable power, the lower the RSF value. Hence relative on the renewable capacity, annual mean of the RSF drops from 0.92, the 2019 Dutch market level, to 0.12, the highest among the four scenarios. And this is the opposite in the cases of RLR and RLSR, which rise from 0.094 to at least 0.69. The changes are owing to the joint effort by the relative decrease of load and the relative increase of renewable power generation. Annual mean of the Dutch normalised load decreases from 0.3 in 2019 to 0.21, the highest in the scenarios, while that of the Dutch normalised renewable generation is five-fold from 0.028 to no less than 0.14.

2.7. CAUSALITY VALIDATION

To further investigate the relevance of the candidate feature variables, as well as the proposed feature integration solutions, to the DAM price, the Granger Causality Test (GCT) is conducted as the final step. In studies containing feature selection processes, e.g. [78, 79, 80], only the features with causal relationship to the target variable shown in GCT results, are taken as predictor variables. In brief, GCT examines if variance of error in the prediction by an estimator is reduced with extra information from the feature variable. Taking the definition of causality made by Granger [81] as an example, variable

Y is regarded causing X if

$$\sigma^2_{GCT}(X|U) < \sigma^2_{GCT}(X|\overline{U-Y}), \tag{2.11}$$

where σ^2_{GCT} is the variance of prediction error in GCT. U is all the information in the universe, and $\overline{U-Y}$ is all the information apart from Y. GCT is available from the *statsmodel* library [82] in Python.

A prerequisite of GCT is that the feature variable should be stationary [81]. If not, the relationship between the feature and the target variables might vary across time. The stationarity can be inspected with the Augmented Dickey Fuller (ADF) test, with the updated details in [83]. ADF test can be done conveniently in Python with the library *statsmodel* [84].

It has been argued that GCT can still be conducted even when the feature variable is not stationary, as long as cointegration between the feature and target variables can be found. Cointegration of two variables means that at least one of their linear combinations are stationary. As in the feature selection procedure put forward by Feng et al. [85], a cointegration test follows if the stationarity test fails. GCT is skipped only when failure occurs in both tests. The cointegration test can be found again from library *statsmodel* [86].

All the candidate features have passed GCT, concluding that the null hypothesis in GCT that there is no causal relationship from the features to the target variable has been negated. Results from the three tests can be found in Table 2.5. In ADF test, only the national non-household gas price F_{gp} is tested positive for the null hypothesis of non-stationarity, while this feature has passed the following cointegration test.

Table 2.5: Results from GCT, ADF and cointegration tests showing the cointegration and causality to the DAM price. Note that Y denotes negation of the null hypotheses that there is no stationarity, cointegration or causality, and N otherwise.

No	Candidate Feature	Stationarity	Cointegration	Causality
1	F_D	Y		Y
2	F_M	Y		Y
3	F_L (F_{nL})	Y		Y
4	F_T	Y		Y
5	F_{gp}	N	Y	Y
6	F_{gen} (F_{ngen})	Y		Y
7	F_{cap} (F_{ncap})	Y		Y
8	F_{RSF}	Y		Y
9	F_{RLR}	Y		Y
10	F_{RLSR}	Y		Y

2.8. SELECTED FEATURE VARIABLES SUMMARY

The selected feature variables are summarised by each proposed solution to the normalisation of DAM market states in Table 2.6. Every solution contains one or two market state indicators. For the NCAP solution, these are the load forecast and the renewable

generation forecast each divided by total installed generation capacity, F_{nL} and F_{ngen}. For the RSF, RLR and RLSR solutions, the market state indicators are F_{RSF}, F_{RLR} and F_{RLSR}, respectively.

It should be noticed that the number of features is not equal to the number of dimensions. Since the goal of this study is to project 24-hour price series, and due to the high correlation between the DAM price and the load on each hour, it is decided to include hourly values, thus in 24 dimensions, for the market state indicators (F_{nL},F_{ngen},F_{RSF},F_{RLR} and F_{RLSR}). If only one dimension is given for each of these features, e.g. via taking the mean over each 24 hours, it would not be possible to investigate the influence of load and generation on certain hours of the day to the hourly DAM price series.

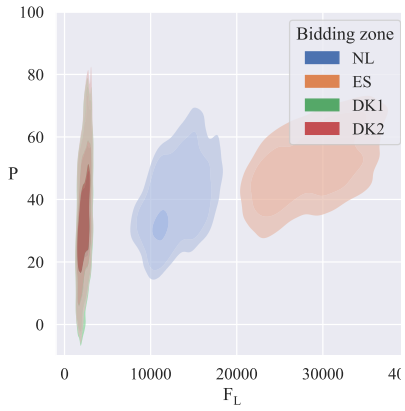
Table 2.6: Summary of the feature variables in the proposed NCAP, RSF, RLR and RLSR feature sets.

Solution	Feature set
NCAP	$\{F_D,F_M,F_T,F_{gp},F_{nL},F_{ngen}\}$
RSF	$\{F_D,F_M,F_T,F_{gp},F_{RSF}\}$
RLR	$\{F_D,F_M,F_T,F_{gp},F_{RLR}\}$
RLSR	$\{F_D,F_M,F_T,F_{gp},F_{RLSR}\}$

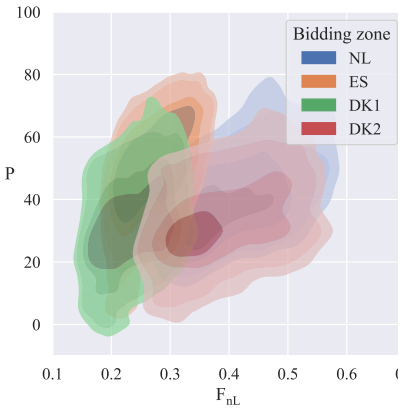
2.9. EFFECTIVENESS OF THE MARKET STATE INDICATORS

Finally, the effectiveness of using the market state indicators to normalise the NL, ES, DK1 and DK2 markets can be explicated. Illustration using KDE-plots is suitable for this task, because of the intuitive display of the relationship between two continuous variables. Figure 3.5 shows how the DAM price distributes with the selected indicators in each of the four markets. When the data sets are not normalised, as in Figure 2.6a, the data points of the four markets are clustered distantly in the plot, which is not suitable for DAM price projection with continuous condition values. When the load forecast is normalised with the total installed generation capacity, the data regions of the four DAMs become similar in scales, overlapping in the center, and each market extends the boundary. An analogous situation can be found in Figure 2.6d. When using the RSF as the market state indicator, the data regions again gather together. Nevertheless, the overlapping area in this case is much smaller, implying that more conditional region has been covered in this manner.

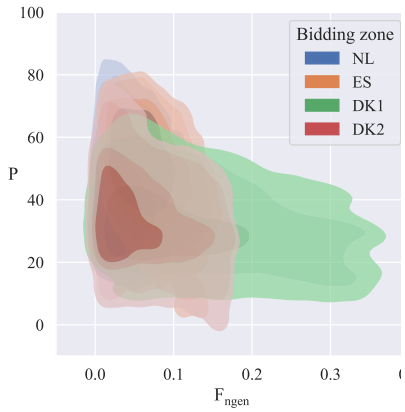
When the market state indicator is (partially) proportional to the renewable generation forecast, i.e. the cases of F_{ngen}, F_{RLR} and F_{RLSR}, the data regions are shrunken towards zero with large overlapping areas in Figure 2.6c, 2.6e and 2.6f. The shape of the data regions conforms with the typical long-tail distributed renewable power generations, e.g. Weibull distribution to model solar and wind power output probabilities [87, 88]. Judging from the large overlapping areas, normalisation with these market state indicators can be considered inferior to those with F_{nL} and F_{RSF}. Among these two, F_{RSF} seems better than F_{nL}.



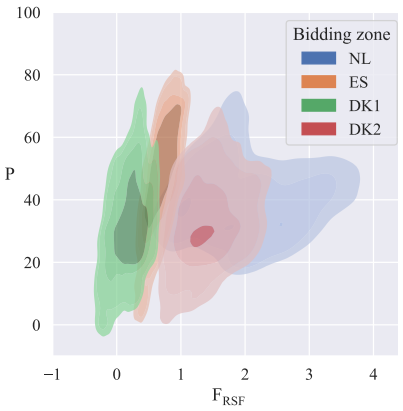
(a) P vs. F_L, without market state normalisation.



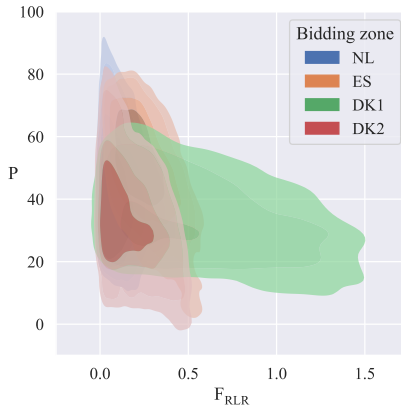
(b) P vs. F_{nL} from the NCAP normalisation solution.



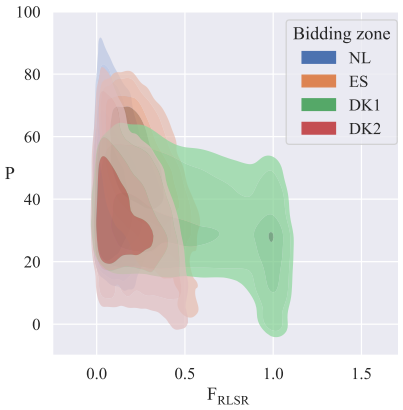
(c) P vs. F_{ngen} from the NCAP normalisation solution.



(d) P vs. F_{RSF} from the RSF normalisation solution.



(e) P vs. F_{RLR} from the RLR normalisation solution.



(f) P vs. F_{RLSR} from the RLSR normalisation solution.

Figure 2.6: DAM prices against selected market state indicators in KDE-plots for each of the NL, ES, DK1 and DK2 markets from 2015 to 2019. Regions from dark to light includes approximately 50%, 70%, 90% and 95% of the data points, respectively.

3

Conditional Wasserstein Generative Adversarial Nets with Gradient Penalty (cWGAN-GP)

The two aforementioned generative methods suitable for the task in this thesis, conditional Generative Adversarial Nets (cGAN) and Conditional Variational Auto-Encoders (CVAE), are both based on neural networks. Before applying the methods to tackling specific problems, it is necessary to understand how neural networks work as learning machines. To this end, this chapter starts with briefing on the history of neural networks. Subsequently, neurons in the neural networks is introduced in its primal form, perceptions as binary classifiers, together with the neuron activation functions. With layers of neurons, how a neural network learns via gradient descent of the neuron weights in back-propagation is elaborated. Improvements developed for the learning process that enable adaptive tuning of the network learning rate are also discussed. Upon this, the cGAN and improvements to the conditional Wasserstein Generative Adversarial Nets with Gradient Penalty (cWGAN-GP) are explained, together with the CVAE and the multivariate Gaussian distribution (mGaus). Finally, a test case on using the three methods to reproduce sinusoidal curves with changeable amplitude and Gaussian noise level is conducted as a heuristic check on feasibility of the methods.

3.1. History of Neural Networks in a Nutshell

The concept of neural networks originates from the work by y Cajal [89], a Spanish neuroscientist, histologist, pathologist and Nobel laureate who discovered that the basic elements of the human brain is the *neuron*. There are approximately 68 billion neurons in the human brain [90]. Each neuron is connected to others via *synapses*, which typ-

ically receive signals from and send signals to neurons with chemical reactions, while transmitting in electrical pulses [91].

The field of neural networks and Artificial Intelligence (AI) commenced when McCulloch and Pitts [92] constructed a number of primitive computational neural networks, which contained neurons that could only be turned on or off, and were not affected by their past states. In addition, they found that networks with different configurations, i.e. variant ways of internal connection, are essentially capable of producing the same results. Hence in our study, only the simple and most popular feed-forward networks are considered.

Although studies in neural networks pioneered in the field of AI before the 21st century, neural networks were outperformed by other machine learning methods till 2010 [93]. With the computational power rocketing in recent years, neural networks have gained advantages. Especially, many research efforts have been made on improving convolutional networks and recurrent networks that lead to applications in industries [94, 95].

3.2. THE NEURONS

As a primal concept, pointed out firstly by Rosenblatt [96], each neuron acts as a *perceptron* in a neural network. Originally, a perceptron can be regarded as a linearly separable binary classifier. To elaborate, let us consider input-output data set with underlying functional relationships $(\mathbf{x}_n, y_n), n = 1, 2, ..., N, y_n \in \{-1, +1\}, \mathbf{x}_n \in \mathbb{R}^l$, where N is the number of samples and l the number of independent variables in each sample. Here the bold symbol $\boldsymbol{x}$ means that it is a vector, containing an array of values. This norm will be followed thorough out this thesis. The task is to find a hyperplane

$$\boldsymbol{\theta}^T \mathbf{x} = 0, \tag{3.1}$$

so that

$$\boldsymbol{\theta}^T \mathbf{x} > 0, \quad \text{if } \mathbf{x} \in \omega_1, \tag{3.2}$$

$$\boldsymbol{\theta}^T \mathbf{x} < 0, \quad \text{if } \mathbf{x} \in \omega_2. \tag{3.3}$$

where $\boldsymbol{\theta}$ are the synaptic weights of the neuron, and ω_1 and ω_2 are the two classes. Equalisation to zero is excluded in the constraints as the classification problem is assumed linearly separable. The cost of each perceptron is defined to penalise for misclassified samples, given as [93]

$$J(\boldsymbol{\theta}) = -\sum_{n:\mathbf{x}_n \in \Upsilon} y_n \boldsymbol{\theta}^T \boldsymbol{x}_n, \tag{3.4}$$

where

$$y_n = \begin{cases} +1, & \text{if } \boldsymbol{x} \in \omega_1, \\ -1, & \text{if } \boldsymbol{x} \in \omega_2. \end{cases} \tag{3.5}$$

And Υ refers to the samples misclassified. The cost is at its minimum 0 when all samples are classified correctly.

As the goal is to minimise the cost, the iterative update rule of the synaptic weights can then be given by taking the negative gradient of Eq. (3.4) [96]

$$\boldsymbol{\theta}^{(i)} = \boldsymbol{\theta}^{(i-1)} + \mu_i \sum_{n:\boldsymbol{x}_n \in \Upsilon} y_n \boldsymbol{x}_n, \quad i = 1,2,3... \tag{3.6}$$

where i is the iteration number, and μ_i is a parameter often referred to as learning rate. The value of learning rate shall be chosen such that the minimum can be reached.

A neural network consists of multiple neurons. A simple example can be seen in Figure 3.1. The synaptic weights are attached to each connection, multiplying the signals from the previous layer. Each neuron sums the signals, together with a constant bias term $\boldsymbol{b}$.

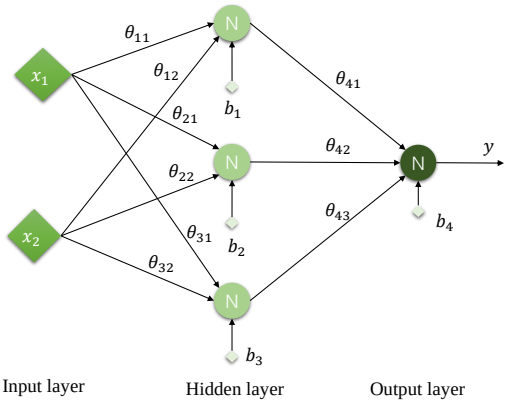
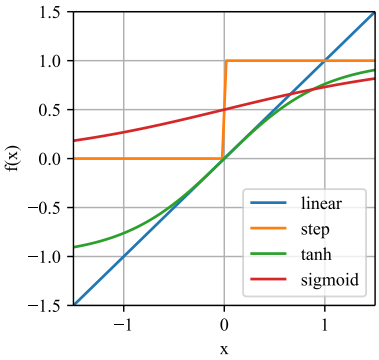


Figure 3.1: A simple feed-forward neural network. The neurons are in circular shape.

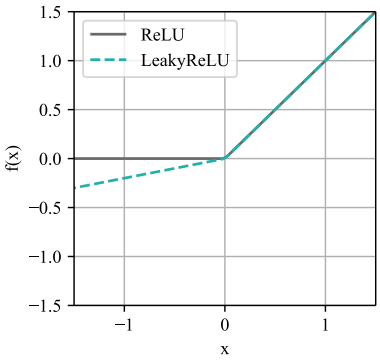
For problems other than linearly separable binary classification, the neurons can be modified by taking other loss (cost) functions or activation functions. The activation function depicts the functional relationship between input to and output from the neuron. In the case above, the activation is linear, transmitting the exact sum of the bias and products of input and corresponding weights. There are other kinds of activation functions, such as the $tanh$ and $sigmoid$ activation functions in Figure 3.2a.

Due to the two practical issues of gradient vanishing/exploding in saturating activation functions and the need for training acceleration, non-saturating activation functions were put forward [97]. Non-saturating activation functions do not have boundaries for their output values. Two outstanding examples are standard rectified linear unit (ReLU) and leaky rectified linear unit (Leaky ReLU), shown in Figure 3.2b.

The ReLU activation function was presented in a paper by Nair and Hinton [98], which demonstrated significant improvements in the image recognition accuracy with ReLU as the activation function in a Convolutional Neural Network (CNN), compared to that of the Guassian binary unit, a saturating activation function. ReLU suppresses input value when it is below zero, and returns the value when input is positive, given as



(a) Standard activation functions.



(b) ReLU and LeakyReLU activation functions.

Figure 3.2: A selection of activation functions.

$$f_{ReLU}(x) = \begin{cases} x & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases} \tag{3.7}$$

Note that the x here refers to the sum of bias and input-weight products, not the input.

The LeakyReLU activation function is a variant of ReLU, with information leakage when the input value is negative. It was well-known in the first place by its implementation in acoustic neural network models [99], which outperformed accuracy in Large-Vocabulary Continuous Speech Recognition (LVCSR), compared to models with $tanh$ and $sigmoid$. The function is

$$f_{LeakyReLU}(x) = \begin{cases} x & \text{if } x \geq 0, \\ \alpha x & \text{if } x < 0. \end{cases} \tag{3.8}$$

where $\alpha \in (0,1)$ is a predetermined parameter.

Among the activation functions, ReLU-based ones are generally considered better in performance, especially for deep learning [100]. It is concluded that ReLU-based activation functions excel because these functions are able to cover the sparsity and dispersion of the data points [97], i.e. they help the neural networks avoid over-fitting while retain necessary information across layers. Therefore, ReLU and LeakyReLU are taken as potential activation functions in this study.

3.3. Gradient Descent and Back-Propagation

In the iterative optimization process, it is expected that the optimum would be reached by applying the gradient descent scheme, similar to the iterations of synaptic weights in Section 3.2. To generalise the process, a loss function can be given as

$$J(\boldsymbol{\theta}) = \sum_{n=1}^{N} L(y_n, f_{\boldsymbol{\theta}}(\boldsymbol{x}_n)), \tag{3.9}$$

where L is the loss function. The trainable parameters are updated at the direction that the loss function decreases most steeply, which is the negative gradient by definition. Thus the update scheme is [101]

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - \mu \frac{dJ}{d\boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}_i}. \tag{3.10}$$

Since there are gradients, the activation function should be differentiable. The step function is therefore not applicable in the scheme.

It would be guaranteed that the global minimum will be found if the loss function is convex, which has only one polar, the global minimum. In terms of non-convex loss functions, the global minimum may never be reached. As can be seen in Figure 3.3, the gradient decent scheme can stop at a local minimum or a saddle point, where the gradient is zero.

To tackle the issue of not reaching the global minimum, several approaches can be taken. One is to accept a local minimum, which is assumed not considerably worse than the global minimum. Randomising the initial synaptic weights is another, so that the global minimum might be captured by training models repeatedly. How some optimization algorithms deal with the issue of searching the global optimum in model training will be further discussed in Section 3.4.

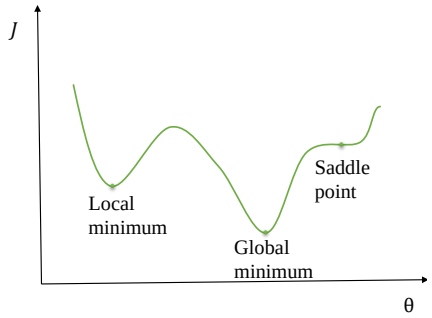


Figure 3.3: An example of non-convex loss function, adapted from [93].

In a feed-forward network, there are multiple layers of neurons. Again, consider training samples $(\boldsymbol{x}_n, y_n), n = 1, 2, ..., N, \mathbf{x}_n \in \mathbb{R}^l$, with $\boldsymbol{x}_n$ as the independent variables and y_n as the dependent variable. Let $\boldsymbol{\theta}_k^r$ denote the synaptic weights to the kth neuron in the rth layer, including the bias term, that is

$$\boldsymbol{\theta}_k^r = [b_k^r, \theta_{k,1}^r, \theta_{k,2}^r, ...], \tag{3.11}$$

and $\hat{y}_{n,k}^r$ is the output of that neuron when training on sample n, $\boldsymbol{z}_n^r$ the input to neurons in that layer, assuming full connection. Each iteration of the synaptic weight optimization comes in two steps.

1. Compute all neuron outputs forwardly.
2. Update the synaptic weights of neurons in every layer backwardly.

The output from each neuron is computed in a forward manner, demonstrated as follows. Firstly, at the first hidden layer $r = 1$, the output is

$$\hat{y}^1_{n,k} = f_{act}({\boldsymbol{\theta}^1_{n,k}}^T \boldsymbol{x}_n), \tag{3.12}$$

where f_{act} is the activation function. For subsequent layers $r = 2, 3...L$, the input to each neuron is

$$\boldsymbol{z}^r_n = [1, \hat{y}^{r-1}_{n,1}, \hat{y}^{r-1}_{n,2}, ..., \hat{y}^{r-1}_{n,K_r}], \tag{3.13}$$

where K_r is the number of neurons in layer r. The output is hence

$$\hat{y}^r_{n,k} = f_{act}({\boldsymbol{\theta}^r_{n,k}}^T \boldsymbol{z}^r_n). \tag{3.14}$$

The neuron output in the output layer $\hat{y}^L_{n,k}$ is the estimate of the dependent variable y_n

$$\hat{y}_n = \hat{y}^L_{n,k}. \tag{3.15}$$

Now that the outputs are acquired, the synaptic weights can then be updated with the loss. According to Eq. (3.9), the loss of the estimate for sample n is $L(y_n, \hat{y}_n)$. Looking at the output layer $r = L$, the output is obtained by

$$\hat{y}^L_{n,k} = f_{act}({\boldsymbol{\theta}^L_{n,k}}^T \boldsymbol{z}^L_n). \tag{3.16}$$

Applying the scheme in Eq. (3.10), the iterative change in the weights is

$$\Delta\boldsymbol{\theta}^L_{n,k} = -\mu \frac{\partial L}{\partial \boldsymbol{\theta}^L_{n,k}} = -\mu \frac{\partial L}{\partial \hat{y}^L_{n,k}} \frac{\partial \hat{y}^L_{n,k}}{\partial \boldsymbol{\theta}^L_{n,k}} = -\mu \frac{\partial L}{\partial \hat{y}} f'_{act} \boldsymbol{z}^L_n. \tag{3.17}$$

For previous layers $r = 1, 2, ...L-1$, the iterative change of neurons in layer r is

$$\Delta\boldsymbol{\theta}^r_{n,k} = -\mu \frac{dL}{d\boldsymbol{\theta}^r_{n,k}} = -\mu \frac{\partial L}{\partial \hat{y}^r_{n,k}} \frac{\partial \hat{y}^r_{n,k}}{\partial \boldsymbol{\theta}^r_{n,k}} = -\mu \frac{\partial L}{\partial \hat{y}^r_{n,k}} f'_{act} \boldsymbol{z}^r_n. \tag{3.18}$$

According to the chain rule, the partial derivative factor can be extended into

$$\frac{\partial L}{\partial \hat{y}^r_{n,k}} = \sum_{i=1}^{K_{r+1}} \frac{\partial L}{\partial \hat{y}^{r+1}_{n,i}} \frac{\partial \hat{y}^{r+1}_{n,i}}{\partial \hat{y}^r_{n,k}}. \tag{3.19}$$

Both $\frac{\partial L}{\partial \hat{y}^L_{n,k}}$ and $\hat{y}^r_{n,k}$ have been calculated using Eq. (3.17) and in the first step. Thus the iterative changes of synaptic weights can only be calculated backwards, layer by layer, hence the term "back-propagation". The synaptic weights are updated for all samples at once. The iterative change is finally

$$\Delta\boldsymbol{\theta}^r_k = \frac{1}{N} \sum_{n=1}^{N} \Delta\boldsymbol{\theta}^r_{n,k}, \quad r = 1, 2, ..., L. \tag{3.20}$$

Here full connection of the network is assumed. The cases of fractional connection can be treated similarly, by handling input to each neuron in the same layer separately. But the presence of this assumption does not affect the validity of Eq. (3.19) and (3.20).

3.4. WEIGHT OPTIMIZATION APPROACHES

Based on the gradient descent back-propagation scheme explained in Section 3.3, advanced weight optimization approaches have been proposed. To train neural networks for LEPP, two well-adopted optimizers, Root Mean Square Propagation (RMSprop) and Adaptive moment estimation (Adam), as well as their prerequisite algorithm, Stochastic Gradient Descent (SGD), will be discussed in this section.

The SGD algorithm is suitable if training time is costly in the study, which normally occurs when the sample set for training is large [102]. In this algorithm, the synaptic weights are updated by SGD on a randomly sampled set with a certain batch size, rather than on the whole sample set. The loss function turns to

$$J_{SGD}(\boldsymbol{\theta}) = L(y_{n_{SGD}}, f_{\boldsymbol{\theta}}(\boldsymbol{x}_{n_{SGD}})). \tag{3.21}$$

where n_{SGD} is a randomly chosen sample number from the N samples. And the iterative change is simply

$$\Delta\boldsymbol{\theta}_k^r = \Delta\boldsymbol{\theta}_{n_{SGD},k}^r, \quad r = 1,2,...,L, \quad n_{SGD} \in \{1,2,...,N\}, \tag{3.22}$$

Following SGD, further improvements have been made on the learning rate. The Adaptive gradient (Adagrad) optimizer has been put forward where learning rate of a weight is adaptive on its own history. In Adagrad algorithm [101], the iterative scheme becomes

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - \frac{\mu}{\sqrt{\boldsymbol{G}_i + \epsilon}}\boldsymbol{g}_i, \tag{3.23}$$

$$\boldsymbol{g}_i = \left.\frac{dJ_{SGD}}{d\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_i}. \tag{3.24}$$

where ϵ is a small value smoother to avoid division by zero, $\boldsymbol{G}_i$ a diagonal matrix with sum of squares of the gradient $\boldsymbol{g}_i$ for a weight, and i the iteration step, given as

$$\boldsymbol{G}_i = \sum_{t=0}^{i} \boldsymbol{g}_i^2. \tag{3.25}$$

The scheme is applied separately to all the weights. As the square of gradient is non-negative, the iterative change of the weights gradually vanishes as the training goes by.

In the Root Mean Square propagation (RMSprop) algorithm, the matrix $\boldsymbol{G_i}$ is replaced by a variable as follows [103],

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - \frac{\mu}{\sqrt{E[\boldsymbol{g}^2]_i + \epsilon}}\boldsymbol{g}_i, \tag{3.26}$$

$$E[\boldsymbol{g}^2]_i = \gamma E[\boldsymbol{g}^2]_{i-1} + (1-\gamma)\boldsymbol{g}_i^2, \quad E[\boldsymbol{g}^2]_0 = 0, \tag{3.27}$$

with $E[\boldsymbol{g}^2]_i$ as a symbolic mean of gradients of all the weights, and γ as a momentum decay parameter. By using gradients in the previous step, the iterative change of the weights stays adaptive, but no longer vanishing along with the training.

Another weight optimization approach, Adam, is often compared with RMSprop. One example is a recent patient classification study, where neural network models are

trained to identify COVID-19 cases from Computed Tomography (CT) images of the patients [104]. The Adam algorithm uses two momentum terms, elucidated as follows [105]. On each training step, there are

$$\boldsymbol{m}_i = \beta_1 \boldsymbol{m}_{i-1} + (1-\beta_1)\boldsymbol{g}_i, \quad \beta_1 \in [0,1), \tag{3.28}$$

$$\boldsymbol{v}_i = \beta_2 \boldsymbol{v}_{i-1} + (1-\beta_2)\boldsymbol{g}_i^2, \quad \beta_2 \in [0,1), \tag{3.29}$$

$$\hat{\boldsymbol{m}}_i = \frac{\boldsymbol{m}_i}{1-\beta_1^i}, \tag{3.30}$$

$$\hat{\boldsymbol{v}}_i = \frac{\boldsymbol{v}_i}{1-\beta_2^i}, \tag{3.31}$$

computed, with m_i and v_i as the biased momentum terms, and $\hat{m}_i$ and $\hat{v}_i$ as bias-corrected estimate terms. β_1 and β_2 are two exponential decay rates. Note that i in β_1^i and β_2^i denotes power. The update scheme is

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - \frac{\mu \hat{\boldsymbol{m}}_i}{\sqrt{\hat{\boldsymbol{v}}_i} + \epsilon}. \tag{3.32}$$

3.5. CONDITIONAL GENERATIVE ADVERSARIAL NETS (CGAN)

Initially put forward by Goodfellow et al. [106], Generative Adversarial Networks (GANs) is a framework in which two neural networks compete against each other. More specifically, one is a generator producing counterfeit samples. The other is a discriminator that judges how likely a provided sample is real.

This framework has been proven capable of generating uni-dimensional data series. To give an example, Wang et al. [107] used GANs to learn from classified weather types and subsequently augment the training data for some uncommon weather types. The framework was capable of producing different realistic samples for the same weather type. Another example is the GANs-based electricity load generation model by Wang et al. [63] that produced residuals upon point forecasts to form different load scenarios. Given these latest successful cases, and that LEPP is a generative task, it can be confidently assumed that GANs would be a suitable method for generating DAM price time series in this study.

For this study, inputs to both generator and discriminator include conditions of the sample, the features that will be selected for generating price series. The latent variables provide the source of randomness to the generator, in order to produce changeable fake samples each time. This is according to the initial concept of conditional GANs (cGAN) framework. Figure 3.4 illustrates the functions of these two networks. Note that this block diagram intends to show the input and output of the networks. The specific training cycle will be explained later.

To explain the framework in detail, let us denote by $\boldsymbol{\zeta}$ the latent variables, or the noise vector, $\boldsymbol{C}$ as the vector of conditions, D, G, $\boldsymbol{\theta}_D$, and $\boldsymbol{\theta}_G$ as the discriminator & generator neural networks and theirs weights, respectively. The noise is usually set with a standard normal distribution

$$\boldsymbol{\zeta} \sim N(0,1). \tag{3.33}$$

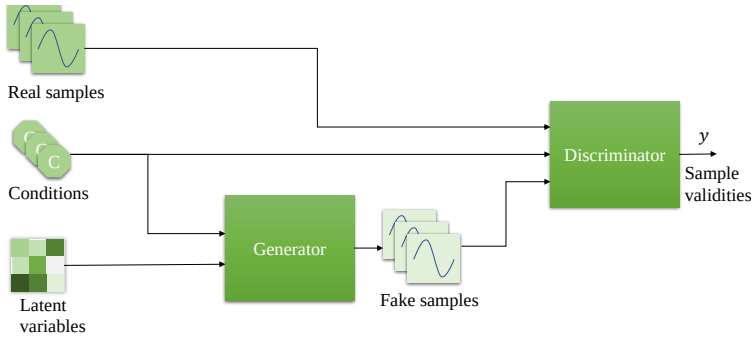


Figure 3.4: The GANs framework including the conditions as input.

Let us reiterate the roles of the two networks. The generator produces fake samples

$$\boldsymbol{s}_g = G(\boldsymbol{\zeta}, \boldsymbol{C}; \boldsymbol{\theta}_G), \tag{3.34}$$

where $\boldsymbol{s}_g$ is the generated samples. And the discriminator evaluates the validity of the samples

$$v = D(\boldsymbol{s}, \boldsymbol{C}; \boldsymbol{\theta}_D), \quad v \in [0,1], \tag{3.35}$$

where v is the validity of sample $\boldsymbol{s}$. The loss of the discriminator depends on the scores it gives to real and fake samples,

$$L_D = -\mathbb{E}_{\boldsymbol{s}_r}[log(D(\boldsymbol{s}_r, \boldsymbol{C}; \boldsymbol{\theta}_D))] - \mathbb{E}_{\boldsymbol{\zeta}}[log(1 - D(\boldsymbol{s}_g, \boldsymbol{C}; \boldsymbol{\theta}_D))], \tag{3.36}$$

where $\boldsymbol{s}_r$ is the real samples. The first term is the negative average logarithmic scores the discriminator gives to all real samples, which decreases when the scores are higher. The second term is the negative average logarithmic difference between unity and the scores that the discriminator gives to all generated samples, which decreases when the scores are lower. The loss function of the generator includes output from the discriminator, given as

$$L_G = \mathbb{E}_{\boldsymbol{\zeta}}[log(1 - D(G(\boldsymbol{\zeta}, \boldsymbol{C}; \boldsymbol{\theta}_G), \boldsymbol{C}; \boldsymbol{\theta}_D))]. \tag{3.37}$$

There is only one term in generator loss that decreases when the generated samples receive higher scores from the discriminator. Logarithmic functions are in both losses because they are originated from the idea of cross-entropy [106].

The generator and the discriminator get trained in cycles. With the loss functions above, each cycle can now be seen as a step towards solving the following optimization problem,

$$\min_{\boldsymbol{\theta}_G} \max_{\boldsymbol{\theta}_D} \mathbb{E}_{\boldsymbol{s}_r}[log(D(\boldsymbol{s}_r, \boldsymbol{C}; \boldsymbol{\theta}_D))] + \mathbb{E}_{\boldsymbol{\zeta}}[log(1 - D(G(\boldsymbol{\zeta}, \boldsymbol{C}; \boldsymbol{\theta}_G), \boldsymbol{C}; \boldsymbol{\theta}_D))]. \tag{3.38}$$

3.6. IMPROVEMENTS OF THE GAN (CGAN) FRAMEWORK

In practice, there are several issues with the cGAN and GAN frameworks regarding the training stability, mainly on the exploding/vanishing gradient of the loss function with

respect to the weights [108]. A set of improvements as follows has been published, formulating the conditional Wasserstein Generative Adversarial Nets with Gradient Penalty (cWGAN-GP) [109].

1. LeakyReLU is applied as activation functions in all layers of the two networks, instead of the $Sigmoid$ function.

2. The logarithmic function has an issue that the gradient is infinite at 0. Thus it is removed from the losses. And the optimization problem is

$$\min_{\boldsymbol{\theta}_G}\max_{\boldsymbol{\theta}_D}\mathbb{E}_{\boldsymbol{s}_r}[D(\boldsymbol{s}_r,\boldsymbol{C};\boldsymbol{\theta}_D)]+\mathbb{E}_{\boldsymbol{\zeta}}[1-D(G(\boldsymbol{\zeta},\boldsymbol{C};\boldsymbol{\theta}_G),\boldsymbol{C};\boldsymbol{\theta}_D)]. \tag{3.39}$$

3. A gradient penalty can be applied as

$$GP=\lambda\mathbb{E}_{\hat{\boldsymbol{s}}}[\|\nabla_{\hat{\boldsymbol{s}}}D(\hat{\boldsymbol{s}},\boldsymbol{C};\boldsymbol{\theta}_D)\|_2-1]^2, \tag{3.40}$$

$$\hat{\boldsymbol{s}}=\rho\boldsymbol{s}_g+(1-\rho)\boldsymbol{s}_r, \tag{3.41}$$

$$\rho\sim U[0,1], \tag{3.42}$$

where λ is a hyper-parameter deciding the magnitude of the gradient penalty, $\boldsymbol{s}_g$ and $\boldsymbol{s}_r$ here are real and generated samples for the same condition $\boldsymbol{C}$ values, and ρ is a standard uniform-distributed random variable.

The optimization problem is finally

$$\min_{\boldsymbol{\theta}_G}\max_{\boldsymbol{\theta}_D}\mathbb{E}_{\boldsymbol{s}_r}[D(\boldsymbol{s}_r,\boldsymbol{C};\boldsymbol{\theta}_D)]+\mathbb{E}_{\boldsymbol{\zeta}}[1-D(G(\boldsymbol{\zeta},\boldsymbol{C};\boldsymbol{\theta}_G),\boldsymbol{C};\boldsymbol{\theta}_D)]-GP. \tag{3.43}$$

And the pseudo-code of the cWGAN-GP framework can be found as follows, adapted from [109]. Note that the optimizer and the weight initialisation approach is not specified here. Candidate optimizers are RMSprop and Adam optimization approaches introduced in Section 3.4. The optimizer calculates the updated weights from the gradients

of the loss with respect to the weights, as shown in Eq. (3.18).

Algorithm 3.6.1: cWGAN-GP($\boldsymbol{\theta}_G, \boldsymbol{\theta}_D$)

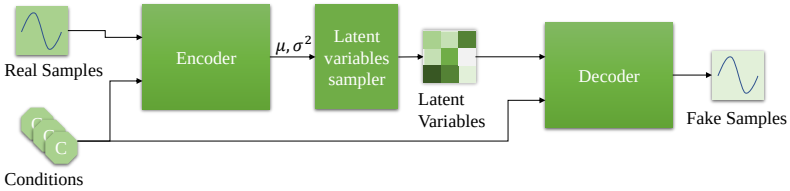
Require: λ : gradient penalty coefficient.
Require: n_D, n_G : no. of training on the networks per cycle.
Require: m : batch size, α : learning rate of the optimizer.
Randomly initialize $\boldsymbol{\theta}_G, \boldsymbol{\theta}_D$
while $\boldsymbol{\theta}_G, \boldsymbol{\theta}_D$ have not converged **do**
  for $t = 1, 2, ..., n_D$ **do**
    for $i = 1, 2, ..., m$ **do**
      Sample real data $\boldsymbol{s}_r, \boldsymbol{C}$, latent variables $\boldsymbol{\zeta}$, and random variable ρ.
      $\boldsymbol{s}_g \leftarrow G(\boldsymbol{\zeta}, \boldsymbol{C}; \boldsymbol{\theta}_G)$
      $\hat{\boldsymbol{s}} \leftarrow \rho \boldsymbol{s}_g + (1-\rho)\boldsymbol{s}_r$
      $L_D^{(i)} \leftarrow D(\boldsymbol{s}_g, \boldsymbol{C}) - D(\boldsymbol{s}_r, \boldsymbol{C}) + \lambda(\|\nabla_{\hat{\boldsymbol{s}}} D(\hat{\boldsymbol{s}}, \boldsymbol{C}; \boldsymbol{\theta}_D)\|_2 - 1)^2$
    end for
    $\boldsymbol{\theta}_D \leftarrow \text{optimizer}(\nabla_{\boldsymbol{\theta}_D} \frac{1}{m} \sum_{i=1}^{m} L_D^{(i)}; \alpha)$
  end for
  for $t = 1, 2, ..., n_G$ **do**
    Sample m latent variables $\boldsymbol{\zeta}^{(i)}, i = 1, 2, ..., m$
    $\boldsymbol{\theta}_G \leftarrow \text{optimizer}(\nabla_{\boldsymbol{\theta}_G} \frac{1}{m} \sum_{i=1}^{m} -D(G(\boldsymbol{\zeta}^{(i)}, \boldsymbol{C}; \boldsymbol{\theta}_G), \boldsymbol{C}; \boldsymbol{\theta}_D); \alpha)$
  end for
end while

3.7. OTHER POTENTIAL METHODS

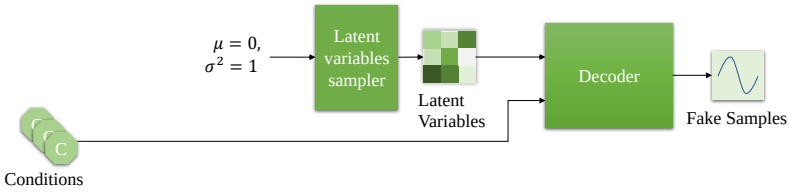
Another network-based generative method based on neural networks is the Variational Auto-Encoder (VAE), which is often brought up with the GAN [106]. Similarly, conditions can be included as input to the network, resulting in the Conditional VAE (CVAE) framework. The CVAE functions differently in training and sample generation. During network training, as in Figure 3.5a, the latent variables are sampled from mean and variance values designated by the encoder. And the decoder takes the sampled latent variables and conditions to generate fake samples [110, 111]. When generating samples, the CVAE framework works without real samples, as in Figure 3.5b. The latent variables are sampled from the standard normal distribution so that samples with the learned general interrelationship of the sample and conditional dimensions can be generated. Conditions are given to the decoder with the latent variables. Note that the fake samples generated is not criticised by any means in the CVAE framework, unlike the case of cGAN.

Moreover, the multivariate Gaussian distribution (mGaus) is taken as an example of statistical methods. In this method, it is assumed that the prices and the conditions distribute normally, given as

$$\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \tag{3.44}$$



(a) In training, the latent variables are sampled from distributions designated by the encoder.



(b) In sample generation, the latent variables are sampled from the standard normal distribution.

Figure 3.5: The CVAE framework.

where

$$\boldsymbol{X} = [\boldsymbol{P}, \boldsymbol{C}]. \tag{3.45}$$

$\boldsymbol{P}$ and $\boldsymbol{C}$ are price and condition vectors, $\boldsymbol{\mu}$ the mean of each dimension in $\boldsymbol{X}$, and $\boldsymbol{\Sigma}$ the co-variance matrix of the values on the dimensions,

$$\boldsymbol{\mu} = [\boldsymbol{\mu_P}, \boldsymbol{\mu_C}], \tag{3.46}$$

$$\boldsymbol{\Sigma} = [\boldsymbol{\Sigma_{PP}}\ \boldsymbol{\Sigma_{PC}}, \boldsymbol{\Sigma_{CP}}\ \boldsymbol{\Sigma_{CC}}]. \tag{3.47}$$

To generate samples from the conditions, the price vector, as put by Eaton [112], is

$$(\boldsymbol{P}|\boldsymbol{C} = \boldsymbol{c}) \sim \mathcal{N}(\overline{\boldsymbol{\mu}}_{\boldsymbol{P}}, \overline{\boldsymbol{\Sigma}}_{\boldsymbol{P}}), \tag{3.48}$$

where $\overline{\boldsymbol{\mu}}_{\boldsymbol{P}}$ and $\overline{\boldsymbol{\Sigma}}_{\boldsymbol{P}}$ are conditional mean and covariance matrix of the price $\boldsymbol{P}$, calculated by

$$\overline{\boldsymbol{\mu}}_{\boldsymbol{P}} = \boldsymbol{\mu_P} + \boldsymbol{\Sigma_{PC}}\boldsymbol{\Sigma}_{\boldsymbol{CC}}^{-1}(\boldsymbol{c} - \boldsymbol{\mu_C}), \tag{3.49}$$

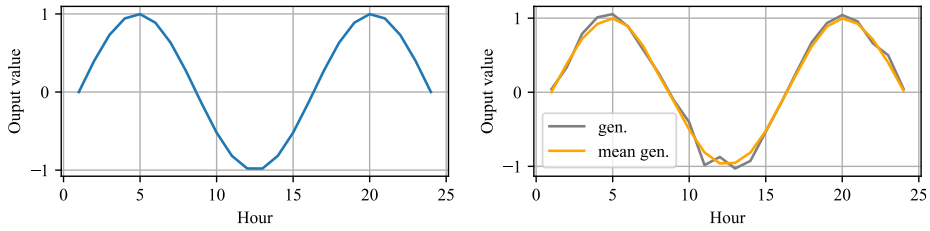
$$\overline{\boldsymbol{\Sigma}}_{\boldsymbol{P}} = \boldsymbol{\Sigma_{PP}} - \boldsymbol{\Sigma_{PC}}\boldsymbol{\Sigma}_{\boldsymbol{CC}}^{-1}\boldsymbol{\Sigma_C P}. \tag{3.50}$$

3.8. TEST CASE ON REPRODUCING SINUSOIDAL CURVES

To have a heuristic check on feasibility of the introduced cWGAN-GP, CVAE and mGaus methods, a test case of reproducing sinusoidal curves with different periodic amplitudes and levels of Gaussian noise. The reason not to test directly on the price data is the difficulty in evaluate the generated series with the irregular-shaped price samples. By testing with this case, standard metrics of signal similarity or even visual examination can be used to basically evaluate the model behavior. The sinusoidal curve to be reproduced is given by

$$y_h = A\sin(\frac{3\pi}{24}h) + noise_h, \quad h = 1, 2, ..., 24, \quad noise \sim \mathcal{N}(0, v), \tag{3.51}$$

where A is the amplitude of the sinusoidal wave, and v the standard deviation of the Gaussian noise. A24 outputs represent the price values of each hour in a day. The factor of h is set as $3\pi/24$ such that the curve has two peaks and one valley during the 24 hours typically in daily DAM price curves. An example can be seen in Figure 3.6a.



(a) An example of target sine curves in the test case, $A = 1$, $v = 0$.

(b) One sample and the mean sample generated by the cWGAN-GP model, $A = 1$, $v = 0.025$.

Figure 3.6: Example of target curve and generated samples in the test case.

Table 3.1: Conditions in the test case and their variable spaces.

Condition	Variable space
A	$\{0.7, 0.85, 1\}$
v	$\{0.025, 0.075\}$

The output values in Equation 3.51 is dependent on two conditions, the sine wave amplitude A and noise variance v, verbalising the daily and hourly price patterns. In the test case, the conditions are designed to variate in discrete values, as listed in Table 3.1. For each of the $3 \times 2 = 6$ combinations, 100 samples are generated to training the networks with.

The evaluation of generated samples consists of two scores, one for the curve, and the other for the noise. For each combination of conditions, the network generate $S = 1000$ samples. In sample s, with condition combination c, and hour h, the generated value is $\hat{y}^s_{c,h}$. The mean generated curve of a condition combination is

$$\overline{y}_{c,h} = \frac{1}{S}\sum_{s=1}^{S} \hat{y}^s_{c,h}. \tag{3.52}$$

The mean curve is compared with the sine wave from Equation 3.6a, using Dynamic Time Warping (DTW). DTW is a measure of similarity between two sequences with certain (two sample steps) allowance for the phase shift [113]. This is more suitable when evaluating with emphases on the signal shape and inter-dependencies on the sample steps. In our case, these are preferred over reproducing the exact timing. Hence, DTW is adopted as the curvature score. Firstly the DTW cost function is defined as

$$c : F \times F, \tag{3.53}$$

where F is the feature space for two samples x and y, such that c is smaller when x and y are similar, and larger when opposite. The samples x and y has dimensions of N and

M, respectively. A warping path $p_l = (n_l, m_l) \in [1:N] \times [1:M]$ for $l \in [1:L]$, $L = M + N$ is optimized with the following constraints,

1. Boundaries: $p_1 = (1,1)$, $p_l = (N, M)$.
2. Forward rule: $n_1 \leq n_2 \leq ... \leq n_L$, and $m_1 \leq m_2 \leq ... \leq m_L$.
3. Step warping rule: $p_{l+1} - p_l \in \{(1,0),(0,1),(1,1)\}$, for $l = 1,2,...,L-1$.

The total cost of the warping path is given by

$$c_p(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{l=1}^{L} D(x_{n_l}, y_{m_l}), \tag{3.54}$$

where D is the distance between data points x_{n_l} and y_{n_l}. For instance, it can be the Euclidean distance. Finally the DTW distance is given by optimizing Equation 3.54.

$$DTW(\boldsymbol{X}, \boldsymbol{Y}) = c_{p^*}(\boldsymbol{X}, \boldsymbol{Y}). \tag{3.55}$$

where c_{p^*} is the solution to the following optimization problem,

$$\min_p c_p(\boldsymbol{X}, \boldsymbol{Y}) \tag{3.56}$$

It can be seen from the DTW principles above, this metric is able to compare sequences with different speed and phases, in contrast to the point to point measurement by the Euclidean distance. Same with the Euclidean distance, the DTW measurement values are in unit aligned with the measured data. The curvature score is

$$k_{curve,n,c} = DTW(\boldsymbol{y}_{n,c}, \overline{\boldsymbol{y}}_{n,c}), \quad n = 1,2,...,N, \tag{3.57}$$

$$\boldsymbol{y}_{n,c} = [y_{n,c,1}, y_{n,c,2}, ..., y_{n,c,24}], \quad \overline{\boldsymbol{y}}_c = [\overline{y}_{n,c,1}, \overline{y}_{n,c,2}, ..., \overline{y}_{n,c,24}]. \tag{3.58}$$

where $k_{curve,c}$ is the curve score of the nth among N generated samples on condition combination c.

Noise of the generated samples is evaluated by taking their residuals from the mean curve, given by

$$r^s_{n,c,h} = y^s_{n,c,h} - \overline{y}_{c,h}. \tag{3.59}$$

And the standard deviation of the generated samples for each c is

$$\hat{v}_{n,c} = std(r^s_{n,c,h})\Big|_{c=c}, \quad h = 1,2,...,24, \quad s = 1,2,...,S, \tag{3.60}$$

where std denotes standard deviation. The noise score is the absolute difference between the variance in conditions c and the measured value

$$k_{noise,n,c} = |v_c - \hat{v}_c|. \tag{3.61}$$

Finally, the evaluation scores of the curve and noise are given as the root mean square of the scores for all samples for all conditions

$$K_{curve} = RMS(k_{curve,n,c}) \tag{3.62}$$

$$K_{noise} = RMS(k_{noise,n,c}) \tag{3.63}$$

where RMS stands for the root mean square function. The RMS is taken to penalise for particular high scores. Regarding both scores, the lower the score value, the better is the model being evaluated.

To implement the methods, a default set of parameters need to be determined for each method. For cWGAN-GP, a suitable reference is the work by Wang et al. [63], which projects electricity daily load series with the same framework. In their study, the number of neurons in the hidden layer is set to 256. That of the output layer is 24, as the target is hourly load trace in a day. The hyper-parameter λ for gradient penalty is set to 10. And the generator G is trained once after the discriminator D has been trained for five times. According to this reference, the default set of parameters for the cWGAN-GP framework is determined, listed in Table 3.2. The RSF is selected as the default market state indicator for its favorable effectiveness in market state normalisation among all the indicators considered in this study, as visible in Figure 3.5. The layers here denote the hidden layers in the neural network. And each layer has the same amount of neurons. Dimensions of the hidden parts are identical for both the discriminator and the generator networks.

Table 3.2: Default parameters of the cWGAN-GP method for the test case.

Setting	Value
Market state indicator	RSF
Training time	30 minutes
Optimizer	RMSprop
Number of neuron layers	2
Number of neurons per layer	256
Training batch size	140
Number discriminator train(s) per cycle	5
Number of generator train(s) per cycle	1
Initial learning rate of discriminator optimizer	0.001
Initial learning rate of generator optimizer	0.001
GP magnitude parameter λ	10
Activation function in hidden layers of discriminator	LeakyReLU
Activation function in hidden layers of generator	LeakyReLU
Activation function in the output layer of discriminator	Linear
Activation function in the output layer of generator	Sigmoid

The neural networks used to build CVAE models are created with parameters aligned with the cWGAN-GP models. Both encoder and decoder networks have two layers, containing 128 neurons each. And they are optimized with RMSprop. The dimension of latent variables is selected as 12, halved the number of hours in a day. This is chosen because the latent dimension should be smaller than that of the price series, in order to avoid direct transmission of the price values as latent means. And with fewer latent dimensions than the price series, the encoder and the decoder are obliged to extract the underlying (inter-dimensional) relationships.

Models trained or fitted with the cWGAN-GP, CVAE and mGaus methods are able to

reproduce the sine curve and the noise. Figure 3.6b shows one sample generated by the cWGAN-GP model. The sample follows the sinusoidal curve in shape, and a superposed noise can be noticed. For each combination of the conditions, 1000 samples are generated by each model. Evaluating the generated samples, mean of the curvature score K_{curve} and noise score K_{noise} over all conditions can be found in Table 3.3. The multivariate Gaussian model is exceptionally better than the others on the curve score. And the cWGAN-GP model is the worst in both scores. While the samples generated by the cWGAN-GP model are visibly good enough in Figure 3.6b, it can be said that all the three models are proven feasible at least for this basic task.

Table 3.3: Mean of the curvature and noise scores of samples generated by the models, over all conditions.

Method	K_{curve}	K_{noise}
cWGAN-GP	0.53	0.32
CVAE	0.36	0.026
Multivariate Gaussian	0.085	0.026

4

Model Training, Evaluation and Projection

In this chapter, practical issues in model evaluation, training, and projection are dealt with, enabling the LEPP to the Dutch 2050 energy scenarios. The evaluation metrics measure the model performance in two aspects, price value distribution and hourly inter-dependencies. The price value distribution is assessed visually with the Quantile-to-Quantile (Q-Q) plot, and numerically with the Kolmogorov-Smirnov (K-S) distance. The hourly inter-dependencies are appraised by Auto-Encoder (AE) tester trained to regenerate DAM price series in the data set. The appraisal is on how well the AE tester can regenerate the input sample, indicating the similarity of hourly inter-dependencies. About the model training, splitting the data set for training and testing, as well as the weight initialisation process, is discussed. How the discriminator and generator losses behave during model training is explained. And search spaces of the model parameters are put forward for optimisation. Multiple models are trained preliminarily to investigate the choice of split random state and select a promising combination of parameters. Finally, to prepare the conditions for the projection, the reasonings behind taking the annual mean of the market state indicator as the value on every hour, using the five-year mean daily temperature and selecting certain fuel prices for the most frequent final generating technology are vindicated.

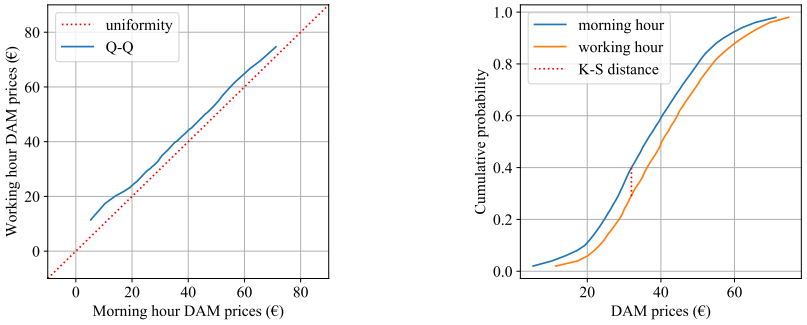
4.1. Evaluation Metrics

Most machine learning methods train models directly on the chosen evaluation metric. As the generator loss depends on training of the discriminator, no explicit performance of the model is available within the cWGAN-GP training procedure. Whereas we regard a prerequisite to LEPP that the model should be solid on **regenerating the test set**. Therefore, a set of evaluation procedures are put forward to examine DAM prices generated with the conditions of the test samples on overall and conditional price distributions, as well as hourly price inter-dependencies. To evaluate the generative models, number of

generated samples should be much larger than those in the test set. In this study, 100 samples will be generated for each test sample.

The tests used in the evaluation procedure include Quantile-Quantile (Q-Q) plots that visualises the relationship between distributions of the training and generated sample set, Kolmogorov-Smirnov (K-S) distance finding to what extent the two distributions are similar, and Auto-Encoders (AE) to reveal the conformity on hourly price inter-dependencies. The visualisation of Q-Q plots provides straightforward graphical comparisons of the distributions of two sample sets. By showing the corresponding values for the same quantiles, Q-Q plots display the conformity of two distributions on the whole range [114]. Since the Q-Q plot is a graphical metric, it is used only when detailed examination is necessary. A non-parametric goodness-of-fit metric, the KS distance, is often used to evaluate how similar two distributions are [115, 116, 117]. The distance is given by taking the maximum absolute difference in the cumulative distribution plot of the two distribution.

To demonstrate the metrics on price value distribution, an example case of comparing DAM prices in morning hours (2-9 a.m.) and working hours (10-17) is given as follows. The Q-Q plot and K-S distance of the DAM price distributions in the two categories are shown in Figure 4.1. From the Q-Q plot in Figure 4.1a, it can be seen that on every quantile, the working hour DAM price is higher than that of the morning hour. By comparing with the uniformity (diagonal) line, it can be seen that the difference peaks at the lower price region. The K-S distance is derived from the cumulative probability plot. The K-S distance is the maximal distance in probability for a certain variable value, as shown in Figure 4.1b. Note that the K-S distance is non-dimensional, therefore suitable for evaluating distributions of two variable in the same scale. Aside from the variable values, the significance of the K-S distance also depends on the number of data points. In this study, the test set have the same size, 20% of the combined data set. Thus, the K-S distance can be used for model evaluation in this study.



(a) An example of the Q-Q plot.

(b) An example of the K-S distance.

Figure 4.1: Q-Q plot and K-S distance of an example case showing distributions of DAM prices in morning hours (2-9) and working hours (10-17).

It has been shown that Auto-Encoders (AE) are capable of learning the dimensional inter-dependencies of data series [118]. Unlike VAEs, AEs do not have variable latent codes, as illustrated in Figure 4.2. In this framework, each input sample is fed to the

encoder that gives out latent codes. And the latent codes are taken by the decoder to produce the output sample as a regeneration of the input sample. It is expected that the regenerated sample should be as close to the real one as possible with no phase shift wanted. Hence the reconstruction error, Euclidean distance between each input and output sample, is used here as the training loss, given as

$$E_{AE} = \sqrt{\sum_{i=1}^{24} (x_{in,i} - x_{out,i})^2} \tag{4.1}$$

where $x_{in,i}$ and $x_{out,i}$ are the input and output prices of hour i in a DAM price series. In the training, the loss is taken as the mean over the training batch. And the AE score is the E_{AE} mean over the test set. This metric, however, is with the unit same with the sample value, €. Sample values might affect the evaluation. For instance, a test set with relatively smaller values might lead to lesser reconstruction error, given the same embodied inter-dependency level. Consequently, this raise a requirement to the choice of test set that it should be ordinary enough to represent the whole data set. In response, the split of training and test sets will be repeated and examined. Generally, the lower the reconstruction error, the more similar the dimensional inter-dependencies of the sample series is to those in the combined data set.

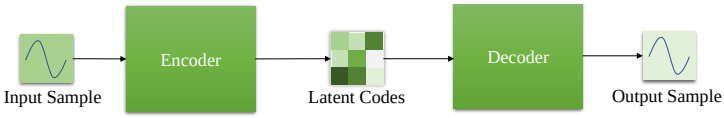


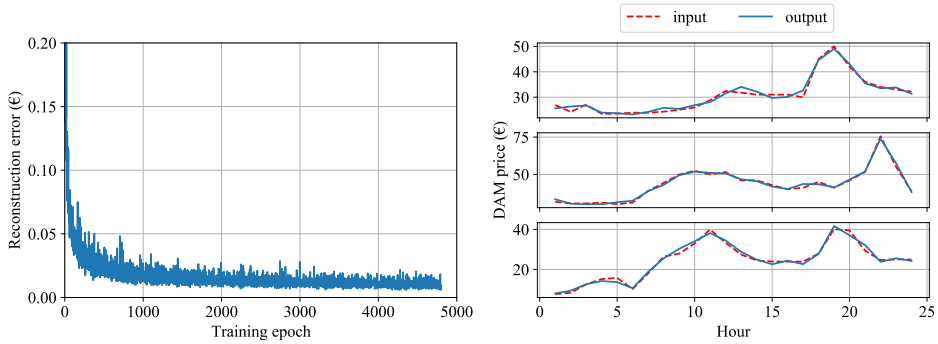
Figure 4.2: The AE framework.

Table 4.1: Parameters of the AE tester model. Layer(s) here denotes hidden layer(s). Hidden parts of the encoder and decoder have the same size. And each layer has the same amount of neurons.

Setting	Value
Training time	10 minutes
Optimizer	RMSprop
Number of neuron layers	2
Number of neurons per layer	128
Training batch size	140
Initial learning rate of the optimizer	0.001
Activation function in hidden layers of encoder and decoder	LeakyReLU
Activation function in the output layer of encoder and decoder	Linear

To train the encoder and decoder, the parameters in Table 4.1 are adopted. And the model is trained on all price series in the combined data set. As it is a tester model, the parameters do not need to be optimised, as long as the regeneration of the input sample is acceptable. As can be seen in Figure 4.3a, the training of the AE tester model is successful, with the reconstruction error converging towards zero. And Figure 4.3b

demonstrates three DAM price series regeneration of different price patterns and scales. Visibly, the regeneration (output samples) acceptably resembles the input samples.



(a) Reconstruction error during training.

(b) Three examples of DAM price series regeneration.

Figure 4.3: Training and regeneration examples of the AE tester model.

4.2. THE TRAINING PROCEDURE

To train the cWGAN-GP model, several details in the training procedure need to be specified. One is splitting the data set for training and testing. A standard approach in evaluating time series predictor is using k-fold cross-validation to fully exploit the data [119]. The k-fold cross-validation splits the data set into k sections and each time using one of the sections as the test set, the remaining as the training set. In this study, however, the one-time split is adopted for two reasons. Firstly, the five-year data set of four DAMs is sufficiently large. If k-fold cross-validation is chosen, though all of the data set participates in training, the large data set would be used for k-1 times, incurring unnecessary computational loads. Secondly, unlike other prediction tasks, similar value distribution and reliable dimensional inter-dependency are desirable, rather than small point-to-point errors. Hence the choice of training and test sets shall be carefully examined, instead of leaving it to an automated process. Specifically, the choice should be ordinary that no particular good or bad performance is presented as a consequence of over-fitting or under-fitting.

Another issue in training is the weight initialisation process, mentioned in Section 3.6. In our models, neurons in adjacent layers are densely connected. And before training, the neuron-to-neuron weights are initialised with the uniform Glorot initialiser and the bias weights are set as zeros, as default from the *Keras* library in *Python* [120]. The uniform Glorot initialiser samples random values from a uniform distribution between $[-L_{glorot}, L_{glorot}]$. The magnitude L_{glorot} is given as

$$L_{glorot} = \sqrt{6/(fan_{in} + fan_{out})}, \tag{4.2}$$

where fan_{in} and fan_{out} are numbers of dimensions in layer input and output, respectively. Other details unspecified follow standard approaches in the *Keras* and *Tensorflow* libraries.

After splitting the randomly reordered data set into 80% training set and 20% test set and initialising the weights, a cWGAN-GP model with default parameters in Table 3.2 is built and trained using packages *Keras* and *Tensorflow* in *Python* (CPU: AMD Ryzen 4800H; RAM: 16Gb). The losses returned by the discriminator and generator during the training are shown in Figure 4.4. Note that the mean of every five epochs is taken as the generator loss in the Figure for synchronised comparison. The large discriminator loss at training epoch 0 is due to the initial value of the gradient penalty from Eq. (3.10) that the discriminator is temporarily unable to recognise any difference between real and fake samples. The adversarial nature of the framework can be observed in the contrary trend that when the generator loss decreases, the discriminator loss increases and vice versa. After epoch 500, the two losses gradually converge to steady levels, indicating stable learning progresses of the two networks.

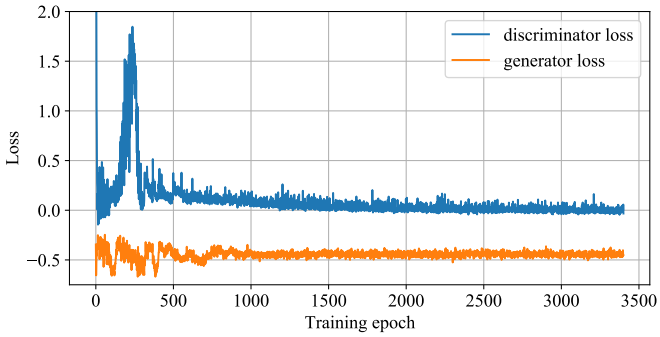


Figure 4.4: Discriminator and generator losses of the default cWGAN-GP model during training.

The given 30-minute is enough to train an acceptable model. As presented in Figure 4.5, the model performance improves considerably till epoch 500. After this, the K-S distance fluctuates between 0.2 and 0.5. And the AE score drops slowly. With the dwindled improvement rates, an extra unit of training time only benefits little.

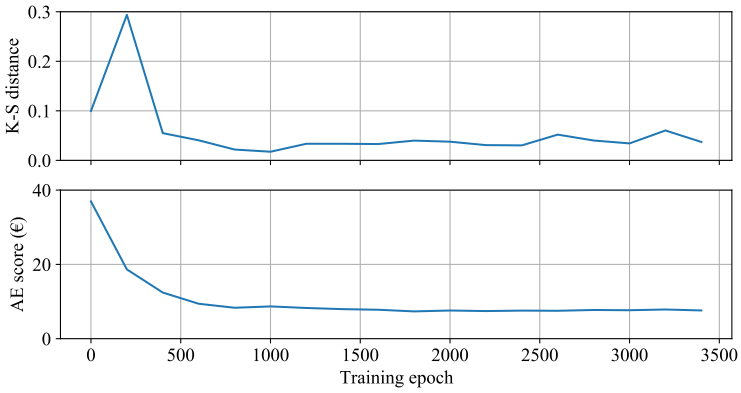


Figure 4.5: Model performance of the default cWGAN-GP model during training.

The default parameters are to be optimised for price projection in this study, especially the parameters that are directly relevant to the model's learning ability. Three of such parameters are selected to be optimised, namely the optimizer, number of layers in the model and number of neurons per layer. Reasonable search spaces of the three parameters are drawn in Table 4.2. The best combination of parameters will be used in further investigations.

Table 4.2: Parameter search space for the cWGAN-GP model.

Parameter	Search space
Optimizer	{RMSprop, Adam}
Number of neuron layers	{1,2,3}
Number of neurons per layer	{64, 128, 256, 512}

To address the aforementioned issues on splitting the data set and parameter search, multiple models have been trained preliminarily and evaluated with the proposed metrics. The random state is a parameter in the *train_test_split* function in the *sklearn* library, which is used in this study. Although the function splits the data set randomly, designating the random state makes the function produces the same training and test sets every time. Performance of models by different training sets is shown in Table 4.3. A normal distribution that has the same mean and variance for each hour in the whole data set is taken as a reference, whose K-S distance is taken by comparing against random state 0 test set. As a sanity check, it can be seen that the normal distribution has the best value distribution but much worse in the AE score. This is supposedly due to the independent sampling of DAM price values for every hour. And the AE tester does not recognise the series, resulting in high reconstruction errors. As for the random state, it does influence the model performance. From the five experiments in the table, 0 seems to be an ordinary value, which is favorable to avoid both over-fitting and under-fitting.

Table 4.3: Model performance under different random state of splitting the data set. Parameters are default as in Table 3.2. A normal distribution on each hour with corresponding mean and variance from the data set is included as a reference.

Random state	K-S distance	AE score (€)
Normal distribution	0.022	55.29
0	0.038	7.69
1	0.040	7.46
2	0.046	7.67
3	0.043	7.49
4	0.024	7.58

In the parameter search, models are trained with all possible combinations of parameters from the search spaces in Table 4.2. Top 5 performing models on the K-S distance and the AE score are displayed in Table 4.4 and 4.5, respectively. Colored in the tables, two combinations gives high performance on both metrics. Surprisingly, the simplest network structure with only one layer of 64 neurons performs similarly to the default

two layers of 128 neurons. This could mean that the bigger model is still quite potential after being trained for 30 minutes. For this reason and the similar performance, the two layers of 128 neurons is selected as the network structure for LEPP under Dutch 2050 energy scenarios.

Table 4.4: Top 5 parameter combinations by the K-S distance. Random state = 0.

Optimizer	No. neurons	No. layers	K-S distance	AE score (€)
RMSprop	64	3	0.026	8.06
Adam	256	1	0.026	11.04
RMSprop	64	1	0.035	7.59
RMSprop	64	2	0.036	7.93
RMSprop	128	2	0.037	7.60

Table 4.5: Top 5 parameter combinations by the AE score. Random state = 0.

Optimizer	No. neurons	No. layers	K-S distance	AE score (€)
Adam	64	2	0.051	7.42
RMSprop	64	1	0.035	7.59
RMSprop	128	2	0.037	7.60
RMSprop	64	2	0.036	7.93
RMSprop	128	1	0.054	8.00

4.3. CONDITIONS FOR PROJECTION UNDER DUTCH 2050 ENERGY SCENARIOS

Prior to the price projection, it is imperative to determine the conditions under Dutch 2050 energy scenarios. Viewing from the power capacity and generation portfolios in Figure 1.3a and 1.3b, as well as the features required in Table 2.6, three conditions are still pending, viz. hourly market state indicator values, daily national mean temperature and fuel cost of the most frequent final generating technology. As the features of month of the year and day of the week are expected to handle the DAM price seasonalities, with the cWGAN-GP model learning the hourly DAM price inter-dependencies, variation of the market state indicator is considered unnecessary in price projection. Hence, the annual mean values in Table 2.4 will be taken for every hour in 2050. Nonetheless, the corresponding 24 dimensions conserve the feasibility to study variable renewable generation-to-load relationship within a day. In terms of the temperature, Although the climate change is undoubtedly taking place, the interest of this study lies more in the renewable penetration. For this reason, the daily Dutch national mean temperature in 2050 takes the mean over 2015-2019 on each day of the year.

The fuel cost of the most frequent final generating technology in 2050 needs to be chosen cautiously. The main difficulty is that hydrogen power generation substantially present in the RM, NM and IM scenarios has never been put into large operation. Regarding the fuel, two hydrogen production technologies are envisaged for the carbon-

neutral future, green and blue hydrogen [121]. The green hydrogen refers to hydrogen synthesised with electricity from renewable energy sources, which is ergo naturally "carbon free". While the blue hydrogen technologies focus on capturing the carbon in the exhaust gas from the hydrogen production with fossil fuels, typically coal or natural gas. Among these two technologies, blue hydrogen falls behind in terms of eliminating carbon emission, as neither the efficiency of carbon capture is 100% nor leakage of carbon storage can always be prevented. For this reason, green hydrogen has been anticipated as the most promising future of hydrogen production and carrier of energy generated from renewables, e.g. by [122] and [123].

Even to green hydrogen, costs in hydrogen synthesis are not negligible. For instance, the use of water in hydrogen synthesis would be significant. A study has shown that in scenario RM, hydrogen technologies can take more than 30% in the total water consumption for heat and electricity [124]. Furthermore, several studies have concluded cost estimates of hydrogen synthesis. Kayfeci et al. [125] have found that cost of hydrogen synthesis using solar and wind power would be no less than 5.8 and 5.1 $/kg. The cost, however, is highly related to the energy output from the generator, taking its costs into account. And given certain wind conditions and suitable wind turbines, a research has found that the hydrogen cost can drop to 1.4 $/kg [126]. Using the specific energy of 120 MJ/kg for hydrogen [127], and the currency rate of 1.12 $ per €, the hydrogen cost can be calculated as 0.037 €/KWh. In the data set, the feature of the fuel cost of the most frequent final generating technology (natural gas so far) ranges from 0.020 to 0.034 €/KWh. To avoid using values foreign to the data set, 0.034 €/KWh is taken as the hydrogen cost in scenarios RM, NM and IM, where hydrogen can be regarded as the most frequent final generating technology. Whereas in the EC scenario, gas still fills in. Hence this condition is set at the mean of Dutch non-household gas price from 2015 to 2019, 0.023 €/KWh. Again, the use of current and historical mean values is due to the study focus on the shift in energy portfolios, rather than the changes in fuel costs brought by policies or technological advancements.

5

Results and Discussion

With the practical issues dealt with in the previous chapter, models can be trained for LEPP under the Dutch 2050 energy scenarios. Prior to making projections, choices in market state normalisation and projection method need to be examined. This chapter starts with a comparison of model performance by different market state normalisation solutions presented in this study, where the necessity of including data from other markets is also demonstrated. Subsequently, the selected projection methods, cWGAN-GP, CVAE and mGaus are evaluated not only by the metrics proposed in Chapter 4, but also with visualisation of their sample generation under given conditions. In this evaluation, CVAE is ruled out for its immoderate price variance. Moving on to LEPP under Dutch 2050 energy scenarios, projections by the cWGAN-GP and mGaus models are analysed with two ETM-based models as references. The analyses are firstly on selected general price characteristics, including annual mean, mean daily standard deviation and AE score. By creating a new EC scenario with hydrogen as the most frequent final generating technology, instead of natural gas, we explore how this feature can influence the general price characteristics. Enhanced box-plots showing distributions of projected DAM prices under different temporal conditions (day of week and month of year) are discussed. And DAM price series projected on the third and the fourth Tuesdays in February 2050 are plotted to have a direct assessment. Moreover, to showcase how the generated price series can be utilised, generic profit potential of a energy storage system that trades in the market is investigated by designing and solving a linear optimization problem. Mean potential daily profit is shown to offer another aspect in assessing the price characteristics. Finally, the chapter ends with discussions on the presented results.

5.1. Market State Normalisation

To evaluate the effects on generated DAM price series quality by the proposed market state indicators, as well as the necessity to include data from the ES, DK1 and DK2 markets, cWGAN-GP models are trained with different data sets. Other than the combined NL, ES, DK1 and DK2 data sets with market state indicators, another is created by normalising only the Dutch data with the RSF. And using the metrics proposed in Section

4.1, the evaluation is done on the test set under split random state 0, as recommended in Section 4.2.

Firstly, distributions of the generated DAM prices (all 24 hours) are visualised in the Q-Q plot in Figure 5.1. When the cWGAN-GP model is trained only using the Dutch data and given conditions from the test set, it overestimates the DAM price, especially on the high price region. This is verified with the numerical metrics in Table 5.1 that in both the K-S distance and the AE score, the model performs much worse under foreign conditions. Like those under Dutch 2050 energy scenarios, some of the conditions from the test set are unseen in the Dutch data. Hence, given that the annual mean of market state indicators under Dutch 2050 energy scenarios are covered not in the NL but in the ES, DK1 and DK2 data, it can confirmed that inclusion of data from the ES, DK1 and DK2 markets is necessary for reliable the LEPP.

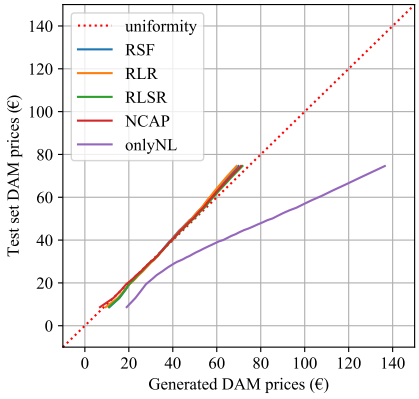


Figure 5.1: Q-Q plot of DAM prices generated by cWGAN-GP models trained with different market state normalisation solutions. "onlyNL" stands for the model trained with only Dutch data normalised with the RSF. Evaluation is on the test set, split random state = 0.

Table 5.1: cWGAN-GP model performance using different market state normalisation solutions. Evaluation is on the test set, split random state = 0.

Market state normalisation solution	**K-S distance**	**AE score (€)**
RSF	0.037	7.60
RLR	0.038	7.50
RLSR	0.019	7.80
NCAP	0.027	8.86
RSF with only NL data	0.388	15.965

Among the market state normalisation solutions, the one with the RLSR leads in K-S distance, in other words the price distribution. Examining the RLSR in Figure 5.1, slightly bigger discrepancies to the uniformity occur in the lower price region, often when the power generation from renewables is relatively high. Presumably, the discrepancies are owing to the increasing variance in the DAM price when the RLSR approaching 1, as

plotted in Figure 2.6f, which adds difficulties for the model to project DAM prices similar with the test set. Nevertheless, the better performance compared to the RLR suggests that capping the RLR at 1 helps the model learn better in the other region. This confirms the expectation that the DAM price is no longer influenced much by the renewable generation-load relationship when the RLR exceeds 1, i.e. when the demand can be satisfied by renewable power in the forecast.

Regarding the RSF, RLR and NCAP solutions, comparable performance can be observed. In Figure 5.1, distributions of DAM prices generated by these models are well in line with that in the test set. It is interesting to see that the cWGAN-GP model trained with RLR is close to the model with RSF. This implies that the additional information on total installed renewable generation capacity is redundant, or its relationship to the DAM price can be inferred from other features. The NCAP model excels in the K-S distance, while lagging in the AE score, which could mean that feeding the renewable power generation forecast and the load forecast separately does help the model predict DAM price values better. But the extra 24 dimensions hampers the model in learning hourly inter-dependencies of the DAM price series. Overall, it can be stated that all the proposed market state indicators normalise market states effectively. Prioritising the hourly inter-dependency and for more included information, RSF is chosen as the market state indicator in the subsequent analyses.

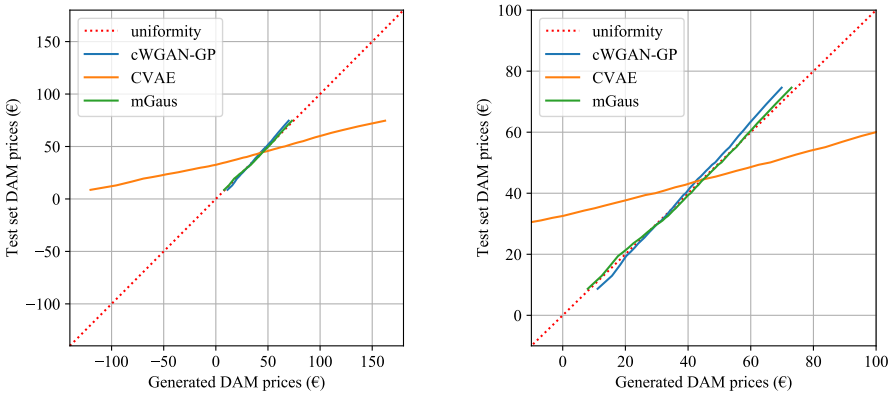
5.2. SELECTED PRICE PROJECTION METHODS

In Chapter 3, CVAE and mGaus are introduced as alternatives to cWGAN-GP. To compare the three methods, CVAE and mGaus models are created and trained on the training set with 0 as the split random state. Performance of these models can be found in Table 5.2 and Figure 5.2. From the K-S distance, it can be told that distribution of the CVAE generated prices is much incongruous with the test samples. This is also shown in the Q-Q plot in Figure 5.2a that CVAE quantiles are way off. Though intersecting the uniformity around the median, the variance is significantly amplified. The mGaus model performs well in the price distribution, with the K-S distance denoting the imperfection in assuming normal distributions. Regarding hourly inter-dependencies, the mGaus model is expected to be worse than others in the AE score as the mGaus hourly prices are independently sampled from the conditional means and variances. The CVAE model, however, is not even as good as the mGaus model, which may be explained that no reasonable inter-dependencies can be recognised from the hourly prices with immoderate variances.

Table 5.2: Model performance of different projection methods. RSF normalisation is adopted. Evaluation is on the test set, split random state = 0.

Price projection method	K-S distance	AE score (€)
cWGAN-GP	0.037	7.60
CVAE	0.406	11.93
mGaus	0.029	9.23

To give an intuitive idea, DAM prices series generated by the cWGAN-GP, CVAE and mGaus model are illustrated in Figure 5.3. Again, as in Figure 5.3a, prices in the CVAE se-



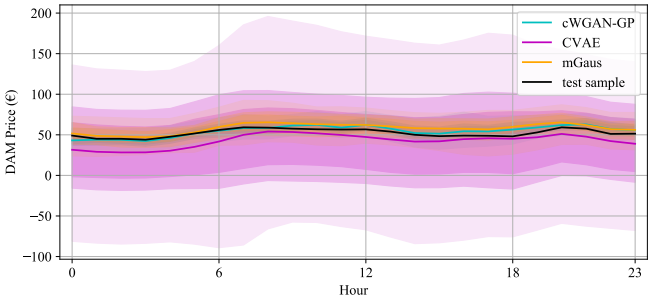
(a) All CVAE quantiles.

(b) Fit to examine the cWGAN-GP and mGaus models.

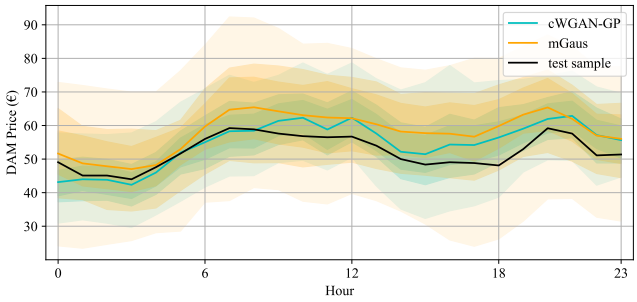
Figure 5.2: Q-Q plots of DAM prices generated with different projection methods. RSF normalisation is adopted. Evaluation is on the test set, split random state = 0.

5

ries excessively vary from less than -70 € to over 180 €. The cWGAN-GP and mGaus series can be better examined in Figure 5.3b with the CVAE series removed. On the conditions of this sample, both models generate price series close to the test sample, which is covered in the middle 50% of the cWGAN-GP series and 70% of the mGaus series. And the mean curves of both methods resemble the test sample, though the cWGAN-GP mean curve conforms better. Notably, variance of the price changes per hour, especially in the cWGAN-GP series. For instance in the figure, colored area of the cWGAN-GP series in the afternoon is wider than that in the morning. This can be explained that the cWGAN-GP model is capable of learning the relationship between price volatility and the given conditions, which is seemingly absent in the mGaus model. Moreover, colored area of the cWGAN-GP series is consistently narrower than that of the mGaus series, showing that the cWGAN-GP model projects more precisely than sampling from the conditional Gaussian distribution. Overall, the cWGAN-GP and mGaus series overlap much, and therefore will be taken in the LEPP under the Dutch 2050 energy scenarios.



(a) cWGAN-GP, CVAE and mGaus generated DAM price series.



(b) cWGAN-GP and mGaus generated DAM price series.

Figure 5.3: Generated DAM price series with conditions from a test sample. Colored areas from light to dark show 95%, 70% and 50% of the price around the median, respectively. Colored lines are hourly mean values. RSF normalisation is adopted.

5.3. PROJECTION UNDER DUTCH 2050 ENERGY SCENARIOS

Projections of the DAM price under Dutch 2050 energy scenarios can be done with conditions including the five-year mean daily temperature and gas price, an estimation of the hydrogen price from Section 4.3 and annual mean of the market state indicator in Table 2.4. In order to have references on the Dutch DAM price in 2050, the results from the Energy Transition Model (ETM) [38] are included. In this tool, detailed specifications of the energy system, such as the energy demand, supply and efficiencies, are taken to form scenarios, then the tool can derive electricity prices. This tool is used in two ways. The first way is adopting the ii3050 scenarios incorporated in the model, which take complex predictions on the developments of the energy infrastructure under Dutch 2050 energy scenarios [45]. The other way is scaling up the supply and demand according to the scenario values. And other parameters remain in 2019 values. This is to avoid further assumptions about the context. To clarify, the generative cWGAN-GP and mGaus models project 100 daily series for each day in the 2050, while the ETM gives one.

To have an overview on the price projections, general characteristics of the generated DAM price series, the annual mean and the mean daily standard deviation, can be found in Table 5.3 and 5.4. Let us firstly look at the annual means. Comparing the two ETM projections, annual means in the ii3050 RM and NM scenarios are significantly lower, which implies that the infrastructural developments would lead to further decrease when the

responsibility to meet the climate targets is appointed to the country or the regions. Interestingly, the opposite can be seen in the EC scenario, where sharing the responsibility with other European countries seemingly abates the decrease in the price level. Without such information on the energy infrastructure, the projections made by the cWGAN-GP model are not far from those by ETM direct scaling, with absolute differences less than 10 €. In the cWGAN-GP projections, the annual mean in the EC scenario, where gas is still the most frequent final generating technology, is saliently lower than the other scenarios. This is also the case in the mGaus projections. But these projections have higher mean values than projections by other methods. Except for the mGaus model, all the presented models agree that the annual mean DAM price will more or less decline from the 2019 level.

Table 5.3: Annual mean of the DAM price projected under Dutch 2050 energy scenarios.

Scenario	**Annual mean of the DAM price (€)**				
	ETM direct scaling	ETM ii3050	cWGAN-GP	mGaus	NL 2019
RM	30.35	14.76	38.38	49.92	41.22
NM	29.41	13.75	38.09	49.79	
EC	31.89	36.26	30.77	37.62	
IM	35.69	35.42	41.11	51.06	

Table 5.4: Mean daily standard deviation of the DAM price projected under Dutch 2050 energy scenarios.

Scenario	**Mean daily standard deviation of the DAM price (€)**				
	ETM direct scaling	ETM ii3050	cWGAN-GP	mGaus	NL 2019
RM	10.43	14.52	6.15	7.21	7.50
NM	10.58	13.51	6.16	7.19	
EC	8.42	21.75	4.75	6.49	
IM	8.05	18.86	5.97	7.19	

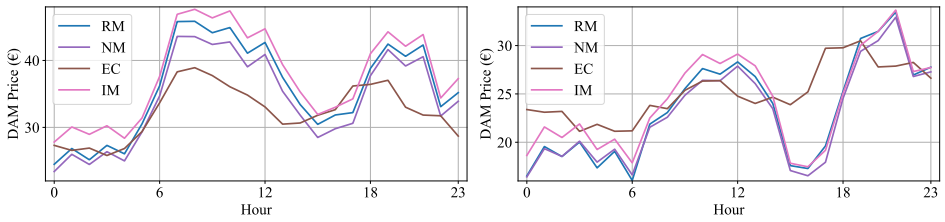
Regarding the mean daily standard deviation in Table 5.4, the ETM ii3050 projections are again eye-catching. These values in the RM, NM and IM scenarios are doubled from the 2019 level, and tripled in the EC scenario. In the same scenario, however, the cWGAN-GP and the mGaus models project lower mean daily standard deviation. Overall, considering developments in the energy infrastructure, the DAM price in ETM ii3050 projections are much more volatile than those by ETM direct scaling. A dissidence appears that ETMs expect enlarged price volatility while the cWGAN-GP and the mGaus models foretell otherwise.

To visualise the peculiarity of projections under the EC scenario, Figure 5.4 shows hourly mean prices on two sample days projected by the cWGAN-GP model. Both the price level and the hourly variation are quite different on the two days. Though the EC daily mean price is not always lower than the other scenarios, the price volatility seems considerably dampened, resulting in outstanding mean price shapes. The peculiar annual means and mean daily standard deviations projected by the cWGAN-GP and the mGaus models under the EC scenario seem to be caused from using the typical Dutch

natural gas price as the fuel cost of the final generating technology. To substantiate this, additional projections are done under a fabricated scenario with the same EC conditions, except that the typical Dutch gas price is substituted with the hydrogen price estimate. Let us name this scenario EC hydrogen. General characteristics of the DAM price in these projections are listed in Table 5.5. Comparing with corresponding values under other scenarios in Table 5.3 and 5.4, it can be concluded that the peculiar values are indeed a consequence from the bigger portion of electricity generation from natural gas.

Table 5.5: General characteristics of the DAM price projected under the fabricated EC hydrogen scenario.

General characteristics	cWGAN-GP	mGaus
Annual mean (€)	40.46	50.75
Mean daily standard deviation (€)	6.01	7.20
AE score ((€))	8.75	8.95



(a) Sample day 1.

(b) Sample day 2.

Figure 5.4: Hourly mean prices of the RM, NM, EC and IM scenarios on two sample days generated by the cWGAN-GP model.

Furthermore, the projections are evaluated with the AE tester to investigate the quality in hourly inter-dependencies. As in Table 5.6 and 5.2, the mGaus model performs consistently with reproducing the test set, which can be owing to the independent sampling from the hourly means and variances, suggesting that the means and variances projected by the mGaus model are not worse than those in testing. This is not the case with the cWGAN-GP model, with performance close to the mGaus model under the RM, NM and IM scenarios, which might results from the fact that the projections are at the boundary of the feature region, where data points are scarce in the training data. It is proven by the AE scores of the projections under the EC hydrogen scenario in Table 5.5, that when the gas price is substituted by the hydrogen price estimate, the AE score rises from 5.90 to 8.75, a similar level with the other scenarios. ETM projections are regarded less reliable by the AE tester, with much higher AE scores. This hints that with the altered supply and demand, as well as influence from infrastructural developments, the ETMs forecast that DAM price series configuration under Dutch 2050 energy scenarios will be intrinsically different from those in the NL, ES, DK1 and DK2 markets from 2015 to 2019.

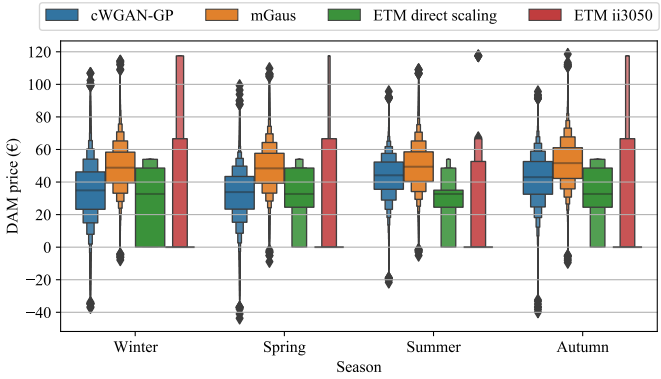
Besides the general characteristics, it is also valuable to dive into the price seasonalities and examine the temporal distribution of the DAM price series projections. For this purpose, the enhanced box-plot introduced in Section 2 again comes into use. Taking the NM scenario as an example, Figure 5.5 demonstrates distribution of projected

Table 5.6: AE score of the DAM price projected under Dutch 2050 energy scenarios.

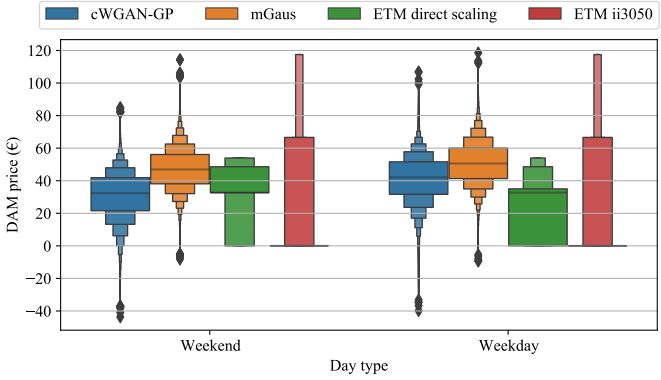
Scenario	AE score of the DAM price (€)			
	ETM direct scaling	ETM ii3050	cWGAN-GP	mGaus
RM	15.43	22.95	8.86	8.96
NM	15.12	21.26	8.90	8.95
EC	11.34	33.03	5.90	8.82
IM	9.75	28.40	8.70	9.00

DAM prices on temporal conditions. For clearer illustration, the days of week are classified into weekends and weekdays, and the months of year are categorised into seasons. DAM prices projected by the cWGAN-GP model have higher means in summer and on weekdays, which are similar with the combined NL, ES, DK1 and DK2 data set plotted earlier in Figures 2.4a and 2.4b. And it is surprising that distribution of the mGaus price does not vary much with the temporal conditions. The ETM prices are discretely distributed, embodied in the lengthy boxes, with a price floor at 0 €. Furthermore, majority of the ETM ii3050 prices are zero, judging from the zero medians shown in the Figure.

Finally, to have a direct assessment, DAM price series projections on the third and the fourth Tuesday in February 2050 are shown in Figure 5.6. With the national mean temperature rising from 3.15 to 5.08 Celsius and other conditions remaining the same, the cWGAN-GP and the mGaus models produces similar price series, varying within smooth boundaries. Hourly variations of vertical position (price level) and the width (price volatility) of the colored areas are recognisable in the figure. Nevertheless, the ETM series are distinguishable from the colored areas, with prices going out of the 95% areas of both cWGAN-GP and mGaus projections on many hours. An explanation to such disparity is that the merit-order mechanism of electricity supply in the ETM assumes a fixed price for each electricity generation technology [128]. Hence the ETM prices jump among certain values, as in the figures. Moreover, in these two samples, the non-zero ETM ii3050 prices are often higher than most cWGAN-GP and mGaus prices, meaning that the marginal generation cost of the final generating technologies assumed in the ETM can be regarded implausible, whereas the non-zero ETM direct scaling prices fit much better.

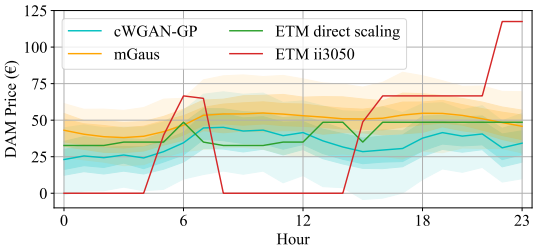


(a) Distribution of projected DAM prices on the seasons under the NM scenario.

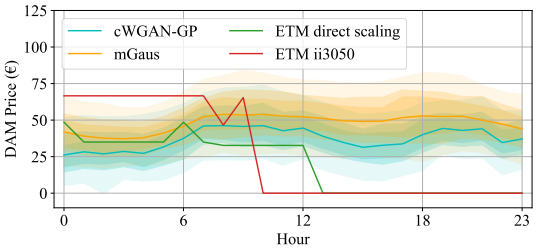


(b) Distribution of projected DAM prices on the weekends and weekdays under the NM scenario.

Figure 5.5: Enhanced box-plots showing the temporal distribution of projected DAM prices. The biggest box in each category includes 50% of the data around the median.



(a) DAM price series projected for the third Tuesday in February 2050. Daily national mean temperature is 3.15 Celsius.



(b) DAM price series projected for the fourth Tuesday in February 2050. Daily national mean temperature is 5.08 Celsius.

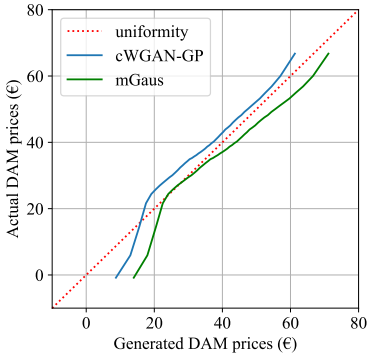
Figure 5.6: DAM price series projection under the NM scenario on the third and the fourth Tuesday in February 2050. Projections by the cWGAN-GP and the mGaus models are plotted in KDE areas, from dark to light including 50%, 70% and 95% of the data points around the median.

5.4. PROJECTION UNDER GERMAN 2019 CONDITIONS

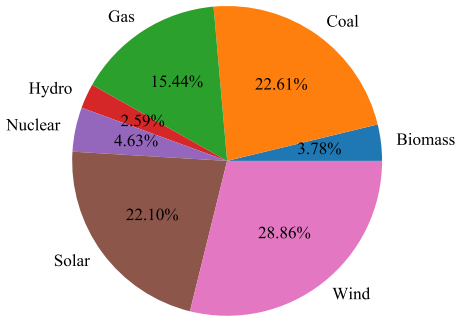
To investigate how effective the models are in projecting DAM price series of a foreign market, German 2019 conditions are acquired. Judging by the relationship between the electricity load and the renewable power generation, German DAM in 2019 (DE 2019) lies between the Dutch DAM in 2019 and Dutch 2050 energy scenarios. For instance, the RSF of the German DAM in 2019 is 0.34, much lower than NL 2019 while higher than those of the 2050 energy scenarios, as in Table 2.4. Therefore, DE 2019 could be a reference of the transitional states of the Dutch DAM. With the German 2019 conditions, each of the cWGAN-GP and mGaus models project the annual DAM price series for 100 times.

Distributions of price values projected by the cWGAN-GP and the mGaus models are quite different from the actual prices, as shown in Figure 5.7a. Quantiles of the projected prices are particularly misaligned in the lower value region, with the maximum discrepancy more than 10 €. The misalignment also embodies in the high K-S distance in Table 5.7. Interestingly, the cWGAN-GP and mGaus quantiles have similar shapes, which implies that the relationship between the DAM price and the conditions in the German market is different from that in the data set of NL, ES, DK1 and DK2 markets. This can be seen in the generation capacity mix in Figure 5.7b. Compared with the Spanish market that has similar capacity share of nuclear energy, the German market has a much larger portion of renewable energy and less gas power. Therefore, facing the severer competition with renewables, the nuclear plants in Germany would have to accept negative DAM

price to maintain in operation more often. Whereas in the Dutch and Danish markets, there is neither such large amount of nuclear energy nor such market squeeze to nuclear plants from the renewables.



(a) Q-Q plot of DAM price projections.



(b) Generation capacity mix of the German DAM in 2019.

Figure 5.7: Q-Q plot of DAM prices projected under German 2019 conditions and generation capacity mix of the German DAM in 2019. RSF normalisation is adopted. Split random state = 0.

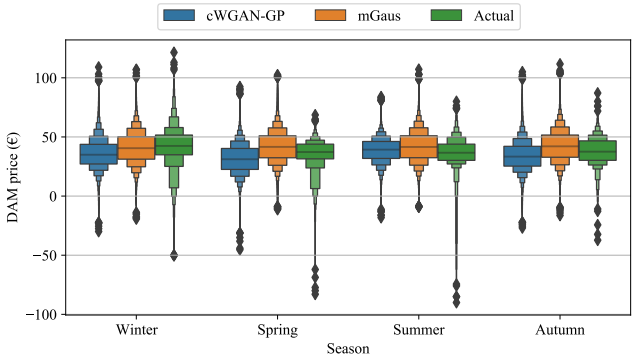
According to the AE scores in Table 5.7, the German 2019 DAM price series projected by the cWGAN-GP model are similar to the actual prices on the hourly inter-dependency. And compared with the AE score of the cWGAN-GP model on the test set in Table 5.2, the AE scores of the cWGAN-GP model on German 2019 conditions and of actual German 2019 price series are lower, suggesting that the German 2019 conditions are close to the region in the training set where the cWGAN-GP model has learned the hourly inter-dependencies well. And the mGaus model is typically graded a higher AE score for the independent sampling of the DAM price for each hour.

Table 5.7: Metric scores and characteristics of the DAM price projected under German 2019 conditions. RSF normalisation is adopted. Split random state = 0.

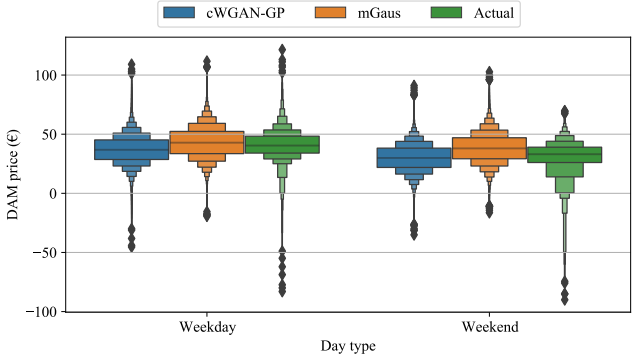
Metrics/characteristics	cWGAN-GP	mGaus	Actual prices
K-S distance	0.144	0.117	0.000
AE score (€)	6.61	9.03	6.23
Annual mean (€)	37.73	34.98	41.71
Mean daily standard deviation (€)	5.82	7.22	8.77

From the last row in Table 5.7, it is shown that the cWGAN-GP prices have a closer mean to the actual German 2019 prices than the mGaus prices. But the cWGAN-GP prices come with a lower mean daily standard deviation. To further examine the price distribution, enhance box-plots of the DAM price under different temporal conditions are shown in Figure 5.8. Across the year, levels of the actual prices are higher in the winter and the autumn, and there are many actual prices below -50 € in the spring and the summer. In the cWGAN-GP projections, the price level in summer is the highest, while all the prices are above -50 €. Similar to the projections under Dutch 2050 energy sce-

narios in Figure 5.5a, the mGaus model still seems insensitive to the month of the year. In terms of the day of the week, all groups of the DAM prices have higher price levels in weekday than in weekend. Again, the case of multiple outliers below -50 € in actual prices is absent in cWGAN-GP and mGaus prices.



(a) Distribution of projected DAM prices per season.



(b) Distribution of projected DAM prices per day type.

Figure 5.8: Enhanced box-plots showing the temporal distribution of DAM prices projected under German 2019 conditions. The biggest box in each category includes 50% of the data around the median.

5.5. GENERIC PROFIT POTENTIAL OF ENERGY STORAGE

To demonstrate how the DAM price projections can be taken advantage of, an energy storage system is designed to earn profit by energy trading. The storage system has a volume of 4 MWh, and power rating of 1 MW. Minimum input/output power is 0 MW. The storage system can ramp up/down instantly, with 100% power efficiency. To have an independent analysis on each daily series, the initial storage state is set at 0 MWh. Lastly, it is assumed that the storage volume and power rating is small enough to the market that the DAM price is inelastic to its bids. By trading on the DAM, the storage operator buy/sell energy on each of the 24 hours. In summary, the daily profit potential of the

storage system can be calculated by solving the linear optimization problem below.

$$\max_{t} \quad \text{Potential daily profit} = \sum_{i=1}^{24} -P_i t_i, \tag{5.1}$$

subject to:

$$0 \leq t_1 \leq 1, \tag{5.2}$$

$$-1 \leq t_i \leq 1, \quad \forall i = 2, ..., 24, \tag{5.3}$$

$$0 \leq \sum_{1}^{i} t_i \leq 4, \quad \forall i = 2, ..., 24, \tag{5.4}$$

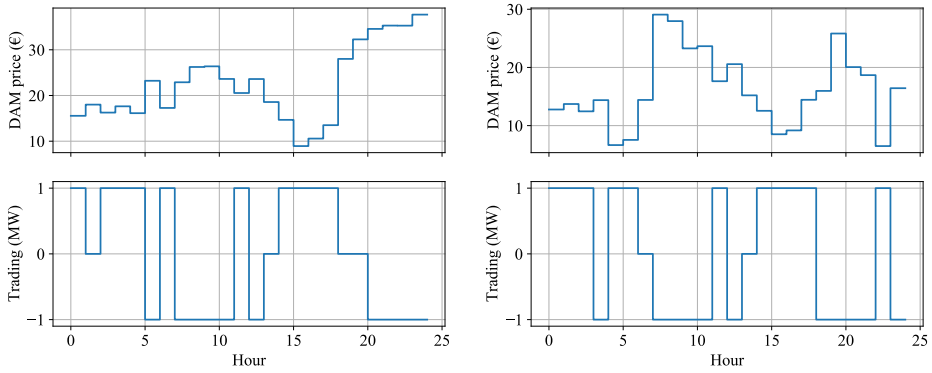
where P_i is the projected DAM price at hour i, and t_i is the energy transaction of that hour, positive for buying while negative for selling. It should be noted that, under the assumption of inelastic price, the ratio between the volume and the power rating is more important than the actual volume and power rating. Due to the linear optimization problem, the potential daily profit is proportional to the volume or the power rating, as long as the ratio remains the same. For example, the potential daily profit will be one-tenth if the the volume is 0.4 MWh and the power rating is 0.1 MW.

Regarding the effect of Volume-Power rating (V-P) ratio, mean potential daily profit of systems with different volumes are acquired in Table 5.8. The cWGAN-GP projections are used. The mean potential daily profits under the RM and NM scenarios are quite similar, which are slightly higher than that under the IM scenario. The EC scenario yields the least mean potential profit, which is in accordance with the result of mean daily standard deviation in Table 5.4. As expected, larger storage volumes lead to increases in the mean potential daily profit, especially under the RM, NM and IM scenarios. However, limits exist in these growths. The mean potential daily profits only increase less than 1 € with the volume enlarging from 8 to 12 MWh. It is foreseeable that these values would be close to those with an unlimited storage volume, or larger than 24 MWh in this case. Figure 5.9 shows two examples of the optimized tradings given by solving the problem in Eq. (5.1) to Eq. (5.4). It is clearly visible that the optimized storage system always tries to buy at low and sell at high, considering both the long and short-term price variation. As the tradings are optimised on the 24 hours together, the tradings are all at full capacity to obtain maximum profits.

Table 5.8: Mean potential daily profit by different storage volumes. The cWGAN-GP projections are used in this analysis.

Scenario	**Mean potential daily profit (€)**		
Volume (MWh)	4	8	12
RM	98.14	115.51	115.91
NM	98.52	115.91	116.32
EC	69.77	83.31	83.53
IM	95.95	112.75	113.08

The mean potential daily profits using projections from the selected methods can be found in Table 5.9. Looking at each method separately, the mean potential daily profits follows the mean daily standard deviation under the scenarios in Table 5.4. For instance,



(a) Projection and trading sample 1.

(b) Projection and trading sample 2.

Figure 5.9: Optimised tradings with corresponding DAM price series on two sample days. The DAM price samples are from the cWGAN-GP projections under the NM scenario.

like the mean daily standard deviation, mean potential daily profit of the ETM ii3050 series is the lowest under the NM scenario and the highest under the EC scenario. When comparing across the projection methods, high standard deviation does not necessarily lead to high daily profit. For instance, under the NM scenario, mean daily standard deviation of the ETM ii3050 series is about doubled of the cWGAN-GP series. But the latter yields a higher mean potential daily profit. This can be explained by the fact that the ETM ii3050 series contains many consecutive zero prices, which cannot be exploited for profit. Even though the standard deviation of the ETM ii3050 series is higher, the more smoothly and constantly changing cWGAN-GP series is shown to be more beneficial to storage systems.

Table 5.9: Mean potential daily profit by different projection methods.

Scenario	**Mean potential daily profit (€)**				
Method	ETM direct scaling	ETM ii3050	cWGAN-GP	mGaus	NL 2019
RM	102.41	100.79	98.14	101.01	108.50
NM	97.08	94.17	98.52	101.31	
EC	79.06	192.04	69.77	91.99	
IM	79.23	171.38	95.95	100.95	

With the NL 2019 level as a reference, only the ETM ii3050 model projects boosted mean potential daily profits under the EC and IM scenarios, while it is widely predicted to be suppressed in other results. According to the cWGAN-GP and mGaus results, the suppression would be exacerbated under the EC scenario, in which natural gas is still the most frequent final generating technology. Finally, without information on the energy infrastructure, mean potential daily profits of the cWGAN-GP and mGaus series, under the RM, NM and IM scenarios, are quite similar, which is not the case for the ETM models.

5.6. DISCUSSION FROM PRICE PROJECTION RESULTS

Having an overview on this study, we have put forward Long-term Electricity Price Projection (LEPP), an unattended study field, to investigate the effect of different energy transition pathways on the electricity spot market prices. Indirectly related to the industrial competitiveness while directly linked to the social welfare and capital investment incentives in the energy industry, such projections of electricity spot market prices is of great values to every major stakeholder in the energy transition.

With that said, the focus of LEPP in this study is not on prophesying precise price values, but rather on exposing the relationship between the price and the scenarios, as well as exploring and validating potential methodologies for this purpose. Hence the 2050 scenarios are simply seen as constraints applied to this inference problem, where part of a multidimensional variable is known and used to estimate distributions of the unknown part. For instance, to obtain price projections on any condition state between 2019 level to that in 2050 scenarios, conditions can be substituted with the desired state. And for this reason, no assumption on the transition process is necessary in this study.

This study has projected Dutch DAM prices under unprecedented conditions including an order of magnitude increase in demand and prominent ratios of generation and capacity from renewables. This was previously viable only through sophisticated physical and procedural modelling like ABM methods. Whereas in this study, an unsupervised generative framework cWGAN-GP was applied to train deep neural networks, without building any price or electricity mechanisms. Such accomplishment was enabled by normalisation of energy systems with proposed system sustainability. These indicators measure the hourly renewable power generation and electricity load in relative scales, with which data from multiple European countries (NL, DK and ES) are taken as homogeneous. For then on, the conditions in Dutch 2050 energy scenarios are no longer unprecedented in the data set, as the high renewable generation to load ratio and possible high temperature are covered in the combined data set. Using the additional data in training the models, they able to make projections under the scenarios.

The projections are a result from the model's capability of generalising the relationship between DAM prices and conditions. This is because even though the once unprecedented conditions are now covered, they are not fully covered by any sample year in the data set. The model still needs to extrapolate the effect of the scenario-specified conditions on month of year, day of week, national mean temperature and fuel price of the most frequent final generating technology. Market state indicator values in 2050 are set as the annual mean calculated from the energy portfolios in the scenarios. This is because no information on the hourly load and renewable power generation can be extracted from the energy scenarios, and that they are expected to be too different in 2050 to be sampled from the data set reasonably.

Nevertheless, beyond this gap are the potentials for the methods in this study to be used as a model of the DAM, as the models are sensitive to the hourly market state. To give an example, Figure 5.10 shows the DAM price series projected by the cWGAN-GP model under the NM scenario on the third Tuesday in February 2050 with and without the RSF at 4 and 16 hours changed to 0.92 from 0.014. As expected, mean prices at the two hours rise. The neighbouring hours are also affected, with mean prices from 0 to 5 and from 13 to 17 get significantly higher. It can be seen that aside from the two hours,

the price-raising effect is more on the hours previous to the change in RSF. A possible explanation is that the load and renewable power generation forecasts are accessible to the Program Responsible Parties (PRP) when submitting the bids, and therefore the consumers tend to compensate the high cost by buying more electricity in advance, hence raising the prices up in the ex-ante hours. With the advancements in the Information and Communications (ICT), one potential application is creating digital twins that can virtually simulate the functions and responses of energy systems for decision support [129]. Owing to this sensitivity to the hourly market state, methods in this study might contribute in the digital twin by building stochastic black-box models of DAM to be used alone or integrated with other white box models that contain topology, components and parameters of the energy system.

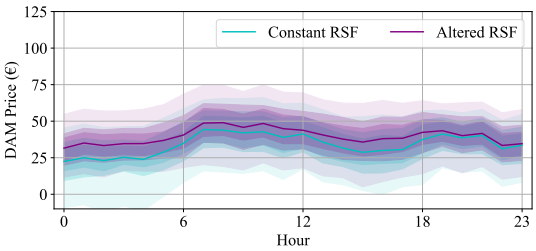


Figure 5.10: DAM price series projection by the cWGAN-GP model under the NM scenario on the third Tuesday in February 2050, plotted in KDE areas, from dark to light including 50%, 70% and 95% of the data points around the median. Constant RSF series are generated with RSF constant at 0.014, the same with those in Figure 5.6a. Altered RSF denotes that the RSF is set to 0.92 at 4 and 16 hours.

In this study, the cWGAN-GP model is compared with CVAE and mGaus models, as well as exogenous ETM models. It clearly outperformed the CVAE models in projection quality. The CVAE model produced unrestrained prices with unreasonable variances. This could be due to its framework structure that emphasises on the proximity to the training sample and predetermined variance level (default at 1 after standardisation). Besides, models trained with such framework might gives projection values in a fitted distribution, but lacking in dimensional inter-dependencies, as demonstrated in this study. The mGaus model projects conservatively around the most approximate training sample. This is no surprise as the mean of such multivariate Gaussian distribution is calculated at the highest possible value. Thus the model is incapable of adapting to the unseen conditions. For instance, as illustrated in this study, under the NM scenario, the mGaus model seems indifferent to temporal conditions, the month of year and the day of week. When compared with the ETM results, The cWGAN-GP model produces series in smoother boundaries and more continuous distributions. And it is regarded more reliable in hourly price inter-dependencies than the ETM models by the AE tester. This shall be attributed to the fact that the cWGAN-GP model learns from the real price series, and ETM gives electricity price via a merit order mechanism, where a certain fixed marginal cost is assumed for each generating technology. Moreover, the non-zero ETM ii3050 prices are often much higher than most cWGAN-GP prices. This could be due to the different assumptions in this study that, variation of Dutch non-household natural

gas price is not considered, and a published estimate of the hydrogen price close to the highest natural gas price in the data set is adopted. Where as in the ETM ii3050 models, the natural gas price in 2050 is set at a level much higher than the historical prices and hydrogen price is taken at a similar value to natural gas.

There are two identifiable limitations of this study. The first is the missing storage in the electricity production/capacity mix under 2050 energy scenarios. From relevant studies, it can be deduced that the growth in energy storage would lead to decreases in the electricity price level and volatility. As the growth in intermittent solar and wind power continues, storage has been regarded an indispensable part for smoothing the power output and reducing curtailment [130, 131]. Hence, even under the same energy portfolio, more renewable energy and less traditional power would be consumed, lowering the price level. When an energy storage system is used alone for arbitrage, a study has shown that depending on its volume and power rating, peaks and troughs of the electricity price can be smoothed out [132]. In addition, it has been shown that with storage systems smoothing the demand, profitability might be enlarged for power generation technologies that are less flexible but more cost-effective, like nuclear power, which would again suppress the price level [133]. Nevertheless, this model is actually compatible with energy storage if its actions are pre-processed into generation and demand. If an energy storage system is trading in the DAM, when the energy storage is selling, and it is not setting the price, then it can be regarded as a renewable power generation. When it is buying from the DAM, it can be viewed as a load. Hence the energy storage can be integrated in the market state indicators, without touching the model. And secondly, the pan-European market integration and cross-boarder power flow are not taken into consideration. As the market integration has not yet happened and cross-boarder power flow is relatively small both now and in the energy scenarios, it is impossible to consider it without proper physical and policy modelling.

Regarding future studies, a number of suggestions can be made. First of all, the potential of cWGAN-GP model has not yet been reached. Judging from the similar performances of the main model in the study and one with much simpler network dimension, much improvements could be made if more computational resources are put in. Another suggestion can be given on differentiating the scenarios for the cWGAN-GP model and other subsequent models. In the price projection results, price characteristics of the projected DAM series under the RM, NM and IM scenarios are similar, suggesting that the model is unable to project distinctly on the alike energy portfolios with the features considered in this study. Given that the supply and demand volumes are disregarded in the market state normalisation, and that there are various assumptions specific to each scenario, more efforts on gathering features that would better differentiate the scenarios can be made. Last but not least, aside from projection to 2050, it would also be valuable to have a look at the ongoing COVID-19 pandemic. The reason for only including data no later than 2019 is to avoid the COVID-19 pandemic. Thinking reversely, it is possible to investigate whether the relationship between the DAM price and conditions of the energy system, as well as hourly DAM price inter-dependencies, has been affected by the pandemic by using the evaluation methods proposed in this study to compare DAM price series projected with conditions in the pandemic years against real prices, which might unveil deeper societal changes, e.g. the fluctuation of consumers' willingness to

pay for electricity.

6

Conclusion

In conclusion, this study projected Dutch Day-Ahead Market (DAM) price series under 2050 energy scenarios using a neural network model trained in the conditional Wasserstein Generative Adversarial Nets with Gradient Penalty (cWGAN-GP) framework. According to the price projection results, due to the high estimation of hydrogen cost, the price-lowering effect brought by the renewable penetration will be counteracted by adopting hydrogen power as the final generating technology, leading to similar DAM price levels and daily price standard deviations in 2050 to those in 2019.

Compared with a basic statistical series generation method, sampling from multivariate Gaussian distributions, the cWGAN-GP model did not lack much in the metric of price value distribution, while actually performing better in the sensitivity to temporal conditions, which are day of the week and month of the year. And in comparison with another generative neural network framework, Conditional Variational Auto-Encoder (CVAE), the cWGAN-GP model was more capable of retaining the price variance in a reasonable range. Although the cWGAN-GP projections differed with the Energy Transition Model (ETM) that takes developments of the energy infrastructure into account, the cWGAN-GP model could generate more realistic DAM price series that variate continuously and sometimes go below zero. To further elucidate the conclusions of this study, the following paragraphs will review how this study has answered the five sub-level research questions proposed in Chapter 1.

What factors should be considered as features for price series projection?

From the DAM clearing mechanism and the result of a frequency analysis of the DAM price series, it was determined to project the DAM price of 24 hours each day. A literature survey was conducted, and day of the week, month of the year, total hourly load forecast, national daily mean temperature, fuel cost of the most frequent final generating technology, hourly renewable power generation forecast and total installed capacity of renewables were preliminarily selected as candidate features. An exploratory data analysis illustrates the linear relationship between the candidate features and the DAM price, where the positive correlation between the total hourly load forecast and the DAM price

was identified. Subsequently, feature importance is analysed via Principal Component Analysis (PCA) and Random Forest (RF) regression, but only confirmed the necessity to include information from all the candidate features together with a causality validation.

Three of the candidate features, total hourly load forecast, hourly renewable generation forecast and total installed capacity of renewables, were used to form market state indicators proposed in this study to normalise the DAM. Four normalisation solutions were presented, namely Renewable Scarcity Factor (RSF), Renewable-Load Ratio (RLR), Renewable Load Satisfaction Ratio (RLSR) and Normalising with total installed generation CAPacity (NCAP). The RSF is the difference between the total hourly load forecast and hourly renewable generation forecast, divided by the total installed capacity of renewables. The RLR is the hourly renewable generation forecast divided by the total hourly load forecast. The RLSR was derived by capping the RLR at 1. And the NCAP solution contains all the three candidate features normalised with the total installed generation capacity.

The market state indicators were proposed because the Dutch data alone are not enough to make projections under 2050 scenarios. Particularly, the renewable penetrations in the scenarios are unprecedented in the Dutch data. To this end, data from Spanish (ES) and Danish (DK1 and DK2) markets were included. And by using the market state indicators instead of the three candidate features, data from all the chosen markets became adjacent and comparable in scale, regardless of their contrasting market sizes.

Which method is most suitable for this task, and how can it be effectively applied?

This study focused on the data-driven methods, which do not build models of components and their functions in a system but try to extract the intrinsic relationship between the target variable and the selected features from data. In view of this, as mentioned earlier, generative methods, including sampling from the multivariate Gaussian distribution, the cWGAN-GP and the CVAE, were adopted in this study. For the multivariate Gaussian distribution, the DAM price and the features were aligned in a matrix. Co-variance between each pair of the dimension was calculated to acquire the mean and variance of the price on each hour of the day under certain conditions. To apply the cWGAN-GP and the CVAE methods that are based on neural networks, the data set was randomly split into a training set and a test set as the standard machine learning approach. And dimensions of the neural network and the choice of parameter optimization algorithm were optimized. In this study, each neural network model was trained for 30 minutes on the author's computing platform. And among the three methods, as mentioned in the beginning, the cWGAN-GP was outstanding for its balance between generated price distribution and hourly inter-dependencies.

Do data from other bidding zones help in projecting future Dutch DAM price series?

Data from other bidding zones were shown to be indispensable in projecting future Dutch DAM price series. This is reflected in two results. One is that when only the Dutch data was used to train the cWGAN-GP model, significant deficiencies in both generated price distribution and price hourly inter-dependencies could be observed in the model performance. The other is that with the data from other bidding zones, the cWGAN-GP

model projected German 2019 DAM prices series with an absolute difference less than 5 € per quantile when the price is above 20 € and price hourly inter-dependencies slightly better than projecting on the test set, though renewable penetration of the German market is much intenser than that in the Netherlands in 2019.

What methods provide reliable validation of the projected price series?

Inspired by the common validation approaches of Generative Adversarial Nets (GAN) applications, this study regarded price value distribution and hourly price inter-dependencies as two aspects of series quality. The price value distribution was generally evaluated with the Kolmogorov-Smirnov distance, the maximum quantile difference between the cumulative distribution curve of the generated and real samples. When a more intuitive demonstration is desired, quantile-quantile plots come in hand to illustrate the distributions in detail. In terms of the price hourly inter-dependency, an Auto-Encoder (AE) tester model that is often used to detect anomalies in data series was trained to regenerate the price series. And the price hourly inter-dependency is assessed with the mean Euclidean distance between the projected price series and their counterparts in the AE regeneration.

What would be valuable applications of the projector model?

It can be envisaged that the model capable of projecting the Dutch 2050 DAM price series would be valuable for various decision support and market analyses. Among the possible applications, this study explored using the projector model to project the German 2019 DAM price series and to calculate the generic profit potential of energy storage in 2050. Despite the more prominent presence of nuclear energy and relatively high renewable penetration in the German market, DAM price series projected by the cWGAN-GP model were not far off in the price value distribution and good in hourly price inter-dependencies. And the generic profit potential of energy storage was calculated by solving a linear optimization problem on the projected price series. It was demonstrated that high daily standard deviation of the DAM price series do not always result in high generic profit potential, for instance, on the ETM series where the DAM price is continuously zero from time to time.

Overall, the DAM price series under Dutch 2050 energy scenarios can be projected by a neural network model trained with the cWGAN-GP framework. The reliability of the projection is substantiated by reproducing the test set and demonstrated with projecting on the German 2019 conditions. Such a model has a lot of potentials. To name a few, it can be a stochastic digital twin of the DAM or integrated into a co-simulation of the energy system for its sensitivity to the hourly variable market state, which is not thoroughly investigated in this study. And more computational resources and efforts in feature engineering may help the model differentiate better the energy scenarios. Lastly, by comparing the actual price series and projections, the generative model might provide insights into the societal changes, for instance, under the COVID-19 pandemic. With these said, nevertheless, it is always welcome to implement other frameworks to pursue better LEPP results.

REFERENCES

[1] A. Atkeson and P. J. Kehoe, "The transition to a new economy after the second industrial revolution," National Bureau of Economic Research, Working Paper 8676, December 2001. [Online]. Available: http://www.nber.org/papers/w8676

[2] IEA, "Key world energy statistics 2020," 2020. [Online]. Available: https://www.iea.org/reports/key-world-energy-statistics-2020

[3] R. Zhang and S. Fujimori, "The role of transport electrification in global climate change mitigation scenarios," *Environmental Research Letters*, vol. 15, no. 3, p. 034019, feb 2020.

[4] IEA, "Global ev outlook 2020," 2020. [Online]. Available: https://www.iea.org/reports/global-ev-outlook-2020

[5] IPCC, *Climate Change 2014: Synthesis Report*, Geneva, Switzerland, 2014, p. 151. [Online]. Available: https://archive.ipcc.ch/report/ar5/syr/

[6] European Commission, "European climate law," Mar 2020. [Online]. Available: https://ec.europa.eu/clima/policies/eu-climate-action/law_en

[7] H. Ritchie and M. Roser, "Renewable energy," *Our World in Data*, 2020, https://ourworldindata.org/renewable-energy.

[8] J. Serrano González and C. Álvarez Alonso, "Industrial electricity prices in spain: A discussion in the context of the european internal energy market," *Energy Policy*, vol. 148, p. 111930, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301421520306418

[9] G. Mordue, "Electricity prices and industrial competitiveness: A case study of final assembly automobile manufacturing in the united states and canada," *Energy Policy*, vol. 111, pp. 32–40, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301421517305645

[10] H. Ai, S. Xiong, K. Li, and P. Jia, "Electricity price and industrial green productivity: Does the "low-electricity price trap" exist?" *Energy*, vol. 207, p. 118239, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544220313463

[11] P. Cramton, "Lessons from the 2021 texas electricity crisis," *Working paper*, 6 2021. [Online]. Available: http://www.cramton.umd.edu/papers2020-2024/cramton-lessons-from-the-2021-texas-electricity-crisis.pdf

[12] TenneT, "Dutch regulation," 2021. [Online]. Available: https://www.tennet.eu/e-insights/regulation-of-the-electricity-price/dutch-regulation/

[13] F. Tanrisever, K. Derinkuyu, and M. Heeren, "Forecasting electricity infeed for distribution system networks: An analysis of the dutch case," *Energy*, vol. 58, pp. 247–257, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544213004428

[14] European Energy Exchange AG, "Power derivatives market," 2021. [Online]. Available: https://www.eex.com/en/markets/power-derivatives-market

[15] F. Tanrisever, K. Derinkuyu, and G. Jongen, "Organization and functioning of liberalized electricity markets: An overview of the dutch market," *Renewable and Sustainable Energy Reviews*, vol. 51, pp. 1363–1374, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032115006668

[16] P. Sorknæs, S. R. Djørup, H. Lund, and J. Z. Thellufsen, "Quantifying the influence of wind power and photovoltaic on future electricity market prices," *Energy Conversion and Management*, vol. 180, pp. 312–324, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0196890418312512

[17] T. Rintamäki, A. S. Siddiqui, and A. Salo, "Does renewable energy generation decrease the volatility of electricity prices? an analysis of denmark and germany," *Energy Economics*, vol. 62, pp. 270–282, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140988317300063

[18] B. Fanzeres, S. Ahmed, and A. Street, "Robust strategic bidding in auction-based markets," *European Journal of Operational Research*, vol. 272, no. 3, pp. 1158–1172, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0377221718306325

[19] P. Cramton, "Electricity market design," *Oxford Review of Economic Policy*, vol. 33, no. 4, 2017.

[20] TenneT, "Annual market update 2020: electricity market insights," Tech. Rep., 03 2021. [Online]. Available: https://www.tennet.eu/company/publications/technical-publications/

[21] D. W. Bunn, "Modelling prices in competitive electricity markets," 2004.

[22] M. Bucksteeg, S. Spiecker, and C. Weber, "Impact of coordinated capacity mechanisms on the european power market," *The Energy Journal*, 2019. [Online]. Available: http://www.iaee.org/en/publications/ejarticle.aspx?id=3333

[23] W. Huang, N. Zhang, C. Kang, M. Li, and M. Huo, "From demand response to integrated demand response: review and prospect of research and application," *Protection and Control of Modern Power Systems*, vol. 4, pp. 1–13, 2019.

[24] X. Yan and N. A. Chowdhury, "Mid-term electricity market clearing price forecasting: A hybrid lssvm and armax approach," *International Journal of Electrical Power Energy Systems*, vol. 53, pp. 20–26, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061513001658

[25] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169207014001083

[26] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Applied Energy*, vol. 221, pp. 386 – 405, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S030626191830196X

[27] D. Keles, J. Scelle, F. Paraschiv, and W. Fichtner, "Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks," *Applied Energy*, vol. 162, pp. 218–230, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261915012039

[28] K. Lindberg, P. Seljom, H. Madsen, D. Fischer, and M. Korpås, "Long-term electricity load forecasting: Current and future trends," *Utilities Policy*, vol. 58, pp. 102 – 119, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957178719300116

[29] S. Backe, P. C. del Granado, A. Tomasgard, D. Pinel, M. Korpast, and K. B. Lindberg, "Towards zero emission neighbourhoods: Implications for the power system," in *2018 15th International Conference on the European Energy Market (EEM)*, 2018, pp. 1–6.

[30] T. Boßmann and I. Staffell, "The shape of future electricity demand: Exploring load curves in 2050s germany and britain," *Energy*, vol. 90, pp. 1317–1333, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544215008385

[31] I. k. Nti, A. Adekoya, O. Nyarko-Boateng, and M. Teimah, "Electricity load forecasting: a systematic review," *Journal of Electrical Systems and Information Technology*, vol. 7, 09 2020.

[32] CE Delft, "Energy and electricity price scenarios 2020-2023-2030," Tech. Rep., 2017. [Online]. Available: https://cedelft.eu/wp-content/uploads/sites/2/2021/04/CE_Delft_3H58_Energy_and_electricity_price_scenarios_DEF.pdf

[33] C. Perez-Linkenheil, "Update december 2019 – eu energy outlook 2050," Feb 2020. [Online]. Available: https://blog.energybrainpool.com/en/update-december-2019-eu-energy-outlook-2050/

[34] F. Leuthold, H. Weigt, and C. von Hirschhausen, "A large-scale spatial optimization model of the european electricity market," *Networks and Spatial Economics*, vol. 12, pp. 75–107, 03 2012.

[35] C. Fraunholz, E. Kraft, D. Keles, and W. Fichtner, "Advanced price forecasting in agent-based electricity market simulation," *Applied Energy*, vol. 290, p. 116688, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261921002142

[36] C. Macal, "Everything you need to know about agent-based modelling and simulation," *Journal of Simulation*, vol. 10, pp. 144–156, 2016.

[37] E. J. Chappin, L. J. de Vries, J. C. Richstein, P. Bhagwat, K. Iychettira, and S. Khan, "Simulating climate and energy policy with agent-based modelling: The energy modelling laboratory (emlab)," *Environmental Modelling Software*, vol. 96, pp. 421–431, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364815216310301

[38] Quintel, "Create your own energy future," 2021. [Online]. Available: https://energytransitionmodel.com/

[39] Émile Jean Louis Chappin, "Simulating energy transitions," Ph.D. dissertation, Delft University of Technology, Delft, 2011.

[40] B. O'Neill, T. Carter, K. Ebi, P. Harrison, E. Kemp-Benedict, K. Kok, E. Kriegler, B. Preston, K. Riahi, J. Sillmann, B. Ruijven, D. Vuuren, D. Carlisle, C. Conde, J. Fuglestvedt, C. Green, T. Hasegawa, J. Leininger, S. Monteith, and R. Pichs-Madruga, "Achievements and needs for the climate change scenario framework," *Nature Climate Change*, p. 1074–1084, 12 2020.

[41] L. Zahradníčková and E. Vacík, "Scenarios as a strong support for strategic planning," *Procedia Engineering*, vol. 69, pp. 665–669, 2014, 24th DAAAM International Symposium on Intelligent Manufacturing and Automation, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877705814002860

[42] Government of the Netherlands, "Climate act," 2019. [Online]. Available: https://www.government.nl/topics/climate-change/climate-policy

[43] CE Delft, "Energy scenarios for the Netherlands to 2030," Tech. Rep., 2014. [Online]. Available: https://www.cedelft.eu/en/publications/1493/energy-scenarios-for-2030

[44] TNO Energy, "TNO energy visio 2030," Tech. Rep., 2016. [Online]. Available: https://www.tno.nl/media/9604/tno_energyvision2030.pdf

[45] Netbeheer, "Het Energiesysteem van de Toekomst Integrale Infrastructuurverkenning 2030-2050," 2021. [Online]. Available: https://www.netbeheernederland.nl/dossiers/toekomstscenarios-64

[46] Berenschot and Kalavasta, "Klimaatneutrale energiescenario's 2050," Tech. Rep., 2020. [Online]. Available: https://www.berenschot.nl/vast/zoeken/?zoeken_term=klimaatneutrale%20energies

[47] ENTSO-E, "Central collection and publication of electricity generation, transportation and consumption data and information for the pan-european market," 2021. [Online]. Available: https://transparency.entsoe.eu/dashboard/show

[48] epexspot, "Newsroom," 2021. [Online]. Available: https://www.epexspot.com/en/newsroom

[49] G. Wolff and S. Feuerriegel, "Short-term dynamics of day-ahead and intraday electricity prices," *International Journal of Energy Sector Management*, vol. 11 No. 4, pp. pp. 557–573, 2017.

[50] G. P. Herrera, M. Constantino, B. M. Tabak, H. Pistori, J.-J. Su, and A. Naranpanawa, "Long-term forecast of energy commodities price using machine learning," *Energy*, vol. 179, pp. 214–221, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S036054421930708X

[51] H. Sangrody, N. Zhou, S. Tutun, B. Khorramdel, M. Motalleb, and M. Sarailoo, "Long term forecasting using machine learning methods," in *2018 IEEE Power and Energy Conference at Illinois (PECI)*, 2018, pp. 1–5.

[52] A. Borji, "Pros and cons of gan evaluation measures," *Computer Vision and Image Understanding*, vol. 179, pp. 41–65, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1077314218304272

[53] S. Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete fourier transform," *IEEE Transactions on Communication Technology*, vol. 19, no. 5, pp. 628–634, 1971.

[54] The Numpy community, "numpy.fft.fft manual," 2021. [Online]. Available: https://numpy.org/doc/stable/reference/generated/numpy.fft.fft.html

[55] K. Mayer and S. Trück, "Electricity markets around the world," *Journal of Commodity Markets*, vol. 9, pp. 77–100, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405851318300059

[56] Z. Khan, G. Iyer, P. Patel, S. Kim, M. Hejazi, C. Burleyson, and M. Wise, "Impacts of long-term temperature change and variability on electricity investments," *Nature Communications*, vol. 12, no. 1, p. 1643, Mar 2021. [Online]. Available: https://doi.org/10.1038/s41467-021-21785-1

[57] Y. Li, W. Pizer, and L. wu, "Climate change and residential electricity consumption in the yangtze river delta, china," *Proceedings of the National Academy of Sciences*, vol. 116, p. 201804667, 12 2018.

[58] I. Staffell and S. Pfenninger, "The increasing impact of weather on electricity supply and demand," *Energy*, vol. 145, pp. 65–78, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544217320844

[59] Open Power System Data, "A free and open data platform for power system modelling," 2021. [Online]. Available: https://open-power-system-data.org/

[60] C. Zhang, R. Li, H. Shi, and F. Li, "Deep learning for day-ahead electricity price forecasting," *IET Smart Grid*, vol. 3, no. 4, pp. 462–469, 2020.

[61] H. Mosbah and M. El-hawary, "Hourly electricity price forecasting for the next month using multilayer neural network," *Canadian Journal of Electrical and Computer Engineering*, vol. 39, no. 4, pp. 283–291, 2016.

[62] Eurostat, "Natural gas price statistics," 2021. [Online]. Available: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Natural_gas_price_statistics#Natural_gas_prices_for_non-household_consumers

[63] Y. Wang, G. Hug, Z. Liu, and N. Zhang, "Modeling load forecast uncertainty using generative adversarial networks," *Electric Power Systems Research*, vol. 189, p. 106732, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378779620305356

[64] Y. Zhang, Q. Ai, F. Xiao, R. Hao, and T. Lu, "Typical wind power scenario generation for multiple wind farms using conditional improved wasserstein generative adversarial network," *International Journal of Electrical Power Energy Systems*, vol. 114, p. 105388, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061519305010

[65] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, no. null, p. 1157–1182, Mar. 2003.

[66] A. L. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artificial Intelligence*, vol. 97, no. 1, pp. 245–271, 1997, relevance. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0004370297000635

[67] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, no. 1, pp. 273–324, 1997, relevance. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S000437029700043X

[68] J. W. Tukey, "The Future of Data Analysis," *The Annals of Mathematical Statistics*, vol. 33, no. 1, pp. 1 – 67, 1962. [Online]. Available: https://doi.org/10.1214/aoms/1177704711

[69] S. M. Stigler, "Francis Galton's Account of the Invention of Correlation," *Statistical Science*, vol. 4, no. 2, pp. 73 – 79, 1989. [Online]. Available: https://doi.org/10.1214/ss/1177012580

[70] I. S. Abramson, "On Bandwidth Variation in Kernel Estimates-A Square Root Law," *The Annals of Statistics*, vol. 10, no. 4, pp. 1217 – 1223, 1982. [Online]. Available: https://doi.org/10.1214/aos/1176345986

[71] H. Hofmann, H. Wickham, and K. Kafadar, "Letter-value plots: Boxplots for large data," *Journal of Computational and Graphical Statistics*, vol. 26, no. 3, pp. 469–477, 2017. [Online]. Available: https://doi.org/10.1080/10618600.2017.1305277

[72] X. Kong, X. Liu, R. Shi, and K. Y. Lee, "Wind speed prediction using reduced support vector machines with feature selection," *Neurocomputing*, vol. 169, pp. 449–456, 2015, learning for Visual Semantic Understanding in Big Data ESANN 2014 Industrial Data Processing and Analysis. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231215005044

[73] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1, pp. 37–52, 1987, proceedings of the Multivariate Statistical Workshop for Geologists and Geochemists. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0169743987800849

[74] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[75] M. Schonlau and R. Y. Zou, "The random forest algorithm for statistical learning," *The Stata Journal*, vol. 20, no. 1, pp. 3–29, 2020. [Online]. Available: https://doi.org/10.1177/1536867X20909688

[76] scikit-learn, "Scikit-learn, machine learning in python," 2021. [Online]. Available: https://scikit-learn.org/stable/

[77] Eurostat, "Renewable energy statistics: statistics explained," https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Renewable_energy_statistics.

[78] D. Liu, D. Niu, H. Wang, and L. Fan, "Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm," *Renewable Energy*, vol. 62, pp. 592–597, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960148113004138

[79] W. Sun, M. Liu, and Y. Liang, "Wind speed forecasting based on feemd and lssvm optimized by the bat algorithm," *Energies*, vol. 8, no. 7, pp. 6585–6607, 2015. [Online]. Available: https://www.mdpi.com/1996-1073/8/7/6585

[80] Q. Wu and C. Peng, "Wind power grid connected capacity prediction using lssvm optimized by the bat algorithm," *Energies*, vol. 8, no. 12, pp. 14 346–14 360, 2015. [Online]. Available: https://www.mdpi.com/1996-1073/8/12/12428

[81] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969. [Online]. Available: http://www.jstor.org/stable/1912791

[82] statsmodel, "Four tests for granger non causality of 2 time series," https://www.statsmodels.org/stable/generated/statsmodels.tsa.stattools.grangercausalitytests.html?highlight=granger#statsmodels.tsa.stattools.grangercausalitytests.

[83] J. G. MacKinnon, "Critical Values For Cointegration Tests," Economics Department, Queen's University, Working Paper 1227, Jan. 2010. [Online]. Available: https://ideas.repec.org/p/qed/wpaper/1227.html

[84] statsmodel, "Augmented dickey-fuller unit root test," https://www.statsmodels.org/stable/generated/statsmodels.tsa.stattools.adfuller.html.

[85] C. Feng, M. Cui, B.-M. Hodge, and J. Zhang, "A data-driven multi-model methodology with deep feature selection for short-term wind forecasting," *Applied Energy*, vol. 190, pp. 1245–1257, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261917300508

[86] statsmodel, "Test for no-cointegration of a univariate equation," https://www.statsmodels.org/stable/generated/statsmodels.tsa.stattools.coint.html?highlight=coint.

[87] O. M. Kam, S. Noël, H. Ramenah, P. Kasser, and C. Tanougast, "Comparative weibull distribution methods for reliable global solar irradiance assessment in france areas," *Renewable Energy*, vol. 165, pp. 194–210, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960148120317584

[88] K. Azad, M. Rasul, P. Halder, and J. Sutariya, "Assessment of wind energy prospect by weibull distribution for prospective wind sites in australia," *Energy Procedia*, vol. 160, pp. 348–355, 2019, 2nd International Conference on Energy and Power, ICEP2018, 13–15 December 2018, Sydney, Australia. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1876610219312573

[89] S. R. y Cajal, *Histologie du système nerveux de l'homme & des vertébrés: Cervelet, cerveau moyen, rétine, couche optique, corps strié, écorce cérébrale générale & régionale, grand sympathique*. Paris: A. Maloine, 1911, vol. 2.

[90] F. Azevedo, L. Carvalho, L. Grinberg, J. Farfel, R. Ferretti-Rebustini, R. Leite, W. Filho, R. Lent, and S. Herculano-Houzel, "Equal numbers of neuronal and non-neuronal cells make the human brain an isometrically scaled-up primate brain," *The Journal of comparative neurology*, vol. 513, pp. 532–41, 04 2009.

[91] T. C. Südhof and R. C. Malenka, "Understanding synapses: past, present, and future," *Neuron*, vol. 60, no. 3, pp. 469–476, 2008.

[92] W. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biology*, vol. 52, pp. 99–115, 1990.

[93] S. Theodoridis, "Chapter 18 - neural networks and deep learning," in *Machine Learning (Second Edition)*, second edition ed., S. Theodoridis, Ed. Academic Press, 2020, pp. 901–1038. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128188033000301

[94] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.

[95] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 11 1997. [Online]. Available: https://doi.org/10.1162/neco.1997.9.8.1735

[96] F. Rosenblatt, "Principles of neurodynamics. perceptrons and the theory of brain mechanisms," *American Journal of Psychology*, vol. 76, p. 705, 1963.

[97] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," *arXiv preprint arXiv:1505.00853*, 2015.

[98] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th International Conference on Machine Learning*, 2010, pp. 807–814.

[99] A. L. Maas, "Rectifier nonlinearities improve neural network acoustic models," 2013.

[100] A. F. Agarap, "Deep learning using rectified linear units (relu)," *ArXiv*, vol. abs/1803.08375, 2018.

[101] S. Ruder, "An overview of gradient descent optimization algorithms," 2017.

[102] L. Bottou, *Stochastic Gradient Descent Tricks*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 421–436. [Online]. Available: https://doi.org/10.1007/978-3-642-35289-8_25

[103] G. Hinton, "Neural networks for machine learning, lecture 6a: Overview of mini-batch gradient descent," http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf.

[104] P. Verma, V. Tripathi, and B. Pant, "Comparison of different optimizers implemented on the deep learning architectures for covid-19 classification," *Materials Today: Proceedings*, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214785321013316

[105] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2015.

[106] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, ser. NIPS'14. Cambridge, MA, USA: MIT Press, 2014, p. 2672–2680.

[107] F. Wang, Z. Zhang, C. Liu, Y. Yu, S. Pang, N. Duić, M. Shafie-khah, and J. P. Catalão, "Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting," *Energy Conversion and Management*, vol. 181, pp. 443 – 462, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0196890418313268

[108] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein gan," 2017.

[109] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of wasserstein gans," 2017.

[110] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," 2014.

[111] A. Pagnoni, K. Liu, and S. Li, "Conditional variational autoencoder for neural machine translation," *ArXiv*, vol. abs/1812.04405, 2018.

[112] M. Eaton, *Multivariate Statistics: A Vector Space Approach*, ser. Probability and Statistics Series. Wiley, 1983. [Online]. Available: https://books.google.nl/books?id=1CvvAAAAMAAJ

[113] M. Müller, *Dynamic Time Warping*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 69–84. [Online]. Available: https://doi.org/10.1007/978-3-540-74048-3_4

[114] J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey, *Graphical methods for data analysis*. Chapman and Hall/CRC, 2018.

[115] *Kolmogorov–Smirnov Test*. New York, NY: Springer New York, 2008, pp. 283–287. [Online]. Available: https://doi.org/10.1007/978-0-387-32833-1_214

[116] N. V. Smirnov, "Estimate of deviation between empirical distribution functions in two independent samples," *Bulletin Moscow University*, vol. 2, no. 2, pp. 3–16, 1939.

[117] A. Kolmogorov, "Sulla determinazione empirica di una lgge di distribuzione," *Inst. Ital. Attuari, Giorn.*, vol. 4, pp. 83–91, 1933.

[118] C. Wang, S. Tindemans, K. Pan, and P. Palensky, "Detection of false data injection attacks using the autoencoder approach," in *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 2020, pp. 1–6.

[119] C. Bergmeir and J. M. Benítez, "On the use of cross-validation for time series predictor evaluation," *Information Sciences*, vol. 191, pp. 192–213, 2012, data Mining for Software Trustworthiness. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025511006773

[120] Keras, "Keras documentation: Dense layer," n.d. [Online]. Available: https://keras.io/api/layers/core_layers/dense/

[121] S. van Renssen, "The hydrogen solution?" *Nature Climate Change*, vol. 10, p. 799–801, 2020. [Online]. Available: https://www.nature.com/articles/s41558-020-0891-0

[122] S. E. Hosseini and M. A. Wahid, "Hydrogen production from renewable and sustainable energy resources: Promising green energy carrier for clean development," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 850–866, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032115014951

[123] H. Dagdougui, R. Sacile, C. Bersani, and A. Ouammi, "Chapter 2 - hydrogen production and current technologies," in *Hydrogen Infrastructure for Energy Applications*, H. Dagdougui, R. Sacile, C. Bersani, and A. Ouammi, Eds. Academic Press, 2018, pp. 7–21. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128120361000020

[124] C. Kaandorp, N. C. van de Giesen, and E. Abraham, "The water use of heating pathways to 2050: analysis of national and urban energy scenarios," *Environmental Research Letters*, vol. 16, 2021.

[125] M. Kayfeci, A. Keçebaş, and M. Bayat, "Chapter 3 - hydrogen production," in *Solar Hydrogen Production*, F. Calise, M. D. D'Accadia, M. Santarelli, A. Lanzini, and D. Ferrero, Eds. Academic Press, 2019, pp. 45–83. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128148532000035

[126] T. Ayodele and J. Munda, "Potential and economic viability of green hydrogen production by water electrolysis using wind energy resources in south africa," *International Journal of Hydrogen Energy*, vol. 44, no. 33, pp. 17 669–17 687, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360319919319330

[127] B. Sundén, "Chapter 1 - introduction and background," in *Hydrogen, Batteries and Fuel Cells*, B. Sundén, Ed. Academic Press, 2019, pp. 1–13. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128169506000117

[128] Quintel Intelligence, "Merit order in the energy transition model," 2021. [Online]. Available: https://docs.energytransitionmodel.com/main/merit-order/

[129] P. Palensky, M. Cvetkovic, D. Gusain, and A. Joseph, "Digital twins and their use in future power systems [version 1; peer review: 1 approved, 1 approved with reservations]," *Digital Twin*, vol. 1, no. 4, 2021.

[130] X. Dui, G. Zhu, and L. Yao, "Two-stage optimization of battery energy storage capacity to decrease wind power curtailment in grid-connected wind farms," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3296–3305, 2018.

[131] P. Denholm and T. Mai, "Timescales of energy storage needed for reducing renewable energy curtailment," *Renewable Energy*, vol. 130, pp. 388–399, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960148118307316

[132] A. S. A. Awad, J. D. Fuller, T. H. M. EL-Fouly, and M. M. A. Salama, "Impact of energy storage systems on electricity market equilibrium," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 3, pp. 875–885, 2014.

[133] M. Askeland, S. Jaehnert, and M. Korpås, "Equilibrium assessment of storage technologies in a power market with capacity remuneration," *Sustainable Energy Technologies and Assessments*, vol. 31, pp. 228–235, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S221313881730601X